Tatra Mt. Math. Publ. **39** (2008), 125–133

Tatra Mountains
Mathematical Publications

MODIFICATIONS AND EXTENSIONS OF THE METHOD OF VIRTUAL NOISE IN CORRELATED EXPERIMENTAL DESIGN

Andrej Pázman — Juraj Laca

ABSTRACT. We consider a random process (or field) $y(x) = \eta(\theta, x) + \varepsilon(x)$ defined on a set $\mathcal{X}$ and having a covariance function $C(x, z, \beta)$ parametrized by β. The aim is to find an N-point design $\{x_1, \ldots, x_N\} \subset \mathcal{X}$ without replications, and such that θ and β will be estimated (locally) optimally. For that purpose one can use the method of virtual noise, which consists of adding to each $y(x)$ for computational purposes a supplementary independent noise $\varepsilon_\xi(x)$, which is a white noise with variances at each $x \in \mathcal{X}$ depending on a design measure ξ We can approach the optimal design by iterative changes of ξ. We extend the method to the case that also β is estimated, and we present alternative, simpler expressions for the variance of the virtual noise.

1. Introduction

We consider a random process (random field)

$$\big\{y(x); \qquad x \in \mathcal{X}\big\} \tag{1}$$

with mean $E\big[y(x)\big] = \eta(x, \theta)$ and with a covariance function $C(x_1, x_2, \beta)$, which are parametrized by vector parameters $\theta \in \Theta \subset \mathbb{R}^p,\ \beta \in B \subset \mathbb{R}^q$. Since replications of observations in a correlated random process are usually not possible, a design is simply a subset of $\mathcal{X},\ A = \{x_1, \ldots, x_N\}$, with $\#A = N =$ the sample size.

The set $\mathcal{X}$ will be finite (if not, we have to approximate it by a finite net), and the matrix $C(\mathcal{X}, \beta) = \big\{C(x, z, \beta)\big\}_{x,z\in\mathcal{X}}$ is supposed to be nonsingular. Consequently, for every design $A \subset \mathcal{X}$ the matrix $C(A, \beta) = \big\{C(x, z, \beta)\big\}_{x,z\in A}$ is nonsingular as well.

2000 Mathematics Subject Classification: Primary 62K05; Secondary 62M10.
Keywords: spatial statistics, information matrix, optimum design, time series.
Supported by the grant VEGA no. 1/3016/06 and the WTZ project No. 04/2006.

In the case when the model is linear, i.e., $\eta(x,\theta) = f^T(x)\theta$, and the covariance function is known, the design problem has been studied in several papers, (cf. a survey on earlier papers in [4], for some recent results see [7]). The "method of virtual noise" has been presented in the MODA conference 1998, and used algorithmically in [2], [3]. Here we reconsider this method and extend it to the case of unknown β.

The main idea of the method is that instead of (1) we consider, purely for computational purposes, a "virtual random process"

$$\left\{y_\xi(x) = y(x) + \varepsilon_\xi(x); \qquad x \in \mathcal{X}\right\}. \tag{2}$$

The added "virtual noise" $\varepsilon_\xi(x)$ is supposed to be distributed $\mathcal{N}\left(0, \sigma_\xi^2(x)\right)$, uncorrelated, and not depending on $y(x)$. Its variance at x, denoted by $\sigma_\xi^2(x)$, depends on the value $\xi(x)$ of a "design measure" ξ, which by definition can be any probability measure on $\mathcal{X}$. The influence of the virtual noise is the following. We suppose that the process $y_\xi(x)$ is observed at each point of $\mathcal{X}$. If at some $x \in \mathcal{X}$ the variance $\sigma_\xi^2(x)$ is very large, then in fact we receive from $y_\xi(x)$ no information about θ or β. On the opposite, if $\sigma_\xi^2(x) = 0$, then $y_\xi(x)$ equals to $y(x)$ (a.s.). In the other points, where $0 < \sigma_\xi^2(x) < \infty$, the variable $y_\xi(x)$ retains a part of the information which could be obtained from $y(x)$. In particular, if $\sigma_\xi^2(x) = 0$ for $x \in A$ and $\sigma_\xi^2(x) = \infty$ for $x \notin A$, then the measure ξ gives the same effect in model (2), as the design A in the original process (1). So we can optimize ξ instead of optimizing A. The skill of the method is to choose $\sigma_\xi^2(x)$ so that the optimal ξ in (2) corresponds as far as possible to the optimal design A with $\#A = N$ in (1). The advantage of the method of virtual noise is that we can use standard "continuous algorithms", like the gradient algorithm, to compute the optimal design. Another advantage is the universality of the method: In the algorithm we just have to compute derivatives of the criteria function, and we do not need to enter into the structure of the model.

2. The information matrix and the optimality criteria

For normal errors the Fisher information matrix under the design A in model (1) is equal to

$$M(A;\theta,\beta) = E_\theta \left\{ \begin{array}{cc} -\frac{\partial \ln g(y|\theta,\beta)}{\partial\theta\partial\theta^T} & -\frac{\partial \ln g(y|\theta,\beta)}{\partial\theta\partial\beta^T} \\ -\frac{\partial \ln g(y|\theta,\beta)}{\partial\beta\partial\theta^T} & -\frac{\partial \ln g(y|\theta,\beta)}{\partial\beta\partial\beta^T} \end{array} \right\} = \begin{pmatrix} M_I(A;\theta,\beta) & 0 \\ 0 & M_{II}(A;\beta) \end{pmatrix},$$

where $g\big(y \mid \theta, \beta\big)$ is the normal density of the vector of observations on A,

$$f_\theta(x) = \frac{\partial \eta(x,\theta)}{\partial \theta},$$

$$M_I(A;\theta,\beta) = \sum_{x,z \in \mathcal{X}} f_\theta(x)\big[C^{-1}(A,\beta)\big]_{x,z} f_\theta^T(z),$$

$$\big\{M_{II}(A;\beta)\big\}_{ij} = \frac{1}{2} tr \bigg\{ C^{-1}(A,\beta)\frac{\partial C(A,\beta)}{\partial \beta_i} C^{-1}(A,\beta)\frac{\partial C(A,\beta)}{\partial \beta_j} \bigg\}$$

(cf e.g., [5] for details). Similarly, the Fisher information matrix $M(\xi;\theta,\beta)$ in model (2) has block-diagonal components

$$M_I\left(\xi;\theta,\beta\right) = \sum_{x,z \in \mathcal{X}} f_\theta\left(x\right)\left[C^{-1}\left(\xi,\beta\right)\right]_{x,z} f_\theta^T\left(z\right), \tag{3}$$

$$\left\{M_{II}\left(\xi;\beta\right)\right\}_{ij} = \frac{1}{2} tr \left\{ C^{-1}\left(\xi,\beta\right)\frac{\partial C\left(\mathcal{X},\beta\right)}{\partial \beta_i} C^{-1}\left(\xi,\beta\right)\frac{\partial C\left(\mathcal{X},\beta\right)}{\partial \beta_j} \right\}, \tag{4}$$

where $C(\xi,\beta) = C(\mathcal{X},\beta) + \mathrm{diag}\Big\{\sigma_\xi^2(\mathrm{x});\ \mathrm{x} \in \mathcal{X}\Big\}$.

When the variances $\mathrm{Var}\left[\mathrm{y(x)}\right]$ are small for each x, then the mean square error matrix of the MLE for θ,β is approximately equal to the inverse of the information matrix, even when the number of observations is small (cf. [5]). So we take optimality criteria as functions of the information matrix, i.e., in the form $\Phi\big[M(A;\,\theta,\beta)\big]$, or $\Phi\big[M(\xi;\,\theta,\beta)\big]$. Here $\Phi[M] = -\ln\det(M)$, for D-optimality, $\Phi[M] = tr(M^{-1})$ for A-optimality, etc. In general, we shall consider only functions Φ which are antimonotone, $M \geq M^* \Rightarrow \Phi[M] \leq \Phi[M^*]$, and which have a gradient $\big\{\nabla\Phi[M]\big\}_{ij} = \frac{\partial \Phi[M]}{\partial M_{ij}}$ at each nonsingular M. Notice that $M \geq M^*$ means $u^T M u \geq u^T M^* u$, for every u.

We consider minimization of $\Phi\big[M(A;\,\theta,\beta)\big]$ or $\Phi\big[M(\xi;\,\theta,\beta)\big]$ for a fixed (hypothetical) θ,β (local optimality). *Therefore in the rest of the paper we shall usually omit to write the symbols for θ and β.*

3. Alternatives for $\sigma_\xi^2(x)$

We denote $\xi_{\max} = \max_{x\in\mathcal{X}}\xi(x)$ for any design measure ξ. Take a number $\delta \in \Big(0, \frac{1}{N} - \frac{1}{N+1}\Big)$.

Alternative 1:

$$\sigma_\xi^2\left(x\right) = \ln\frac{\max\left\{\xi_{\max} - \frac{1}{N} + \delta, \delta\right\}}{\max\left\{\xi\left(x\right) - \frac{1}{N} + \delta, 0\right\}}\,.$$

Properties of this choice for $\sigma_\xi^2(x)$ are indicated in [6]. Here we give more details.
Alternative 2:

$$\sigma_\xi^2(x) = \frac{\left(\xi(x) - \frac{1}{N}\right)^2}{\max\left\{\xi(x) - \frac{1}{N} + \delta, 0\right\}}.$$

Alternative 3:

$$\sigma_\xi^2(x) = \rho \ln \frac{1/N}{\min\{\xi(x), 1/N\}}$$

with a tuning parameter $\rho > 0$. We could find in [2] that this is a limit form of alternative 1, after some smoothing.
Alternative 4:

$$\sigma_\xi^2(x) = \rho \frac{\left(\xi(x) - \frac{1}{N}\right)^2}{\xi(x)}$$

which is evidently a simplification of the alternative 2. Again $\rho > 0$ is a tuning parameter.

Proposition 1. *All four alternatives have the following properties:*

- P1: *If $\xi(x) = 0$, then $\sigma_\xi^2(x) = \infty$.*
- P2: *If $\xi_{\max} = \frac{1}{N}$ and $\xi(x) = \xi_{\max}$, then $\sigma_\xi^2(x) = 0$.*
  The alternatives 1 and 2 have supplementary properties:
- P3: *If $\xi_{\max} \leq \frac{1}{N} - \delta$, then $\sigma_\xi^2(x) = \infty$ for every $x \in \mathcal{X}$.*
- P4: *If $\xi_{\max} > \frac{1}{N} - \delta$, then the set $B_\xi = \left\{x \in \mathcal{X} : \xi(x) > \frac{1}{N} - \delta\right\}$ has at least one and at most N points, and $\sigma_\xi^2(x) < \infty$ in these points. In all other points of $\mathcal{X}$ we have $\sigma_\xi^2(x) = \infty$.*

As a consequence, for the alternatives 1 and 2 we have

$$M_I(\xi) \leq M_I(B_\xi), \qquad M_{II}(\xi) \leq M_{II}(B_\xi), \tag{5}$$

for every design measure ξ.

Proof. Properties P1-P4 follow from the definitions of $\sigma_\xi^2(x)$. The number of points in B_ξ is limited by N, since $\sum_{x \in \mathcal{X}} \xi(x) = 1$ and $\delta < \frac{1}{N} - \frac{1}{N+1}$.

If $\xi_{\max} \leq \frac{1}{N} - \delta$, according to P3 and Lemma 2 in Appendix we have $C^{-1}(\xi) = 0$, hence $M_I(\xi) = M_{II}(\xi) = 0$. On the other hand, in the case that $\xi_{\max} > \frac{1}{N} - \delta$, P4 implies that $\sigma_\xi^2(x) = \infty$ when $x \notin B_\xi$.
Therefore, according to Lemma 2 we have

$$C^{-1}(\xi) = \begin{pmatrix} \left[C(\xi \mid B_\xi)\right]^{-1} & 0 \\ 0 & 0 \end{pmatrix},$$

where $C(\xi \mid B_\xi)$ is the submatrix of $C(\xi)$ with rows and columns corresponding to the points of the set B_ξ, and, according to Lemma 1 in Appendix, the inequalities (5) hold. □

For any design $A \subset \mathcal{X},\ \#A = N$ we denote $\xi_A(x) = \frac{1}{N}$ if $x \in A$, $\xi_A(x) = 0$ if $x \notin A$. As a consequence of Proposition we have the following theorem.

Theorem 1. *For all 4 alternatives we have*

$$M_I(\xi_A) = M_I(A), \qquad M_{II}(\xi_A) = M_{II}(A)$$

for every A such that $\#A = N$.

For the alternatives 1 and 2 we have

$$\mu \in \arg\min_\xi \Phi\big[M(\xi)\big] \tag{6}$$

if and only if $\mu = \xi_A$, where

$$A \in \arg\min_{B,\,\#B \le N} \Phi\big[M(B)\big], \tag{7}$$

i.e., the original task of experimental design (7) *is equivalent to the modified task* (6).

Although the alternatives 1 and 2 give exact solutions of the original design problem, they still can not be used, since derivatives can be computed only after some complicated smoothing (cf. [2]). On the other hand, for the alternatives 3 and 4 the problems (7) and (6) are not strictly equivalent. However, if ρ is very large, then $\sigma_\xi^2(x)$ is very large, "almost" ∞ if $x \notin B_\xi$. As a consequence, the design B_ξ, having no more than N points, is dominating any design ξ if ρ is sufficiently large (in principle, another ρ for each ξ), and the solution of the problem (7) can be approached by the solution of the problem (6). The choice of ρ is a question of skill: if ρ is too large, then we can obtain false local minima, because the solved problem of minimization is nonconvex.

4. The derivatives and the gradient algorithm

We consider only the alternatives 3 and 4. We shall use the notation $\xi_\lambda = (1-\lambda)\,\xi + \lambda\mu$, where ξ, μ are fixed design measures, and $\lambda \in (0,1)$. We consider the Gâteaux derivatives (in the point ξ and in the direction of μ)

$$\frac{\partial\Phi\big[M(\xi_\lambda)\big]}{\partial\lambda}\,\Big|_{0^+} = \lim_{\lambda \searrow 0} \frac{\Phi\big[M(\xi_\lambda)\big] - \Phi\big[M(\xi)\big]}{\lambda}$$

and similarly, we define $\left[\partial\sigma^2_{\xi_\lambda}(x)/\partial\lambda\right]_{0^+}$. For the alternative 3 we have

$$\frac{\partial\sigma^2_{\xi_\lambda}(x)}{\partial\lambda}\,|_{0^+}= \rho\left[1-\frac{\mu(x)}{\xi(x)}\right]$$

if $\xi(x) < \frac{1}{N}$, or if $\xi(x) = \frac{1}{N}$ and $\mu(x) \leq \frac{1}{N}$, and $\left[\partial\sigma^2_{\xi_\lambda}(x)/\partial\lambda\right]_{0^+} = 0$ for all other values of $\xi(x)$ and $\mu(x)$. For the alternative 4 we have

$$\frac{\partial\sigma^2_{\xi_\lambda}(x)}{\partial\lambda}\,|_{0^+}= \rho\left[1-\left(\frac{1}{N\xi(x)}\right)^2\right]\left[\mu(x)-\xi(x)\right].$$

THEOREM 2. *We can write*

$$\frac{\partial\Phi\left[M(\xi_\lambda)\right]}{\partial\lambda}\,|_{0^+}= \sum_{u\in\mathcal{X}}\left[d_I(u,\xi)+d_{II}(u,\xi)\right]\frac{\partial\sigma^2_{\xi_\lambda}(u)}{\partial\lambda}\,|_{0^+},$$

where

$$\begin{aligned}
d_I(u,\xi) &= -a^T_\xi(u)\nabla_I\Phi\left[M(\xi)\right]a_\xi(u),\\
d_{II}(u,\xi) &= -tr\Big\{\nabla_{II}\Phi\left[M(\xi)\right]\ A(u,\xi)\Big\},\\
a_\xi(u) &\in \mathbb{R}^p,\\
a_\xi(u) &= \sum_{z\in\mathcal{X}}\left\{C^{-1}(\xi)\right\}_{u,z}f(z),\\
A(u,\xi) &\in \mathbb{R}^{q\times q},\\
\left\{A(u,\xi)\right\}_{ij} &= \left\{C^{-1}(\xi)\right\}_{u,\cdot}\frac{\partial C(\mathcal{X})}{\partial\beta_i}C^{-1}(\xi)\frac{\partial C(\mathcal{X})}{\partial\beta_j}\left\{C^{-1}(\xi)\right\}_{\cdot,u}
\end{aligned}$$

and $\nabla_I\Phi\left[M(\xi)\right]$ *and* $\nabla_{II}\Phi\left[M(\xi)\right]$ *are the* $p\times p$ *and the* $q\times q$ *diagonal submatrices of the* $(p+q)\times(p+q)$ *gradient matrix* $\nabla\Phi\left[M(\xi)\right]$. *If* Φ *is convex, then for any* $u\in\mathcal{X}$ *and any design measure* ξ *we have* $d_I(u,\xi)\geq 0$, $d_{II}(u,\xi)\geq 0$.

Proof. Since $M(\xi)$ is a block-diagonal matrix, we can write

$$\frac{\partial\Phi\left[M(\xi_\lambda)\right]}{\partial\lambda}\,|_{0^+} = tr\left\{\nabla_I\Phi\left[M(\xi)\right]\frac{\partial M_I(\xi_\lambda)}{\partial\lambda}\,|_{0^+}\right\}+tr\left\{\nabla_{II}\Phi\left[M(\xi)\right]\frac{\partial M_{II}(\xi_\lambda)}{\partial\lambda}\,|_{0^+}\right\}$$

and the required expressions are obtained by taking the derivatives and using that $\left[\partial C^{-1}(\xi_\lambda)/\partial\lambda\right]_{0^+} = -C^{-1}(\xi_\lambda)\left[\partial C(\xi_\lambda)/\partial\lambda\right]_{0^+}C^{-1}(\xi_\lambda)$. If $\Phi[M]$ is convex, then $\nabla\Phi[M]$, hence also $\nabla_I\Phi[M]$ and $\nabla_{II}\Phi[M]$ are nonpositive definite (cf. [3]). So $d_I(u,\xi)\geq 0$. Further, from the eigenvalue decomposition we have $C^{-1}(\xi) = \sum_k c_k\, v_k v_k^T$ with $c_k\geq 0, v_k\in\mathbb{R}^{\#(\mathcal{X})}$.

Hence

$$d_{II}(u,\xi) = -\sum_k c_k h^T(u,k,\xi)\,\nabla_{II}\Phi\big[M(\xi)\big]h(u,k,\xi) \geq 0$$

with

$$\{h(u,k,\xi)\}_i = \left\{C^{-1}(\xi)\right\}_{u.}\left[\partial C(\mathcal{X})/\partial\beta_i\right]v_k\,.$$

□

The gradient algorithm and examples. We start with the design measure ξ_o, which is uniform on $\mathcal{X}$, and at the nth step we take $\xi_{n+1} = (1-\lambda_n) + \lambda_n\mu_n$, say with $\lambda_n = 1/(n+1)$. The correcting measure μ_n is taken so that the directional derivative of Φ at ξ_n and in the direction of μ_n is minimized. Since this derivative is linear in μ_n, we can take for μ_n the Dirac measure at some x_n which is a solution of

$$x_n \in \arg\min_{x\in\mathcal{X}}\Big\{\big[d_I(x,\xi_n) + d_{II}(x,\xi_n)\big]\big[\partial\sigma^2_{\xi_\lambda}(x)/\partial\lambda\big]_{0^+}\Big\}.$$

We stop the algorithm when the design ξ_n is stabilized and ξ_n is distributed uniformly on some set A with $\#A = N$. If not, we change the tuning parameter ρ, and restart again.

We see that $\big[d_I(x,\xi) + d_{II}(x,\xi)\big]$ *measures the importance* of the point x$\in\mathcal{X}$ when the design ξ is considered. In the particular case of the linear model with homoscedastic uncorrelated observations we obtain the well known important expression $\big[d_I(x,\xi) + d_{II}(x,\xi)\big] = -f^T(x)\,\nabla\Phi\big[M(\xi)\big]f(x)$.

We used the alternatives 3 and 4 to compute all examples of linear models presented in [2] again, and we also computed the locally optimum D-optimal design in the model with $\eta(\theta,x) = \theta_1 + \theta_2 x$, $C(x,z,\beta) = \beta_1\exp\{-\beta_2 x\}$, $\mathcal{X} = \{1,2,\ldots,20\}$, $N = 5$, in the locality of $\beta_1 = 10, \beta_2 = 0.1$ or of $\beta_1 = 1$, $\beta_2 = 1$. We found the algorithm very fast, but different values of ρ must be checked. If ρ is too small, the algorithm does not converge, if it's too large, false minima are detected. It is also more difficult to use the algorithm when the net $\mathcal{X}$ on the interval $< 0, 20 >$ is very dense.

Appendix

Lemma 1. *Let Δ be a p.s.d. matrix. Denote $M_I^*(C) = \sum\limits_{x\in\mathcal{X}} f(x)\left[C^{-1}\right]_{x,z} f^T(z)$, $M_{II}^*(C) = \frac{1}{2}tr\left\{C^{-1}\frac{\partial C(\mathcal{X},\beta)}{\partial\beta}C^{-1}\frac{\partial C(\mathcal{X},\beta)}{\partial\beta^T}\right\}$ for any p.d. matrix C. We have*

$$M_I^*[C+\Delta] \leq M_I^*[C]\,, \quad M_{II}^*[C+\Delta] \leq M_{II}^*[C]\,.$$

P r o o f. We have $[C+\Delta] \geq C \Rightarrow [C+\Delta]^{-1} \leq C^{-1}$ (cf. [1]), which implies the first inequality. Further, for any vector u

$$u^T M_{II}^* [C+\Delta]\, u = \frac{1}{2} tr\Big\{ [C+\Delta]^{-1} Q\, [C+\Delta]^{-1} Q \Big\},$$

where $Q = \sum_i u_i \frac{\partial C(A,\beta)}{\partial \beta_i}$. Since C, Δ are symmetric, and C is positive definite, there is a nonsingular matrix U such that $U^T C U = I$, $U^T \Delta U = \mathrm{diag}\{\lambda_i\}$ with $\lambda_i \geq 0$ (cf. [1], beginning of Section 21.14). Hence

$$\begin{aligned} 2u^T M_{II}^* [C+\Delta]\, u &= \sum_i \frac{1}{1+\lambda_i} \Big\{ U^T Q\, [C+\Delta]^{-1} Q U \Big\}_{ii} \\ &\leq \sum_i \big\{ U^T Q C^{-1} Q U \big\}_{ii} = 2u^T M_{II}^* [C]\, u. \end{aligned}$$

□

Lemma 2. *Let C be a symmetric p.d. $t \times t$ matrix and take $\Delta(k) = \mathrm{diag}\big\{\lambda_i(k);\ i = 1, \dots, t\big\} \geq 0$, such that $\lim_{k\to\infty} \lambda_i(k) = \lambda_i < \infty$; $i = 1, \dots, s$, $\lim_{k\to\infty} \lambda_i(k) = \infty$; $i = s+1, \dots, t$. Then*

$$\lim_{k\to\infty} \big[C + \Delta(k)\big]^{-1} = \begin{pmatrix} \Big[\{C_{i,j} + \lambda_i \delta_{ij}\}_{i,j=1,\dots,s}\Big]^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

P r o o f. Consider a decomposition into blocks

$$C + \Delta(k) = \begin{pmatrix} A & b \\ b^T & c(k) \end{pmatrix},$$

where $c(k) = C_{t,t} + \lambda_t(k) \in \mathbb{R}$. According to [1], Theorem 8.5.11 we have

$$\big[C + \Delta(k)\big]^{-1} = \begin{pmatrix} Q^{-1} & -Q^{-1} b / c(k) \\ -b^T Q^{-1} / c(k) & b^T Q^{-1} b / c^2(k) \end{pmatrix},$$

where $Q = A - bb^T / c(k)$. Take the limit $\lambda_t(k) \to \infty$, but let $\lambda_1(k), \dots, \lambda_{t-1}(k)$ be fixed. We obtain

$$\big[C + \Delta(k)\big]^{-1} \to_{k\to\infty} \begin{pmatrix} A^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Since A is the $(t-1) \times (t-1)$ submatrix of $C + \Delta(k)$, we proceed by induction.

□

REFERENCES

[1] HARVILLE, D. A.: *Matrix Algebra from a Statistician's Perspective*. Springer-Verlag, New York, 1997.

[2] MÜLLER, W. G.—PÁZMAN, A.: *An algorithm for the computation of optimum experimental design under a given covariance structure*, Comput. Statist. **14** (1999), 197–211.

[3] MÜLLER, W. G.—PÁZMAN, A.: *Measures for designs of experiments with correlated errors*, Biometrika **90** (2003), 423–434.

[4] NÄTHER, W.: *Effective Observations of Random Fields*. Teubner Texte zur Mathematik, Band 72, Teubner Verlag, Leipzig, 1985.

[5] PÁZMAN, A.: *Correlated optimum design with a parametrized covariance function*, Report no. 5, Dept. of Statistics and Mathematics, Wirtschaftsuniversität Wien, June 2004, `http://statistik.wu-wien.ac.at/`.

[6] PÁZMAN, A.—MÜLLER, W. G.: *Optimal design of experiments subject to correlated errors*, Statist. & Probab. Letters **52** (2001), 29–34.

[7] UCIŃSKI, D.—ATKINSON, A. C.: *Experimental design for time-dependent models with correlated observations*, Stud. Nonlinear Dyn. Econom. **8** (2004), Berkeley, CA, (electronic).

Received September 14, 2006

Department of Applied Mathematics and Statistics
Faculty of Mathematics, Physics and Informatics
Comenius University
Mlynská dolina
SK-842 48 Bratislava
SLOVAKIA

E-mail: pazman@fmph.uniba.sk
laca@fmph.uniba.sk