Contributions to the theory of optimal experimental design

Mgr. Radoslav Harman, PhD.

Department of Applied Mathematics and Statistics
Faculty of Mathematics, Physics and Informatics
Comenius University Bratislava

May 2011

Outline

- Brief introduction to optimal design of experiments
- Notation used in the statements of results
- Key result 1: A method for evaluating criterion-robustness of designs with respect to a very wide class of criteria
- Key result 2: An inequality for removing “useless” support points in *D*-optimum design algorithms
- Key result 3: Transformation of the problem of approximate **c**-optimality to the linear programming form
- Key result 4: Formulas that facilitate construction of exact optimal designs for models with a product covariance structure of errors

Introduction to optimal design of experiments

- Design of experiments: A collection of methods for selecting a small number of trials (conditions under which the experiment is performed) in order to obtain as much valuable information as possible.

We will consider the linear regression model on an experimental domain $\mathfrak{X}$. For each design point $x \in \mathfrak{X}$ representing experimental conditions, we can observe a random variable

$$Y(x) = \beta' \mathbf{f}(x) + \varepsilon(x),$$

where $\mathbf{f} : \mathfrak{X} \to \mathbb{R}^m$ is a known vector of regression functions, $\beta \in \mathbb{R}^m$ is an unknown vector of parameters, and $\varepsilon(x)$ is an unobservable random error. An “optimal design” is a method of choosing the conditions $x_1, x_2, ... \in \mathfrak{X}$ in an optimal way to gain the largest amount of information about the unknown parameter β of the model.

Notation

- Ξ ... the set of all approximate designs on $\mathfrak{X}$ (finitely supported probabilities on $\mathfrak{X}$)
- $\mathbf{M}(\xi)$ ... the information matrix of $\xi \in \Xi$ (the "larger" $\mathbf{M}(\xi)$ the more "informative" is the design ξ)
- $\Phi : \mathcal{S}_+^m \to [0, \infty)$ ... an optimality criterion (information function, measuring the "size" of information matrices)
  - $\Phi_D(\mathbf{M}) = \det^{1/m}(\mathbf{M})$ ... the criterion of *D*-optimality
  - $\Phi_{\mathbf{c}}(\mathbf{M}) = (\mathbf{c}'\mathbf{M}^-\mathbf{c})^{-1}$ for $\mathbf{c} \in \mathcal{L}(\mathbf{M})$ ... the criterion of **c**-optimality
- $\xi^* = \mathrm{argmax}_{\xi \in \Xi} \Phi(\mathbf{M}(\xi))$ ... the Φ-optimal design
- $\mathrm{eff}(\xi|\Phi) = \Phi(\mathbf{M}(\xi))/\Phi(\mathbf{M}(\xi^*))$ ... the efficiency of ξ (measures how good is ξ compared to the Φ-optimal design ξ^*)

During the last half a century, researchers have amassed a vast number of analytic results about optimal designs for all kinds of regression models with respect to many optimality criteria. There also exist several algorithms for numerical construction of optimal design.

Monograph: A. Pázman "Foundations of optimum experimental design", Springer 1986

Key result 1

Harman R (2004): Minimal efficiency of designs under the class of orthogonally invariant information criteria, Metrika, Volume 60, No. 2, pp. 137-153

Theorem

Let $\xi \in \Xi$ be an arbitrary design, let $\mathbb{O}$ be the set of all information functions depending only on the eigenvalues of the information matrix, and, for all $\mathbf{M} \in \mathcal{S}_+^m$, let $\Phi_{E_k}(\mathbf{M})$ be the sum of the k largest eigenvalues of $\mathbf{M}$, $k = 1, ..., m$. Then

$$\inf_{\Phi \in \mathbb{O}} \mathrm{eff}(\xi|\Phi) = \min_{k=1,\dots,m} \mathrm{eff}(\xi|\Phi_{E_k}).$$

Allows us to calculate a measure of criterion-robustness of a given design, and sometimes construct the most criterion-robust design in the class $\mathbb{O}$ of all reasonable optimality criteria that measure the quality of the designs by a “size” of the confidence ellipsoid, independently on its rotation and a shift.

Key result 1

Related publications of R. Harman:

- Harman R (2004): Lower bounds on efficiency ratios based on Φ_p-optimal designs, in: Proceedings from the conference "Advances in Model-Oriented Design and Analysis" (Heeze, Netherlands 2004), Physica-Verlag, pp. 89-96
- Harman R (2008): Equivalence theorem for Schur optimality of experimental designs, Journal of Statistical Planning and Inference, Volume 138, Issue 4, pp. 1201-1209
- Filová L, Harman R, Klein T (2011): Approximate E-optimal designs for the model of spring balance weighing with a constant bias, Journal of Statistical Planning and Inference, Volume 141, Issue 7, pp. 2480-2488
- Work in progress with L. Filová and J. Michaliková.

Cited in: Canadian Journal of Statistics, Kybernetika, Collecting Spatial Data: Optimum Design of Experiments for Random Fields (Springer), Boletín de Estadística e Investigación Operativa, etc.

Key result 2

Harman R, Pronzato L (2007): Improvements on removing non-optimal support points in D-optimum design algorithms, Statistics & Probability Letters, Volume 77, Issue 1, pp. 90-94

Theorem

Let $\xi \in \Xi$, let $\mathbf{M}(\xi)$ be regular and let $d(x,\xi) = \mathbf{f}'(x)\mathbf{M}^{-1}(\xi)\mathbf{f}(x)$ for all $x \in \mathfrak{X}$. Let $\epsilon = \max_{x \in \mathfrak{X}} d(x,\xi) - m$. If $x_0 \in \mathfrak{X}$ is such that

$$d(x_0,\xi) < m\left(1 + \epsilon/2 - \sqrt{\epsilon(4 - 4/m + \epsilon)}/2\right),$$

then x_0 does not support any D-optimal design.

Enables removing a large amount of “useless” support points from the experimental domain which speeds up the *D*-optimum design algorithms (e.g., the multiplicative algorithm). The improvement in efficiency can in some cases be several orders of magnitude. It can also be used to speed up algorithms for constructing minimal volume ellipsoid containing a given set of points.

Key result 2

Related publications of R. Harman:

- Harman R (2003): A method how to delete points which do not support a D-optimal design, Tatra Mt. Math. Publ. 26, pp. 59-67
- Harman R, Trnovská M (2009): Approximate D-optimal designs of experiments on the convex hull of a finite set of information matrices, Mathematica Slovaca 59, No. 6, pp. 693–704

Cited in: Annals of Statistics (2x), Computational Statistics and Data Analysis (2x), Journal of Statistical Planning and Inference (2x), Statistics and Computing, An Introduction to Optimal Designs for Social and Biomedical Research (Wiley), etc.

Key result 3

Harman R, Jurík T (2008): Computing c-optimal experimental designs using the simplex method of linear programming, Computational Statistics & Data Analysis, Volume 53, Issue 2, pp. 247-254

Theorem

Let $\xi \in \Xi$, $\mathbf{c} \in \mathbb{R}^m$ *and let* $\mathbf{F} = (\mathbf{f}_1, ..., \mathbf{f}_{2k})$, *where* $\mathbf{f}_j = \mathbf{f}(x_j), \mathbf{f}_{j+k} = -\mathbf{f}(x_j)$ *for* $j \in \{1, \ldots, k\}$. *Then the design* ξ *is* $\mathbf{c}$*-optimal if and only if* $\xi(x_j) = \alpha_j + \alpha_{j+k}$ *for all* $j = 1, ..., k$ *and for some solution* $(\alpha', h)' \in \mathbb{R}^{2k+1}$ *of the linear programming problem:*

$$\max \left\{ h \middle| \begin{pmatrix} \mathbf{F} & -\mathbf{c} \\ \mathbf{1}'_{2k} & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ h \end{pmatrix} = \begin{pmatrix} \mathbf{0}_m \\ 1 \end{pmatrix}, \alpha \geq \mathbf{0}_{2k}, h \geq 0 \right\}.$$

Thus, the results of LP can be used to formulate statements about **c**-optimality. Permits using the algorithms of LP to find the **c**-optimal designs, which overcome other **c**-optimal design algorithms.

Cited in: Journal of Statistical Planning and Inference, Computational Optimization and Applications, Signal Processing Letters, Proceedings of Compstat 2010

Key result 4

Harman R, Štulajter F (2011): Optimal sampling designs for the Brownian motion with a quadratic drift, Journal of Statistical Planning and Inference, Volume 141, Issue 8, pp. 2750–2758

Theorem

Consider a linear regression model $X(t_i) = \sum_{j=1}^{m} \beta_j f_j(t_i) + \epsilon(t_i)$ *where* $i = 1, ..., n$, $t_1 < \cdots < t_n$ *and* $\mathrm{cov}(\epsilon(t_{i_1}), \epsilon(t_{i_2})) = u(t_{i_1})v(t_{i_2})$ *for* $i_1 < i_2$*. Then the element* (r, s) *of the information matrix for the parameter* $\beta = (\beta_1, ..., \beta_m)'$ *is given by*

$$(\mathbf{M})_{r,s} = \frac{f_r(t_1)f_s(t_1)}{u(t_1)v(t_1)} + \sum_{i=2}^{n} \frac{\left(\frac{f_r(t_i)}{v(t_i)} - \frac{f_r(t_{i-1})}{v(t_{i-1})}\right)\left(\frac{f_s(t_i)}{v(t_i)} - \frac{f_s(t_{i-1})}{v(t_{i-1})}\right)}{\left(\frac{u(t_i)}{v(t_i)} - \frac{u(t_{i-1})}{v(t_{i-1})}\right)}.$$

Key result 4

Simplifies constructing exact optimal designs for the models with a specific correlation structure of errors. It has been used to prove exact optimality of equidistant sampling designs in several models, involving the Wiener and Ornstein-Uhlenbeck covariance structure of errors.

Related publication of R. Harman:

- Harman R, Štulajter F (2009): Optimality of equidistant sampling designs for a nonstationary Ornstein-Uhlenbeck process, In: Proceedings of the 6th St.Petersburg Workshop on Simulation, pp. 1097-1101
- Work in progress with F. Štulajter and V. Lacko

Thank you for attention.