Towards an Equivalence Theorem for Computer Simulation Experiments (Abstract)

Werner G. Müller*

February 17, 2009

For the development of design theory for experiments with independent observations, the so called equivalence theorem of Kiefer and Wolfowitz (1961), henceforth KWEQ, and its extensions, has played a major role. It allows quick checks, whether given designs are optimal and led to the development of efficient algorithms for constructing good designs. One of the key aspects of the KWEQ is the establishment of the equivalence of optimal designs between two criteria of optimality, one related to parameter estimation, the other related to prediction (classically between D- and G-optimality).

Unfortunately, since in the analysis of computer simulation data it become standard to operate under the assumptions of dependent errors (cf. Sacks et al. 1989), most of the conditions of the KWEQ are not met. In this paper we intend to present the main obstacles towards establishing a similar relationship for this case and investigate some possible avenues of solutions. Non-additivity and nonconvexity are amongst these obstacles, which have recently been reviewed by Müller and Stehlik, 2009a. Furthermore the concept of Fisher information is conveniently used as a basis for designing efficient experiments. However, if the output stems from computer simulations they are often approximated as realisations of correlated random fields. Consequently, the conditions under which Fisher information may be suitable must be restated, which was undertaken in Stehlik and Müller, 2008.

Contrasting estimating underlying parameters and predicting a random field at given unsampled settings or over a continuous region, the second task usually leads to minimizing the maximum kriging variance at these sites or over a region $\mathcal{X}$, interpreted as the unconditional mean squared prediction error for the best linear unbiased predictor, i.e.

$$\min_{\xi} \max_{x \in \mathcal{X}} E[(\hat{y}(x|\xi) - y(x))^2]. \tag{1}$$

If the covariance parameters are estimated from the same dataset, the resulting additional uncertainty needs to enter the criterion. This uncertainty is frequently approximated by the trace $\mathrm{tr}\left\{M_\theta^{-1}\mathrm{Var}[\partial\hat{y}(x_0)/\partial\theta]\right\}$, cf. Harville and Jeske (1992) Consequently, Zimmerman (2006) regards

$$\min_{\xi} \max_{x \in \mathcal{X}} \left\{\mathrm{Var}[\hat{y}(x)] + \mathrm{tr}\left\{M_\theta^{-1}\mathrm{Var}[\partial\hat{y}(x)/\partial\theta]\right\}\right\} \tag{2}$$

---

*Department of Applied Statistics, Johannes-Kepler-University Linz Freistädter Straße 315, A-4040 Linz, Austria, email:Werner.Mueller@jku.at

as the design problem, which he terms EK-(empirical kriging-)optimality. These designs are much more demanding to achieve than parameter estimation designs, since they require embedded optimizations over the candidate sets. For parameter estimation, on the other hand, particularly for determining a D-optimal design, Müller and Stehlik (2009b) have suggested a compound criterion (according to Läuter, 1976):

$$\bar{\Phi}'[\xi|\alpha] = |M(\xi)|^{\alpha} \cdot |M'(\xi)|^{(1-\alpha)}, \tag{3}$$

which consist of respective determinants corresponding to trend and covariance parameters, α being a weighing factor. We intend to investigate further how these criteria relate in the spirit of the KWEQ.

References

Sacks J, Welch WJ, Mitchell TJ and Wynn HP 1989 Design and analysis of computer experiments *Statistical Science* **4**(4) p 409-435

Harville DA, Jeske DR. Mean squared error of estimation or prediction under a general linear model. *Journal of the American Statistical Association* 1992; **87**:724-731.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, 12:363–366.

Läuter, E. (1976). Optimal multipurpose designs for regression models. *Mathematische Operationsforschung und Statistik, Series Statistics*, 7:51–68.

Müller W.G. and Stehlík M. (2009). Issues in the Optimal Design of Computer Simulation Experiments, *Applied Stochastic Models in Business and Industry*, DOI: 10.1002/asmb.740.

Müller W.G. and Stehlík M. (2009). Compound Optimal Spatial Designs, *Environmetrics*, forthcoming.

Stehlík M. and Müller W.G. (2008). Fisher information in the design of computer simulation experiments, accepted for *Journal of Physics: Conference Series*, 2008.

Zimmerman DL. Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction. *Environmetrics* 2006; **17**:635–652.