Statistica Sinica **9**(1999), 1083-1088

UNIVERSALLY OPTIMAL DESIGNS FOR COMPUTER EXPERIMENTS

Hongquan Xu

Nankai University and University of Michigan

Abstract: The concept of universal optimality from optimum design theory is introduced into computer experiments, modeled as realizations of stationary Gaussian processes. When the correlation function is a nondecreasing and convex function of a distance measure, it is shown that a design is universally optimal if it is equidistant and of maximum average distance. Examples of universally optimal designs are given with respect to rectangular, Euclidean, Hamming, and Lee distances.

Key words and phrases: Computer experiments, Hamming distance, Lee distance, orthogonal arrays, universally optimal designs.

1. Introduction

A computer experiment is modeled as a realization of a stochastic process, often in the presence of nonlinearity and high dimensional inputs (see Sacks, Welch, Mitchell and Wynn (1989)). In this setting good designs are critical to efficient data analysis and prediction. Many authors have suggested designs, among which two classes of designs are known to be efficient: Latin hypercubes and maximin distance designs. Latin hypercubes, proposed by McKay, Beckman and Conover (1979), are evenly distributed in each one-dimensional projection. To improve their high-dimensional properties, Owen (1992) and Tang (1993) proposed randomized orthogonal arrays and orthogonal array-based Latin hypercubes, respectively. Maximin distance designs, proposed by Johnson, Moore and Ylvisaker (1990), are D-optimal for extremely weak correlations. However, maximin distance designs need not have good projection properties under Euclidean or rectangular distance. To overcome this shortcoming, Morris and Mitchell (1995) suggested restricting designs to Latin hypercubes and proposed use of maximin Latin hypercubes.

The purpose of this paper is to introduce the concept of universal optimality from optimum design theory into computer experiments, and then to exhibit some universally optimal designs with respect to different distance measures. The paper is organized as follows. Section 2 develops a concept of universal optimality based on correlation matrices. Section 3 presents examples of universally optimal designs with respect to rectangular, Euclidean, Hamming, and Lee distances. A brief conclusion is given in Section 4.

2. A Concept of Universal Optimality

The original concept of universal optimality in Kiefer (1975) dealt with information matrices with zero row and column sums. Here we consider correlation matrices with unit diagonal elements.

First we introduce the Bayesian approach of Currin, Mitchell, Morris and Ylvisaker (1991). Let Y be a stationary Gaussian process on a design space Ω, with mean zero and correlation function $\rho(\cdot,\cdot)$. For a design (subset) X of Ω, let Y_X be the vector of $Y(x), x \in X$, and R_X be the correlation matrix of Y_X. It is well known that the posterior process given the observed response Y_X is also Gaussian. The design problem is to minimize the amount of uncertainty in the posterior process by choosing the design X. Taking *entropy* as the measure of uncertainty, Shewry and Wynn (1987) showed that this is equivalent to maximizing the prior entropy of Y_X if Ω is finite. Under the assumption of stationarity, this is the same as the *D–criterion*, i.e., maximization of $\det(R_X)$.

Our concept of universal optimality is based on correlation matrices. Let $C_b = (1-b)I_n + bJ_n$ be an $n \times n$ matrix, where I_n is the identity and J_n consists of all 1's. Let g be a permutation of $\{1,\ldots,n\}$, G_g be the corresponding permutation matrix of g (i.e., G_g is obtained by permuting the columns of I_n according to g), and $G_g^{'}$ be the transpose of G_g. Suppose that $\mathcal{R}_n$is the set of all $n \times n$ correlation matrices (i.e., nonnegative definite matrices with unit diagonal elements) and that $\Phi : \mathcal{R}_n \to (-\infty, +\infty]$ satisfies
(a) $\Phi(\cdot)$ is convex;
(b) $\Phi(C_b)$ is nondecreasing in $0 \le b \le 1$;
(c) $\Phi(R_n) = \Phi(G_g^{'} R_n G_g)$ holds for any permutation g and any $R_n \in \mathcal{R}_n$.
Let $\mathcal{X}_n$ be the set of all n-run designs of Ω.

Definition 1. A design X^* in $\mathcal{X}_n$ is called universally optimal if X^* minimizes $\Phi(R_X)$, $X \in \mathcal{X}_n$, for every Φ satisfying (a), (b) and (c).

Note that $\Phi(R) = -\log\det(R)$ satisfies (a), (b) and (c); therefore, universal optimality covers the D–criterion. In addition, if the correlation function is a strictly decreasing function of a distance measure, then universal optimality covers the maximin distance criterion as well.

The correlation function is often taken to be a function of a distance measure d on Ω. For an n-run design X, let $d_{min}(X) = \min\{d(x,x') : x, x' \in X, x \neq x'\}$ and $d_{ave}(X) = [n(n-1)]^{-1} \sum_{x \neq x'} d(x,x')$ denote the minimum and average distance between all pairs of different points of X, respectively. Clearly, $d_{min}(X) \le d_{ave}(X)$. A design is called *equidistant* if $d_{min}(X) = d_{ave}(X)$.

Proposition 1. *Suppose that the correlation function* $\rho(\cdot,\cdot)$ *is a nonnegative, nonincreasing, and convex function of a distance measure* $d(\cdot,\cdot)$*. If* $X^* \in \mathcal{X}_n$ *satisfies* (i) X^* *is equidistant,* (ii) $d_{ave}(X^*) = \max_{X\in\mathcal{X}_n} d_{ave}(X)$*, then* X^* *is universally optimal.*

Proof. For any $X \in \mathcal{X}_n$, the $n\times n$ correlation matrix of Y_X is $R_X = (\rho(x,x')) = (\rho(d(x,x')))$. By (c) and (a),

$$\Phi(R_X) = (n!)^{-1}\sum_g \Phi(G_g' R_X G_g) \geq \Phi((n!)^{-1}\sum_g G_g' R_X G_g) = \Phi(C_b),$$

where $C_b = (1-b)I_n + bJ_n$ with $b = [n(n-1)]^{-1}\sum_{x\neq x'} \rho(d(x,x'))$. In particular, by (i), $R_{X^*} = C_{b^*}$ where $b^* = \rho(d_{ave}(X^*))$. Then, by the assumption of the correlation function ρ and (ii),

$$b \geq \rho(d_{ave}(X)) \geq \rho(d_{ave}(X^*)) = b^* \geq 0.$$

Thus, by (b), $\Phi(R_X) \geq \Phi(C_b) \geq \Phi(C_{b^*}) = \Phi(R_{X^*})$. This completes the proof.

3. Examples

This section presents some examples of universally optimal designs that come from applying Proposition 1 to the special case $\Omega = \{0,\ldots,q-1\}^s$ with different distance measures. In particular, rectangular, Euclidean, Hamming and Lee distances will be considered. For simplicity in notation, let $F_q = \{0,\ldots,q-1\}$ and $F_q^s = \{0,\ldots,q-1\}^s$. F_q^s is often viewed as a lattice in the unit cube $[0,1]^s$ if F_q is mapped into $[0,1]$ by: $x \mapsto (x+0.5)/q$.

To apply Proposition 1, we require suitable correlation functions. There are many examples of correlations in the literature that are functions of rectangular or Euclidean distance. It is less common to find correlations that are functions of Hamming or Lee distance. Here are some examples. If the random process Y is periodic on F_q^s, or if F_q^s is viewed as a torus lattice, Lee distance is more suitable than rectangular or Euclidean distance. For a two-dimensional example see Martin (1982), who studied designs for stationary torus lattice processes. On the other hand, if the predictors or factors of Y are categorical or qualitative rather than quantitative in nature, correlation functions might best be functions of Hamming distance. Indeed only Hamming distance is studied in classical factorial designs.

Since the focus here is on distance measures rather than on correlation functions, we simply assume that the correlation function is a nonnegative, nonincreasing, and convex function of any distance measure in question. Thus, in order to prove universal optimality, it is sufficient by Proposition 1 to verify that a design is of maximum average distance and equidistant.

Here are some additional definitions and notation for the following examples. Let $a = (a_1, \ldots, a_s)$ and $b = (b_1, \ldots, b_s)$ be two vectors over F_q. Rectangular distance is given by $d_R(a,b) = \sum |a_i - b_i|$; Euclidean distance is $d_E(a,b) = \{\sum (a_i - b_i)^2\}^{1/2}$; Hamming distance is $d_H(a,b) = \text{card}\{i : a_i \neq b_i, 1 \leq i \leq s\}$; and Lee distance is $d_L(a,b) = \sum \min\{|a_i - b_i|, q - |a_i - b_i|\}$. Note that Lee distance depends on q. It is convenient to consider F_q as an integer ring modulo q when dealing with Lee distance. In particular, F_q is a finite field if q is a prime number.

An $n \times s$ matrix A over F_q is called an *orthogonal array* of size n, s constraints, q levels, strength t, and index λ if any set of t columns of A contains all q^t possible row vectors exactly λ times. Such an array is denoted by $OA(n,s,q,t)$. Clearly $n = \lambda q^t$. An $OA(n,s,q,2)$ is saturated if $n - 1 = s(q-1)$. Moreover, an $OA(n,s,n,1)$ is called a *Latin hypercube* here. For convenience, a design is written as a matrix, each row corresponding to a run. Thus, an $n \times s$ matrix over F_q is an n-run design on F_q^s.

A. Rectangular and Euclidean distances

Let X be an $n \times s$ matrix over F_q. It can be shown that $d_{ave}(X) \leq (q-1)ns/[2(n-1)]$ for rectangular distance and $d_{ave}(X) \leq (q-1)\{ns/[2(n-1)]\}^{1/2}$ for Euclidean distance. A necessary condition for equality is that all elements of X are either 0's or $(q-1)$'s and that the number of 0's and of $(q-1)$'s in each column of X are the same. For rectangular distance this condition is also sufficient. It can be shown (see Hamming distance below) that saturated $OA(n, s = n-1, 2, 2)$ are equidistant and of maximum average distance. Thus, by Proposition 1, they are universally optimal with respect to both rectangular and Euclidean distances. In general, by multiplying $OA(n, n-1, 2, 2)$ by $q-1$, one gets a universally optimal design over F_q with $s = n-1$.

B. Hamming distance

The Plotkin bound in coding theory (see Berlekamp (1968), pp.311-315) shows that, for any $n \times s$ matrix X over F_q, $d_{ave}(X) \leq (q-1)ns/[q(n-1)]$, with equality if and only if X is an orthogonal array of strength 1. Using the Plotkin bound, it is easy to show that Latin hypercubes have maximum average distance and are equidistant. Thus, by Proposition 1, they are universally optimal with respect to Hamming distance. Moreover, saturated orthogonal arrays of strength 2, $OA(n,s,q,2)$, are equidistant with respect to Hamming distance (see Lemma 1 of Mukerjee and Wu (1995)). Thus, by the Plotkin bound and Proposition 1, they are universally optimal with respect to Hamming distance.

C. Lee distance

Again, the Plotkin bound shows that, for any $n \times s$ matrix X over F_q, $d_{ave}(X) \leq (q^2-1)ns/[4q(n-1)]$ for odd q and $d_{ave}(X) \leq qns/[4(n-1)]$ for even q. Equality holds if X is an orthogonal array of strength 1. Let A be a Latin hypercube $OA(n = q, s, q, 1)$ or a saturated $OA(n, s, q, 2)$. Use of the Plotkin bound shows that A has maximum average distance. Consider a new design, $B = [A, 2A, \ldots, ((q-1)/2)A] \pmod{q}$. If q is an odd prime number, it can be shown that this $n \times ((q-1)s/2)$ matrix B is equidistant and thus universally optimal with respect to Lee distance over F_q. In particular, there are universally optimal designs over F_q with $n = q^k$ and $s = (n-1)/2$ for any odd prime number q and positive integer k.

Next let A be a saturated $OA(n, n-1, 2, 2)$. Clearly A is universally optimal with respect to Lee distance over F_2 since Lee distance is the same as Hamming distance over F_2. Consider a new design, $B = kA$. It is easy to show that the $n \times (n-1)$ matrix B is equidistant and of maximum average distance and thus universally optimal with respect to Lee distance over F_q with $q = 2k$, k a positive integer. In particular, there are universally optimal designs over F_q with $q = 2k$, $n = 2^l$ and $s = n-1$ for any positive integers k and l. Note that these designs only have two levels.

There is another class of universally optimal designs for $q = 2^k$, k a positive integer. Define a $2^k \times (2^k-1)$ matrix $A = (a_{ij})$ by $a_{ij} \equiv i \times j \pmod{2^k}$, $1 \leq i \leq 2^k, 1 \leq j \leq 2^k - 1$. This design has 2^{k-i} columns of 2^{k-i+1} levels, $i = 1, \ldots, k$. Satyanarayana (1979) pointed out that this design is equidistant and achieves the Plotkin bound. Thus, by Proposition 1, it is universally optimal with respect to Lee distance.

4. Conclusion

It is shown that saturated two-level orthogonal arrays of strength 2 are universally optimal with respect to both rectangular and Euclidean distances. Latin hypercubes and saturated orthogonal arrays of strength 2 are shown to be universally optimal with respect to Hamming distance. Universally optimal designs with respect to Lee distance are also derived from Latin hypercubes and saturated orthogonal arrays.

Although only examples of designs on F_q^s are given in this paper, the generalization to $F_{q_1} \times \cdots \times F_{q_s}$ is straightforward. The application can also handle continuous design spaces. For example, Morris and Mitchell (1995) pointed out that saturated two-level orthogonal arrays of strength 2 are maximin distance designs in the unit cube with respect to both rectangular and Euclidean distances. It is easy to show that they are universally optimal in the unit cube with respect to both rectangular and Euclidean distances.

Acknowledgements

The author wishes to thank Professors C. F. Jeff Wu, Ching-Shui Cheng, Shiyi Shen, Kai-Tai Fang and the referees for their help and constructive suggestions.

References

Berlekamp, E. R. (1968). *Algebraic Coding Theory*. McGraw-Hill, New York.

Currin, C., Mitchell, T., Morris, M. and Ylvisaker, D. (1991). Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *J. Amer. Statist. Assoc.* **86**, 953-963.

Johnson, M. E., Moore, L. M. and Ylvisaker, D. (1990). Minimax and maximin distance designs. *J. Statist. Plann. Inference* **26**, 131-148.

Kiefer, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (Edited by J. N. Srivastava), 333-353. North-Holland, Amsterdam.

Martin, R. J. (1982). Some aspects of experimental design and analysis when errors are correlated. *Biometrika* **69**, 597-612.

McKay, M. D., Beckman, R. J. and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239-245.

Morris, M. D. and Mitchell, T. J. (1995). Exploratory designs for computational experiments. *J. Statist. Plann. Inference* **43**, 381-402.

Mukerjee, R. and Wu, C. F. J. (1995). On the existence of saturated and nearly saturated asymmetrical orthogonal arrays. *Ann. Statist.* **23**, 2102-2115.

Owen, A. B. (1992). Orthogonal arrays for computer experiments, integration and visualization. *Statist. Sinica* **2**, 439-452.

Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P. (1989). Design and analysis of computer experiments (with discussion). *Statist. Sci.* **4**, 409-435.

Satyanarayana, C. (1979). Lee metric codes over integer residue rings. *IEEE Trans. Inform. Theory* **IT-25**, 250-254.

Shewry, M. C. and Wynn, H. P. (1987). Maximum entropy sampling. *J. Appl. Statist.* **14**, 165-170.

Tang, B. (1993). Orthogonal array-based Latin hypercubes. *J. Amer. Statist. Assoc.* **88**, 1392-1397.

Department of Statistics, University of Michigan, Ann Arbor, MI 48109-1027, U.S.A.
E-mail: hqxu@umich.edu

(Received August 1996; accepted December 1998)