The Korean Journal of Applied Statistics (2012) DOI: http://dx.doi.org/10.5351/KJAS.2012.25.3.485
25*(3), 485–501*

Empirical Optimality of Coverage Design Criteria for Space-Filling Designs

Jungmin Baik[1]

[1]Samsung Life Insurance

(Received April 19, 2012; Revised May 3, 2012; Accepted May 9, 2012)

Abstract

This research is to find a design D that minimizes forecast variance in d dimensions over the candidate space $\mathcal{X}$. We want a robust design since we may not know the specific covariance structure. We seek a design that minimizes a coverage criterion and hope that this design will provide a small forecast variance even if the covariance structure is unobservable. The details of an exchange or swapping algorithm and several properties of the parameters of coverage criterion with the unknown correlation structures are discussed.

Keywords: Coverage criteria, distance-based designs, forecast variance, swapping algorithm.

1. Introduction

Recognizing and quantifying the relationship between a dependent variable and independent variables are the two primary goals of statistical experimental design. The construction and analysis of designs, therefore, are two main tasks of the experimental design. The first task focuses on selecting particular sites on which to run an experiment and collect related data. A site is a specific combination of descriptive values of the independent variables. In a given experimental setting, experimenters may have many choices of sites or design along with many objectives. Several important aspects of the experiment go into the choice of design, such as the costs of experiment, the number of sites to be chosen, and the information to be gained through the analysis of the design. The second task involves the application of statistical methods to data from the experiment. The subsequent statistical analysis of a well designed experiment should provide answers to the questions that motivated the experiment. Consequently, the first task of designing the experiment is fundamentally important.

One of the basic experimental design problems is the choice of inputs or sites for an efficient analysis of the data, the "selection-of-sites" problem. Often, when the goal of the experiment is to predict the response variable for particular values of the independent variables, the experimenter needs to find the set of sites of the independent variables that provide the smallest forecast variance from the response variable. Usually, choosing inputs from a continuous space is difficult and often computationally intractable.

Full version of this research was submitted by the author as a PhD dissertation at North Carolina State University at Raleigh in 1999.

[1]Ph.D, Former Samsung Life Insurance Company head of CRM team, E-mail: jungmin.baik@gmail.com

The solutions of traditional experimental design problems often have strict requirements such as the number of factors, the number of levels per factors, and the number of sites in the design. As a remedy, various optimality criteria have been introduced to evaluate the efficiency of designs, quantifying the cost of the experiment and the information gained.

A typical approach for statistical optimal experimental design is to select a particular criterion that measures the merit of a particular design and then optimize this criterion over all possible choices of designs. In particular, the variance of the parameter estimates and predictions depend upon the experimental design; these variances should be as small as possible. The optimality criteria typically used include A-, D-, and G-optimality (see Atkinson and Donev, 1992). A-optimality finds a design that minimizes the average variance of the parameter estimates for a specific parametric model. D-optimality finds a design that minimizes the determinant of the variance-covariance matrix of the parameter estimates of the model. G-optimality finds a design that minimizes the maximum variance of the fitted response over the design region. The usual optimality criteria depend on the model and can be formulated as a function of the variance-covariance matrix of the parameter estimates.

In some fields of science and engineering, computer experiments are used when traditional or physical experiments are not reasonable to perform, such as when it is overly time consuming, expensive, impractical or fundamentally impossible. The primary objective of computer experiments is to estimate the relationship between dependent variables and independent variables from a moderate number of runs so that independent variables can be predicted at untried inputs. In a computer experiment, observations are generated by running a computer code with certain inputs and recording the response. The computer output or response is often deterministic of rerunning the computer code with the same inputs that provide exactly the same results. This lack of random error associated with the observed variable can make computer experiments different from classical physical experiments. Deterministic experiments also differ from simulation experiments (Kleijnen, 1987), since simulations incorporate substantial random error through random number generators or randomization methods. Sacks *et al.* (1989) reviewed recent work on the design and analysis of computer experiments.

Because of the lack of random or replication error, computer experiments do not require the general concepts of an experimental unit; in addition, the lack of random error allows the complexity of the computer model to emerge. For these reasons, the "space-filling" design has been considered. Their construction does not depend on a specific model for the data. Similarly, with the model-based designs for traditional experiments, the experimenters expect to find the model-free space-filling designs for computer experiments.

Two approaches to construct space-filling designs have been proposed. One family of a design criteria are dependent on the specific correlation structure as suggested by Sacks *et al.* (1989) versus another design criteria independent of the assumed covariance function has been proposed based on geometric measures. In this context, the minimax and maximin distance-based designs were introduced (Johnson *et al.*, 1990). Specifically, the minimax designs minimize the maximum of nearest-neighbor distances from the candidate points to those in the design. The minimax designs are ideal space-filling designs based on geometric measures; however, these designs are impractical since the criterion is not well behaved and the minimax design is difficult to obtain even though the criterion itself is easy to compute.

The "coverage" space-filling designs were introduced by Tobias (1995) as a surrogate for the minimax

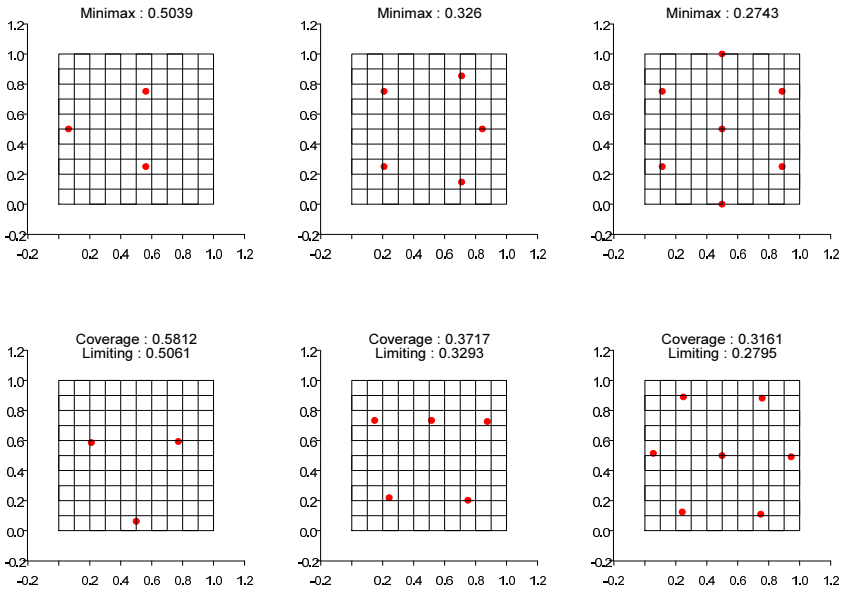

Figure 1.1. Actual Minimax designs and Coverage designs

design. These designs geometrically measure how well a given design covers the space of interest. Coverage designs are easily generated using an "exchange" or "swapping" algorithm. Figure 1.1 shows simple examples for $n = 3, 5, 7$ sites of the minimax designs on a rectangular grid and the final converging coverage designs for the swapping algorithm. The designs in the first row are the actual minimax designs from Johnson *et al.* (1990) and the designs in the second row are coverage designs generated on $2^7 \times 2^7$ grid of candidate points on the unit square. The results of the second row are from the best designs from 5 independent random starts of a swapping search algorithm (with 2,500 nearest-neighbors).

Two different criteria values are given for each of the coverage designs in the second row. One is the coverage criterion with the finite values of parameters labeled as 'Coverage', the other criterion labeled as 'Limiting' is the coverage criterion, which is evaluated at the limiting values of parameters. The minimax criterion is not comparable to the coverage criterion directly, since the coverage criterion in Figure 1.1 calculates the discrete candidate set. Therefore, we consider the limiting coverage criterion as the surrogate of the minimax criterion on the continuum space. Notice that the differences between minimax and limiting coverage are quite small (smaller than .01).

2. Distance-Based Space-Filling Designs and Criteria

Computer experiments are based on models of scientific phenomena and employed across broad areas of science. For example, Currin *et al.* (1991), and Sacks *et al.* (1989) give examples in the field of the electronic circuits. Sacks *et al.* (1989) present applications to chemical kinetics problems. The prediction of the computer output Y for an untried input is treated by modeling the systematic departure of Y from a linear model as a realization of a stochastic process. The

examples listed above show how approaching the problem from a stochastic process model could aid in the computation of efficient designs and suggest several design criteria. The previously mentioned papers suggest several criteria that depend on the covariance structure of the stochastic process. We call the experimental designs that satisfy such criteria covariance-based designs.

The suggested covariance-based criteria for design selection are not appropriate when we do not have much prior information or knowledge on which to base the selection of the covariance structure. An alternative method of choosing a design is to use a geometric criterion that does not depend on the covariance structure. Johnson *et al.* (1990), cited here as JMY, develop the notions of minimax and maximin designs for an arbitrary design set from the space of interest along with a certain distance measure defined on this space.

Minimax distance designs have asymptotically optimal properties when the correlation on the region of interest is a decreasing function of distance. Unfortunately it is often difficult to generate such a design because the minimax criterion is not well behaved, that is, the minimax criterion has many local minima and maxima. In addition, it behaves like a saddle point. Thus, the minimax design is an ideal but possibly impractical distance-based optimal design.

As an approximation, Tobias (1995) discussed "space-filling" coverage designs, where we minimize a criterion that is only a function of the distance between candidate points and design points. The coverage criterion can be interpreted as a measure of how well a set of design point covers the domain of interest. It has the advantage that designs can be easily found using an exchange algorithm. Nychka *et al.* (1998) showed that the coverage design approximately minimizes the average prediction variance.

"Space-Filling" designs are appealing in situations where the experimenter may be unwilling to specify the functional form of the response function but is willing to assume that the surface of prediction is "smooth enough" to justify interpolation (Doehlert, 1970).

Let $\mathcal{C}$ denote a set of possible sites in $\mathcal{X}$ and suppose there is a nonnegative function d on $\mathcal{C} \times \mathcal{C}$ so that

$$d(u, v) = d(v, u), \quad \text{for all } u, v \text{ in } \mathcal{C} \tag{2.1}$$

$$d(u, v) \geq 0 \quad \text{with equality if and only if } u = v. \tag{2.2}$$

Furthermore, if one allows that

$$d(u, v) \leq d(u, t) + d(t, v), \quad \text{for all } u, v \text{ and } t \text{ in } \mathcal{C}, \tag{2.3}$$

then $(\mathcal{C}, d)$ is a metric space.

2.1. Minimax design

Let $d(\cdot, \cdot)$ be a metric on $\mathcal{C} = [0, 1]^p$.

Definition 2.1. *Let $\mathcal{C}$ be a linear metric space. Consider subsets $D \subseteq \mathcal{C}$ with card$(D) = n$, n fixed. Call D^* a minimax distance design if*

$$\min_{D} \max_{x \in \mathcal{C}} d(x, D) = \max_{x \in \mathcal{C}} d(x, D^*) = d^*$$

where $d(x, D) = \min_{u \in D} d(x, u)$

The intuition behind the minimax criterion is that minimax distance designs ensure that all points in $\mathcal{C}$ are not "too" far from the subset D^*. This follows because the minimax criterion attempts to cover the experimental space by locating design points to minimize the maximum distance from any candidate point to the closest design point. Another way of viewing the minimax design is to consider placing a p-dimensional sphere with radius r around each design point. The idea of a minimax design is to place the n points so that the design space is covered by the spheres with a minimal radius.

One illustration of this problem for the two dimensional case may be called the franchise problem. We can think of the design points D as store locations and the candidate set $\mathcal{C}$ as customer locations. From a customer point of view, we would like to locate the stores closest to the customers. A minimax plan of placing fast food stores would guarantee that no customer is too far from one of the franchise stores. The minimax criterion is in fact the largest distance that a customer is from his closest store. For small n, the minimax distance-based designs will generally lie in the interior of the design space.

JMY derived the asymptotic optimality properties of the minimax distance designs. The minimax designs have the asymptotic G-optimality property when the correlation in the space of interest is a decreasing function of distance. JMY showed that minimax designs based on distance also minimize the maximum prediction error when the correlations tend toward zero, that is, independence. The above asymptotic optimality property is proposed by JMY as the Theorem along with more details of the proof. In essence, the minimax designs have an asymptotic optimality when the space of interest has independency.

2.2. Coverage design

Because of difficulties of finding the minimax designs as mentioned before, we consider an approximate geometric criterion to locate design points whose solution may approach optimality. Space-filling "coverage" designs are spatial sampling plans that optimize a distance-based criterion. It is assumed that the designs are the solution to the discrete problem of selecting a subset of points from a larger finite set of candidates. Let $\mathcal{C}$ denote the set of N candidate points and D a subset constituting a design of size n. One measure for the distance between a candidate point $\mathbf{x}$ and a particular design D is

$$d_p(\mathbf{x}, D) = \left(\sum_{\mathbf{u} \in D} ||\mathbf{x} - \mathbf{u}||^p \right)^{\frac{1}{p}}. \tag{2.4}$$

This metric-type measure (2.4) can be thought of as measuring how well the design covers point $\mathbf{x}$.

Definition 2.2. *Let D and $\mathcal{C}$ be a design space and candidate space, respectively. A coverage design satisfies the minimum of*

$$C_{p,q}(D) = \left\{ \sum_{\mathbf{x} \in \mathcal{C}} [d_p(\mathbf{x}, D)]^q \right\}^{\frac{1}{q}},$$

where $p < 0$ and $q > 0$.

Taking $q > 0$, an overall coverage criterion is an L_q average of "coverages" for each candidate point. There are two parameters in coverage criterion, p and q. We can think of p as how relatively close a

design point has to be to affect $d_p(\mathbf{x}, D)$. We set the coverage parameters as $p < 0$ and $q > 0$. Since if we let $p \geq 0$ or $q \leq 0$ then the solution is a trivial design where all points equal the centroid.

With the following Lemma 2.1 and Lemma 2.2, we see that the minimax designs have connections with covering problems similar to coverage designs.

Lemma 2.1. *Let $\mathbf{x} \in \mathcal{C}$. If $p < 0$, then $d_p(\mathbf{x}, D) \to 0$ as $\mathbf{x}$ converges to a member of D.*

Lemma 2.2. *If $p \to -\infty$, then the metric $d_p(\mathbf{x}, D)$ is the minimum distance between the candidate point and the design points.*

Apparently, the coverage criterion measures how close selected design points are to candidate points. We can summarize the behavior of the coverage criterion in the following Theorems.

Theorem 2.1. *The coverage criterion based on the geometric distance converges to the minimax distance design criterion if $p \to -\infty$ and $q \to \infty$.*

By Theorem 2.1, the coverage criterion with $p = -\infty$ and $q = \infty$ is equal to the minimax criterion that minimizes the maximum of nearest-neighbor distances among points in the candidate set to those in the design. Therefore, one can view a coverage design, with relatively big number of finite p and q, as a good surrogate of the minimax design.

Theorem 2.2. *If $p \to -\infty$ and $q = 1$, then the coverage design criterion converges to some design criterion that has asymptotic A-optimal properties when the correlation function satisfies the assumptions of the JMY Theorem.*

Because the coverage designs do not depend on the covariance structure of the process to be sampled, coverage designs are more efficiently computed than designs that are optimal for mean squared error criteria. By using a simple "swapping" or "exchange" algorithm, the coverage designs are easily generated.

2.3. Maximin and spread designs

Definition 2.3. *Let $D \subseteq \mathcal{C}$ with card$(D) = n$, n fixed. Call D° a maximin distance design if*

$$\max_{D} \min_{u,u' \in D} d(u, u') = \min_{u,u' \in D^o} d(u, u') = d^\circ.$$

Let the *index* I° be the number of pairs of sites in D° separated by distance d°, and $D^{\circ\circ}$ denote any maximin set with the smallest index. For simple illustration, let $\mathcal{C} = [0, 1]$, card$(D) = n$, then it is easy to see that the maximin set assigns an equal spacing to n points that include 0 and 1.

The maximin design criterion tries to spread the design points in space to maximize the minimum distance between the pairs of design points. We can think of a maximin design as a packing design since these designs pack the n design points, with their associated spheres with maximum radius, into the design space. Parts of the sphere may be out of D but the design points must be in D. Analogous to the minimax illustration above is the desire of the owners of fast food stores to place them as far apart as possible. A maximin plan for placing the franchises would ensure that no two stores are too close to each other and will not compete with each other.

Tobias (1995) also discussed of another kind of space-filling design criterion or "spread" design. For $u \in D$, $d_p(u, D - \{u\})$ measures how far the design point u is from the rest of the design points.

Definition 2.4. *Let D be a design space. A spread design satisfies the minimum of*

$$S_{p,q}(D) = \left(\sum_{u \in D} (d_p(u, D - \{u\}))^q \right)^{\frac{1}{q}}$$

where $p < 0$ and $q < 0$.

The spread criterion, therefore, is the measure of how close selected design points are to each other. Tobias (1995) describes several properties of a spread design. One of the main properties relates the spread design to the JMY maximin design: if we let $p \to -\infty$ and $q \to -\infty$, the spread design converges to the maximin design.

3. The Point Swapping (Exchange) Algorithm and Details of Coverage Criterion

3.1. The swapping algorithm

The suggested algorithm for finding the optimum designs based on the coverage criterion uses a random starting configuration and decreases the coverage criterion by swapping or exchange a candidate point with a design point. The references for this algorithm are Kennard and Stone (1969), Mitchell (1974) and in connection to distance-based criteria Marengo and Todeschini (1992) and Tobias (1995). Several different kinds of the exchange algorithms are reviewed in Nguyen and Miller (1992), that concentrates on finding discrete D-optimal designs.

The swapping algorithm may be summarized as follows: Let $\mathbf{x}_i$ and $\mathbf{u}_j$ be candidate points and design points, respectively.

(1) Select a random starting design, and compute the vector of initial row sums, $\mathbf{r}$ of the distance matrix and $C_{p,q}(D)$.

(2) For each design point, replace $\mathbf{u}_j$ by $\mathbf{x}_i$, $i = 1, \ldots, N$ and compute N criterion values.

(3) Swap $\mathbf{x}_i$ with the $\mathbf{u}_j$ that produces the largest decrease over the initial criterion. Given this swap recalculate $\mathbf{r}$ and $C_{p,q}(D)$.

(4) Repeat (2)–(3) until no swap can be made.

When the design point is updated, the coverage criterion will always be reduced since the swapping is made only if it reduces the criterion. Since the algorithm stops when no swap gives an improvement, the solution may only be a local optimum. Consequently, it is very important to use several different random starts to improve the chances of finding the global optimum. The more details of the swapping step and the flow chart are in the dissertation paper (Baik, 1999)

3.2. Details of coverage criterion

In the coverage criterion of the design D, $C_{p,q}(D)$, we do not know how large the parameters p and q must be to give a near optimal design. We might hope that moderate number of values will provide a relatively good design. Furthermore, we want to know how many runs (random starts) do we need to obtain a design near the optimum. For answering these questions, we ran a simple simulation. We selected 25 design points from 51×51 (2,601) grid candidate points on the unit square, $\mathcal{C} = [0, 1]^2$, with the number of nearest-neighbors set to 25% of the candidate set. We use the 5×5 grid design for comparison, having minimax criterion value $\sqrt{2}/10 = 0.1414$ for this specific

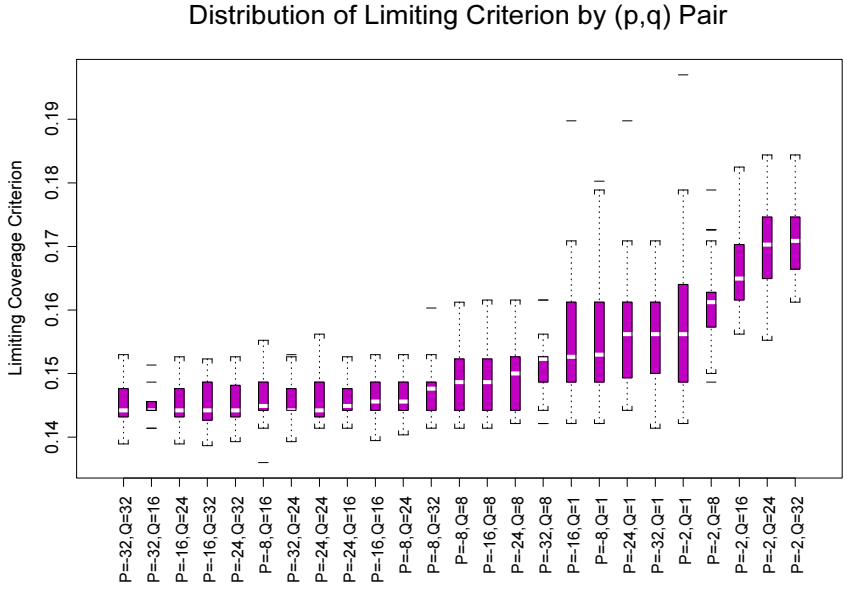

Figure 3.1. Box-plots of 100 random runs at various (p, q) pairs

candidate set, since we do not know the exact minimax 25 point design and the 25 grid point design is known to be the A-optimal design. Using the 5×5 grid points and the three different coverage designs that all have the same criterion values of the grid points to 4 decimal places. These three designs illustrate three different random starts with $p = -32$ and $q = 32$.

Theoretically, larger values of $-p$ and q should lead to better designs. Practically, the size of $-p$ and q are limited for computational reasons. One of the benefits of coverage design over the minimax design is the stability of criterion function. While the minimax criterion is as easy to compute as the coverage criterion, finding the optimal solution is not easy since local optima abound. With finite values of coverage criterion parameter p and q, we are looking for a near optimal design instead of the minimax design. From Theorem 2.1, for large $-p$ and q values we expect that the coverage design is nearly the same as the minimax design that has them asymptotically G-optimal property under certain conditions of the correlation function. Also by Tobias (1995), large $-p$ values with $q = 1$ we can achieve a nearly optimal design. However, the question is how large the parameter values should be to be nearly optimal.

For the various parameter values we use combinations of $-p = \{2, 8, 16, 24, 32\}$ and $q = \{1, 8, 16, 24, 32\}$. We then ran 100 coverage designs from random starts at each (p, q) pair. Figure 3.1 is a box plot of the 100 minimax criteria (limiting coverage criteria) values of the coverage designs at each combination of parameters. From Figure 3.1, we can see that the minimax criteria values from coverage searches with small $-p$ values have more variance than the large $-p$ values. Interestingly, when the q fixed at 1, p values do not seriously influence the criterion function. But small $-p$ values definitely are not a good choices when $q \neq 1$. The functional variability of the criterion is relatively stable around large $-p$ and q values. Consequently, relatively large $-p$ and q values are good choices.

If we take $-p = q$, and q to be a power of 2, then we can remarkably reduce the calculation time of

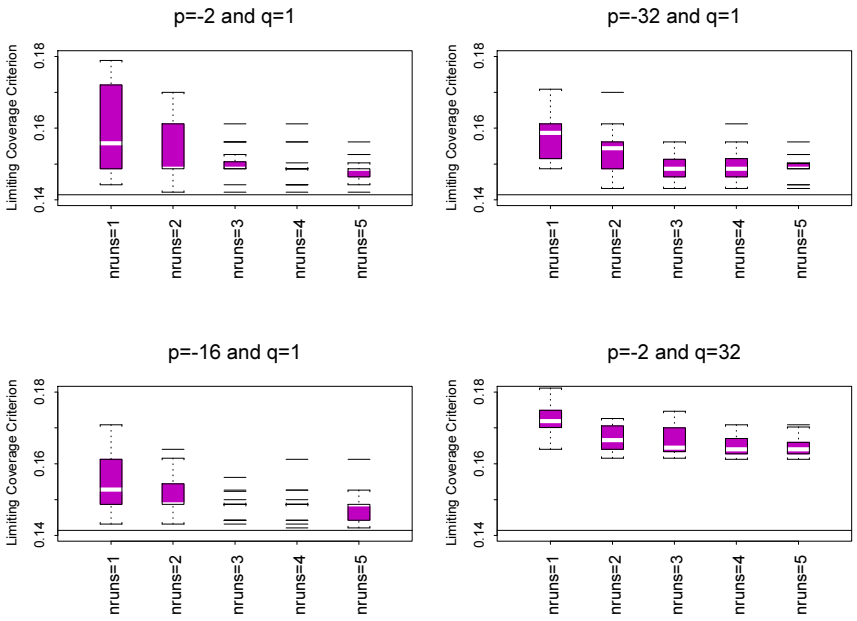

Figure 3.2. Box-plots of the best of the number of runs at several (p, q) pairs

the coverage criterion. Since if p and q have the same values except the sign, the calculation time is cut by removing the $(q/p) = 1$ power inside the coverage criterion (2.2), therefore, we recommend values of (p, q) of $(-16, 16)$ or $(-32, 32)$.

3.3. Rule of thumb for selecting the number of runs

As mentioned before, we do not expect to have a good design with a single random starting design. Therefore, we try to run with several different random starts and choose the best design out of these solutions. For this reason, we want to know the number of runs that are likely to produce a good design. Figure 3.2 and Figure 3.3 show box plots of minimum criterion values for several numbers of runs at several parameter choices. If we choose good (p, q) values, we need only a small number of runs to have a fairly good design.

From Figure 3.2, the cases with $q = 1$ are A-optimal type designs. In these cases, the small $-p$ values have larger variance and bigger median value of criterion than the large $-p$ values through all number of runs. The case with small $-p$ and large q has a large variance and median value of criterion through all number of runs.

From Figure 3.3, the variances and median values of criterion are all smaller than those of Figure 3.2 for up to 5 number of runs. This means that the best of k runs with suggested pairs, $(-32, 32)$ or $(-16, 16)$, has same performance as best of l runs with other pairs (p, q) where $k \ll l$. Even within the suggested pair of cases, the larger $-p$ and q have the smaller variance and median values. Therefore, the difference numbers of runs between $(-32, 32)$ and $(-16, 16)$ for the same performance should be small. Consequently, the design with suggested pairs (p, q) needs only a small number of runs to create the same effect as other pairs design.

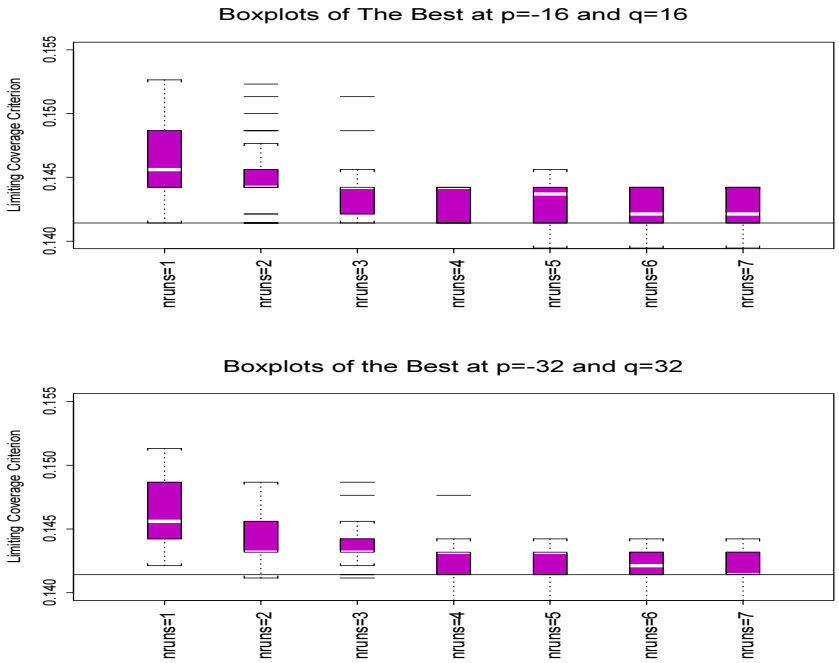

Figure 3.3. Box-plots of the best of the number of runs at suggested (p, q) pairs

4. Forecast Variance Designs

The experimenters usually notice that there exists some degree of relationship among the responses in the space of interest. Often, however, they cannot measure the magnitude or find any form of relationship. In these circumstances, our primary objective is finding designs, D, that minimize forecast variance over the candidate space $\mathcal{X}$. Since we do not have much information about the covariance structure, we want a design that would be robust over a variety of spatial covariance structures. We might hope that a design that covers the space well such as having a small value of the coverage criterion, $C_{p,q}(D)$, might also give small forecast variances for some values of spatial correlation.

In this context, we describe the minimized forecast variance designs and compare them with the original swapping coverage designs over the different levels of correlations. We use the cumulative distribution function of the best number of runs to compare these two designs.

Let us define the covariance structure between x and x' given parameters r and θ, $V(x, x'|r, \theta)$, as

$$V(x, x'|r, \theta) = \sigma^2 \cdot R(x, x'|r, \theta)$$

where we use a correlation structure based on distance

$$R(x, x'|r, \theta) = \exp\left\{-\frac{||x - x'||^r}{\theta}\right\}. \tag{4.1}$$

$R(x, x'|r, \theta)$ is the correlation between x and x', and σ^2 is the process variance. The above correlation decreases as θ decreases for fixed r. We use the above correlation structure throughout this discussion.

Table 4.1. Scheme of The Simulation

Candidate Set		Dimension	
		$d = 2$	$d = 4$
Coarse	C_1^d	$n = 5, 7$ (200)	$n = 5, 7$ (200)
Fine	C_2^d	$n = 5, 7$ (50)	

Note: The number of runs are provided in parentheses.

Let Σ_D be a covariance matrix of the discrete candidate space with the design D. It can be written as follows

$$\Sigma_D^{N\times N} = \begin{bmatrix} \Sigma_{11}^{n\times n} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

where Σ_{11}, Σ_{22} are the partitioned covariance matrices of the n design points and the $N - n$ remaining candidate points (excluding design points), respectively.

Definition 4.1. *A forecast variance design minimizes the average or maximum of the diagonal of the posterior variance, $PV(\Sigma_D)$,*

$$\text{Average Forecast Variance (AFV)} : \min_{D_i\in\mathcal{C}} \left[\frac{\sum_j PV_j(\Sigma_{D_i})}{(N-n)}\right]$$

$$\text{Maximum Forecast Variance (MFV)} : \min_{D_i\in\mathcal{C}} \left[\max_j PV_j(\Sigma_{D_i})\right]$$

where

$$PV_j(\Sigma_D) = \text{diag}_j\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right), \quad j = 1, \ldots, (N-n).$$

4.1. Comparing designs and analysis

With the given correlation structure (4.1), we are going to compare the forecast variance designs with the coverage designs. In simulation, we use two types of candidate sets (coarse and fine) and two kinds of dimensions, 2 ($C^2 = [0, 1]^2$) and 4 ($C^4 = [0, 1]^4$). The coarse set with $\text{card}(C_1^2) = 289$ and fine set with $\text{card}(C_2^2) = 4,225$ are used for 2 dimension case. For 4 dimension case, however, only the coarse candidate set, ($\text{card}(C_1^4) = 1,296$), is used since the fine candidate set is hardly manageable and impractical. Two different numbers of design points ($n = 5$ and $n = 7$) at each dimensions are used with a 25% nearest-neighbors search. In the covariance structure, two levels of r, ($r = 1$ and 2) and four levels of θ, ($\theta = 1, 0.75, 0.25$ and 0.1) are used.

First, we generate initial random designs on the specific candidate set and calculate the three different criteria, coverage, $C_{p,q}$ (for $p = -32$, $q = 32$), AFV, and MFV at various values of r and θ. Table 4.1 shows the scheme of these simulations.

Figure 4.1 and Figure 4.2 display the scatter plots of the $n = 5$ coverage designs and the two cases of the forecast variance designs with matched random designs for the various r and θ.

The top and bottom in each combination of the parameters, illustrate the swapping AFV and the MFV designs (triangle points) from the same random initial designs (circled points), respectively. The x-axes and y-axes represent the coverage and forecast variance criterion. Therefore, the coverage criteria, x-axes, of the random designs of the top and bottom are the same. From the random

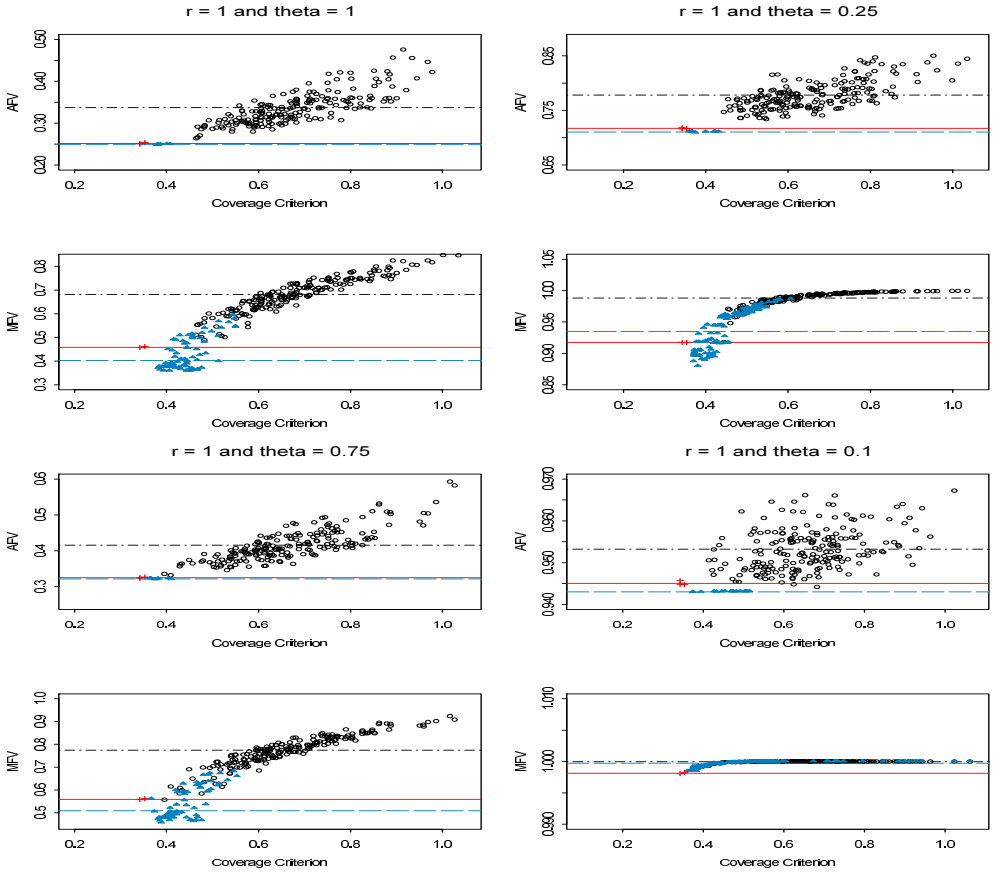

Figure 4.1. The coverage and forecast designs at $n = 5$ with $r = 1$ for coarse

designs (circled points), the converged swapping coverage design criteria (cross points) are plotted at both plots. Similar to the random designs, the swapping coverage criteria at the top and bottom plots are all the same. The horizontal lines represent the mean values of the each design group. The solid line, long broken line, and short dotted line represent the mean of the coverage, forecast variance, and random designs, respectively.

In these Figures, the swapping solutions are superior to random designs. The coverage design is correlated with the forecast variance designs. The correlation between coverage and forecast variance designs decreases as the spatial correlation decreases. The slopes or shapes are changed by r. Now we are interested in the mean of the differences between the swapping coverage designs and the swapping forecast variance designs for various r and θ.

Table 4.2 shows the means and standard errors of paired differences between the swapping coverage and forecast variance designs. Two means are given for each case. One column gives mean of the differences between average forecast variances evaluated for the coverage designs and the converged AFV designs. The other column gives mean of the differences between maximum forecast variances evaluated for the coverage designs and the final swapping MFV designs. From Table 4.2, the ranges

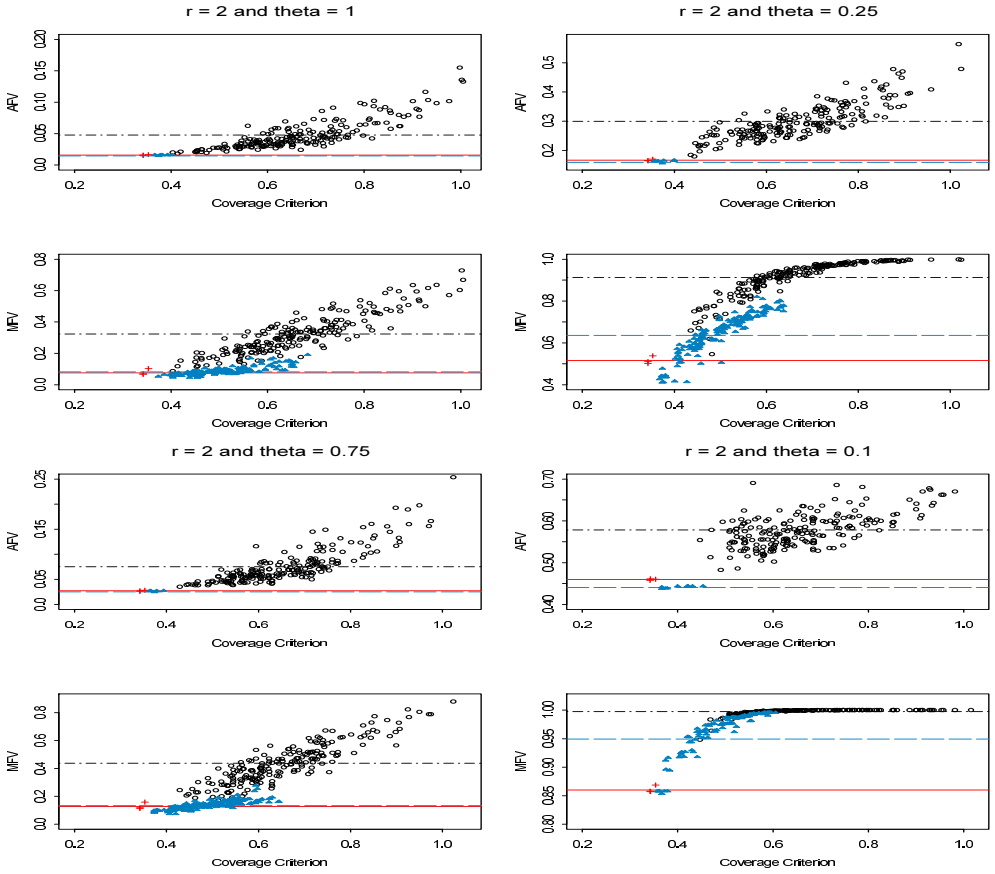

Figure 4.2. The coverage and forecast designs at $n = 5$ with $r = 2$ for coarse

of the absolute mean differences for the AFV designs are smaller than the MFV designs. The AFV designs have small standard errors for all θ values. The variances of the MFV designs are larger than the AFV designs. If the spatial correlation is large ($\theta \geqq 0.75$) then the MFV designs are better than the coverage designs. However, the MFV designs have more variance than the coverage designs. Even if the spatial correlation is small ($\theta \leqq 0.25$), the MFV designs have more variance and are higher than the coverage designs. The MFV designs have a large variance even in the plots from the fine set. Therefore, the swapping MFV designs are not a good choice of design criteria when there exists some level of correlation. In these situations, the coverage designs give more promising results than the MFV designs. These features are consistent throughout the fine candidate case. When we increase the dimension to 4 the above arguments still hold. Interestingly, correlations with the coverage and MFV designs begin to disappear as the spatial correlation decreases.

The mean differences between the coverage designs and the AFV designs are relatively small over the whole range of the parameters. These differences are statistically significant; however, not practically so since the differences are so small. For $d = 4$ case, the results are similar. However, the correlations between the coverage and the forecast variance designs begin to disappear more

Table 4.2. The mean of difference between coverage and forecast in 2 dimension

C_i^2	r	θ	$n=5$		$n=7$	
			AFV	MFV	AFV	MFV
Coarse 200 Runs	1	1	.0028 (.0018)	.0556 (.0518)	.0042 (.0030)	.0781 (.0392)
		.75	.0033 (.0019)	.0504 (.0575)	.0049 (.0036)	.0809 (.0415)
		.25	.0067 (.0012)	−.0175 (.0310)	.0010 (.0053)	.0038 (.0415)
		.1	.0020 (.0004)	−.0016 (.0004)	.0039 (.0021)	−.0029 (.0011)
	2	1	.0012 (.0008)	−.0058 (.0295)	.0016 (.0018)	.0108 (.0162)
		.75	.0019 (.0013)	−.0040 (.0373)	.0026 (.0024)	.0158 (.0266)
		.25	.0073 (.0051)	−.1200 (.0982)	.0123 (.0096)	−.0110 (.1096)
		.1	.0192 (.0026)	−.0891 (.0450)	.0253 (.0116)	−.0567 (.0934)
Fine 50 Runs	1	1	.0027 (.0013)	.0232 (.0314)	.0037 (.0037)	−.0074 (.0223)
		.75	.0027 (.0015)	.0171 (.0400)	.0046 (.0030)	−.0105 (.0225)
		.25	.0032 (.0011)	−.0199 (.0167)	.0094 (.0038)	−.0435 (.0115)
		.1	.0013 (.0002)	−.0013 (.0002)	.0052 (.0023)	−.0041 (.0005)
	2	1	.0011 (.0004)	.0135 (.0202)	.0015 (.0021)	.0080 (.0092)
		.75	.0014 (.0008)	.0089 (.0341)	.0024 (.0032)	.0128 (.0133)
		.25	.0063 (.0027)	.0819 (.0861)	.0106 (.0095)	−.0508 (.0435)
		.1	.0107 (.0034)	−.0332 (.0469)	.0265 (.0105)	−.1236 (.0377)

Note: The standard errors are in the parenthesis.

quickly than $d = 2$ case as the spatial correlation decreases. Therefore, the coverage designs and the AFV designs are relatively close for the whole range of θ. We cannot conclude that the coverage designs have the same effect as the AFV designs when there exists some level of correlation; however, we can compute the coverage designs more easily than the AFV designs.

Due to the nature of the swapping search algorithm and the random initial start, in practice a user may wish to take the best design of a few, say k, solutions. Because of the random initial design, the solution from a swapping algorithm can be viewed as a random sample from a (small) population of solutions, so that the best of k runs has the distribution of the minimum from an *iid* sample of size k. Suppose the best of k solutions from coverage swapping algorithm performs as well as the best of l solutions from forecast variance swapping. Then these two methods can be compared by estimating the ratio, $\beta = l/k$. Therefore, one swapping coverage solution has the same effect as well as the best β forecast variance search.

Let the cumulative distributions of the swapping AFV criteria, $F_x(t)$, and the swapping coverage solutions, $G_x(t)$, be a follows,

$$F(t) = \Pr(D_a \leqq t)$$
$$G(t) = \Pr(D_a \leqq t)$$

where D_a, is the criterion of solution with average forecast criterion. If we assume that the distributions of the two swapping methods (coverage and forecast variance) of designs are same at the best of k^{th} and l^{th} then we can establish the following equation,

$$1 - (1 - G(t))^k \approx 1 - (1 - F(t))^l$$
$$(1 - G(t)) \approx (1 - F(t))^{\frac{l}{k}}.$$

Taking the logarithm of both sides,

$$\ln(1 - G(t)) \approx \beta \cdot \ln(1 - F(t))$$

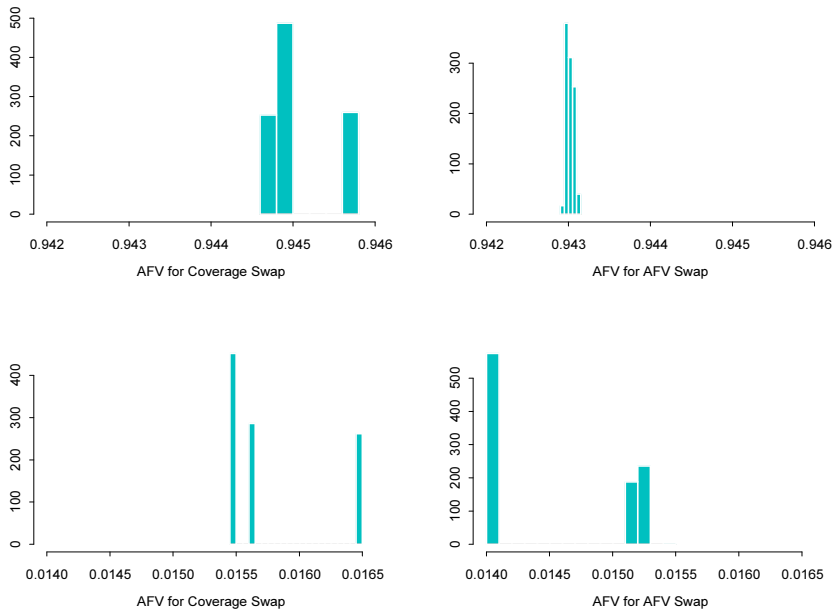

Figure 4.3. The histograms of coverage and AFV designs for $n = 5$

where $\beta = l/k$. To estimate β, from sample values from both F and G, the relevant values of t should be taken from the intersection of their supports. We can use then simple linear regression to estimate β.

To estimate β, we use 1,000 forecast variance designs for $\hat{F}(t)$ and 1,000 coverage designs for $\hat{G}(t)$. The candidate set, C_β, is 17×17 on the unit square ($[0,1]^2$) to calculate several criteria, such as coverage criterion, average and maximum forecast variance criteria. We will consider only two cases of the correlation structures such as low and high relationship. For given correlation structure (4.1), we use $r = 1$, $\theta = 0.1$ and $r = 2$, $\theta = 1$ for low, high correlation, respectively. In addition, we investigate both $n = 5$ and $n = 7$ designs.

Figure 4.3 and Figure 4.4 show the histograms of 1,000 runs of the swapping coverage designs and the AFV designs evaluated for the average forecast variance criteria. In Figure 4.3 and Figure 4.4, the first row is for the low correlation and the bottom row is for the high correlation. For each correlation structure, it shows the histograms of the average forecast variance criteria for the coverage swapping and AFV swapping designs. In the situation with the low correlation, the coverage design and the AFV design distributions do not overlap; however, in the same cases with a high correlation, they share some common t values. With the high correlation for $n = 7$ designs, the β for the AFV criteria between the coverage swapping and the AFV swapping is 3.51. This means that one AFV design has a same effect as well as 3.5 coverage designs for the AFV criterion points of view. However, the convergence time for AFV design is almost 4 times slower than for the coverage design.

In the practical situation, with the huge candidate and design set, AFV designs take significantly longer to compute than coverage designs. Therefore, the swapping coverage designs are a robust design for minimizing the forecast variance when there exists an unobservable covariance structure.

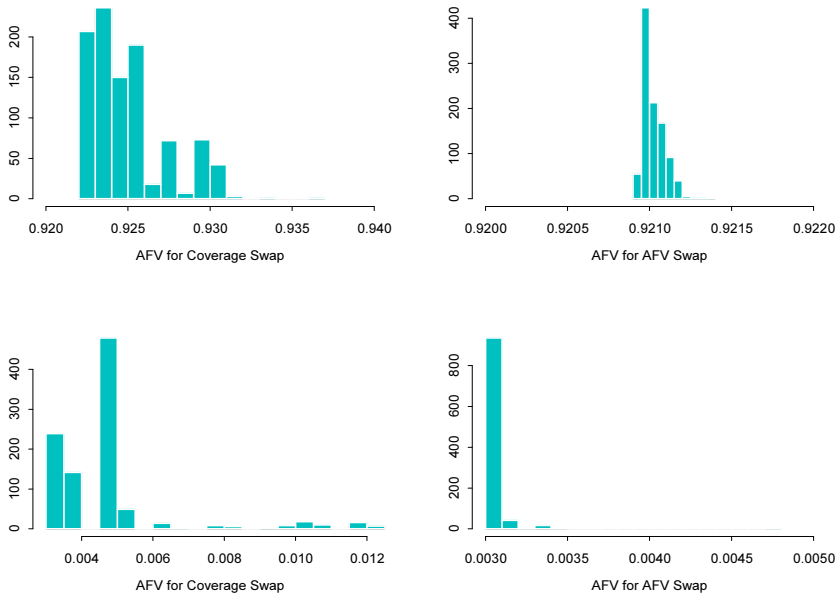

Figure 4.4. The histograms of coverage and AFV designs for $n = 7$

Even if the covariance structure were known, the best of a few number of coverage designs will outperform an AFV design and are faster to compute.

5. Concluding Remarks

This research enables us to generate robust designs when there exist some level of spatial correlation. The swapping algorithm is effective in finding good coverage designs. From the study of the parameters (p, q) of coverage criterion connection with, we show that the pairs of $(-16, 16)$ and $(-32, 32)$ provide relatively good designs. Furthermore, with those pairs, we can reduce the number of random starting runs to achieve the same effectiveness of the designs.

This work has empirically shown that coverage space-filling designs can be used for problems with spatial correlation. Interestingly, the average forecast variance designs and the coverage designs perform similarly in minimizing the average forecast variance through the low and high correlation structures.

We must mention at this point about the degree of penalty of the swapping algorithm due to discretization, comparing numerically to the Newton-type (quasi-Newton) algorithm by using the gradient and Hessian of the coverage criterion. We cannot run a Newton-type algorithm starting from just any random location because if the starting solution is not near a local minimum then it may lead to a stationary point other than a minimum. A sensible scheme is to run the swapping algorithm first and then compute the Hessian and gradient at the design of the solution points to see if the swapping solution might be close to a local or global minimum in the continuum. The more detail about these topics and the results of the practical situation (especially for the typical chemical library) are in Baik (1999).

References

Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*, Oxford Science Publications, New York.

Baik, J. (1999). *Optimization of Distance-Based Criteria for Space-Filling Designs*, Thesis of North Carolina State University, Raleigh.

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, *Journal of the American Statistical Association*, **86**, 953–963.

Doehlert, D. H. (1970). Uniform shell designs, *Journal of the Royal Statistical Society, Series C*, **19**, 231–239.

Johnson, M. E., Moore, L. M. and Ylvisaker, D. (1990). Minimax and maximin distance designs, *Journal of Statistical Planning and Inference*, **26**, 131–148.

Kennard, R. W. and Stone, L. A. (1969). Computer aided design of experiments, *Technometrics*, **11**, 137–148.

Kleijnen, J. P. C. (1987). *Statistical Tools for Simulation Practitioners*, Dekker, New York.

Marengo, E. and Todeschini, R. (1992). A new algorithm for optimal distance-based experimental design, *Chemometrics and Intelligent Laboratory Systems*, **16**, 37–44.

Mitchell, T. J. (1974). An algorithm for the construction of "D-optimal" experimental designs, *Technometrics*, **16**, 203–210.

Nguyen, N.-K. and Miller, A. J. (1992). A review of some exchange algorithms for constructing discrete D-optimal designs, *Computational Statistics & Data Analysis*, **14**, 489–498.

Nychka, D., Piegorsch, W. and Cox, L. (1998). *Case Studies in Environmental Statistics*, Lecture Notes in Statistics, **132**, Springer Verlag, New York.

Sacks, J., Schiller, S. and Welch, W. (1989). Designs for computer experiments, *Technometrics*, **31**, 41–47.

Tobias, R. (1995). SAS QC Software. Volume 2: Usage and Reference. SAS Institute, Inc., Cary, NC.