行政院國家科學委員會專題研究計畫 成果報告

凸面區域之均勻設計及其在電腦實驗上之應用
研究成果報告(精簡版)

計 畫 類 別 ： 個別型
計 畫 編 號 ： NSC 98-2118-M-004-007-
執 行 期 間 ： 98 年 08 月 01 日至 99 年 10 月 31 日
執 行 單 位 ： 國立政治大學統計學系

計 畫 主 持 人 ： 洪英超

計畫參與人員： 碩士班研究生-兼任助理人員：黃國倫
碩士班研究生-兼任助理人員：吳亞倫
碩士班研究生-兼任助理人員：謝至芬
碩士班研究生-兼任助理人員：呂雨築
博士班研究生-兼任助理人員：莊世鐘

報 告 附 件 ： 出席國際會議研究心得報告及發表論文

處 理 方 式 ： 本計畫可公開查詢

中 華 民 國 100 年 01 月 06 日

研究成果報告(精簡版)

目錄

報告內容

（一）前言

We first illustrate two examples from stochastic processing networks that are used to motivate this research.

Motivating Example 1: Consider a queueing system comprised of K infinite capacity first-in-first-out (FIFO) queues in parallel, and each queue corresponds to a different class of job traffic. The class k jobs arrive according to some process with mean rate x_k and are queued up for service, $k = 1, \ldots, K$. At any point in time, the system can be in one of S service modes. When service mode s is used, the first job of queue k receives service with mean rate μ_{sk}, $k = 1, \ldots, K$. Therefore, mode s is associated with the service rate vector $U_s = (\mu_{s1}, ..., \mu_{sK})$, $s = 1, \ldots, S$. This queueing system is known as the Switched Processing System (SPS), which captures the essence of a fundamental resource allocation problem in many modern systems involving heterogeneous processors and multiple classes of job traffic flows (e.g., parallel computing, wireless networking, call centers, flexible manufacturing, etc).

It has been shown in literature (Armony and Bambos 2003, Hung and Michailidis 2008) that, in order to achieve system stability (e.g. the long-term input rate for each queue is equivalent to its long-term service rate), the input rate vector $x = (x_1, \ldots, x_K)$ must lie within the following region:

$$D = \left\{ x \in \mathbb{R}_+^K : x_k < \sum_{s=1}^{S} \omega_s \mu_{sk}, \text{ for all } k = 1, \ldots, K \right\},$$

where $0 \leq \omega_s \leq 1$ for all $s = 1, \ldots, S$, and $\sum_{s=1}^{S} \omega_s = 1$. Note that the linear constraints in D describe that the long-term input rate of each queue k can not exceed its long-term service rate. Further, it can be shown that D is a *convex hull* generated by all service rate vectors U_s and their projections on the axes. An example of such D is shown in the left panel of Figure 1. In addition, a service-mode allocation policy is called a throughput-maximizing policy if it can stabilize the system for all input rate vectors $x \in D$. A natural question followed is, under a particular throughput-maximizing policy π, how the input rate vector x affects the performance measures of interest (such as delay, backlog, etc). In this case, $x_1, \ldots, x_K$ are treated as input factors, while the input domain D is a convex polygon.

Motivating Example 2: Let's consider the same queueing system introduced in Example 1. Suppose now all the input rates are fixed (but still inside the region D), and we consider a throughput-maximizing policy called the "MaxProduct" (Armony and Bambos 2003, Hung and Chang 2008, Hung and Michailidis 2008). This policy employs service mode s^* at time t if

$$s^* = \arg \max_{s=1,\cdots,S} \sum_{k=1}^{K} x_k Y_k(t) \mu_{sk},$$

where $Y_k(t)$ represents the state of queue k (either the number of jobs or the workload) at time t, and x_k is any chosen positive queue weight. Another interesting question is, how the choice of the queue weight vector $x = (x_1, \dots, x_K)$ affects the performance measures of interest. It has been shown that the weight vector x affects the MaxProduct policy only through its directions in $\mathbb{R}_+^K$ (Hung and Michailidis 2008). Therefore, it suffices to consider the vectors x satisfying that $x_1 + \cdots + x_K = 1$. This constraint can be further reduced to $x_1 + \cdots + x_{K-1} \le 1$, which clearly represents a $simplex$ in $\mathbb{R}_+^{K-1}$. In this case, the queue weights $x_1, \dots, x_{K-1}$ are treated as input factors, while the input domain D is a $(K-1)$-simplex. An example of such D is shown in the right panel of Figure 1.

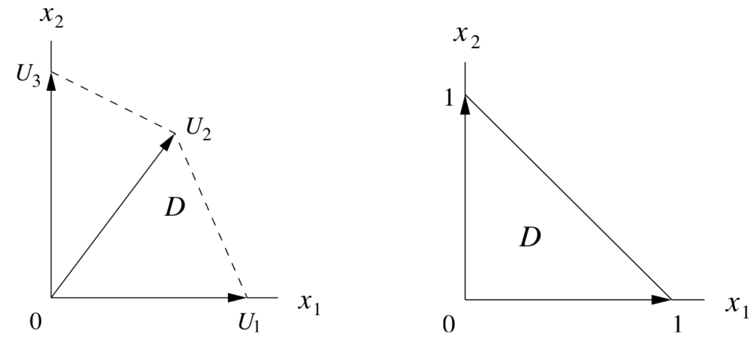

Fig. 1. (Left panel): The domain D of the input rate vectors x for a 2-queue system with three service rate vectors $U_1 = (3, 0)$, $U_2 = (2, 3)$, and $U_3 = (0, 4)$. (Right panel): The domain D of the weight vectors x (in the MaxProduct policy) for a 3-queue system (where $K = 3$ and $x_3 = 1 - x_1 - x_2$).

Note that due to the complex structure of dynamics, it is often hard to obtain the relationship between the input factors and the response measures of interest for such systems. Therefore, this type of problems are often examined through computer simulations (Hung et al. 2003, Hung and Michailidis 2008). However, simulation of such complex systems is expensive in terms of CPU time and the requirement of simulation resources. An natural question is then how one can obtain a comprehensive understanding of system's performance by performing the minimum possible number of simulation trials. This has been a challenging task in the area of design and analysis of computer experiments (DACE).

（二）研究目的

In this study, we propose a new UD method that is suitable for any types of design area under the framework of the so-called number-theoretic method (NTM). The proposed UD method has an important feature that the optimal design is invariant under coordinate rotations and can be properly extended so that lower-dimensional uniformity is also considered (see Section 2 for details). For practical applications, we also develop a methodology to estimate the target region of computer experiments by utilizing the proposed UD method. Note that the target region here represents a subset of the input domain in which the experimental output measure(s) of interest is desired to be produced. For example, let's consider the queueing system introduced in Example 1, where one can specify a target region for the input rate vectors so that the average delay (i.e., the average time waiting until service is first provided) of jobs in each queue does not exceed a prespecified quantity. By keeping the input rates in the target region, the system then provides a commitment to a certain level of quality service. Another example can be found in the recent work done by Ranjan et al. (2008), where a sequential design based on the Gaussian stochastic

（三）文獻探討

The uniform design (UD) was first proposed by Fang and Wang (Fang 1980, Wang and Fang 1981) and has been widely used in computer experiments over the last two decades (Fang and Lin 2003). Its basic idea is to seek input points to be uniformly scattered on the input domain so that the relationships between the response(s) and the input factors can be explored using a reasonable number of experimental trials. Traditional methods have provided solutions to UD for the experiments without restrictions (Fang and Wang 1994), i.e., the design area is or can be reasonably transformed into a unit hypercube $[0, 1]^K$. However, for non-rectangular types of design areas (such as the input areas shown in the two motivating examples), how to best perform the UD is not fully discussed in literature. The work done by Fang et al. (1999b) is closely related to the examples introduced above, wherein they proposed a "simplex method" to perform a stochastic representation of UD over convex polyhedrons. The shortcomings for this type of Monte-Carlo methods are: (i) the represented UD has larger variations; and (ii) they are shown to have relatively low efficiency on approximating the output measures of interest (Fang and Wang 1994).

（四）研究方法

Uniform Design over General Input Domains

The uniform design (UD), first proposed by Wang and Fang in 1980, is one of the space filling designs (Box and Drapper 1987, Cheng and Li 1995, Hickernell 1999, Wu et al. 2000) that seeks input points to be uniformly scattered on the input region D. Its basic idea is introduced in the following. Suppose we would like to choose a set of n experiment points $\mathcal{P} = \{p_1, \ldots, p_n\}$ that are uniformly scattered on an $identifiable$ input domain D, $D \subset \mathbb{R}^K$. Let M be a measure of uniformity of $\mathcal{P}$ such that smaller M corresponds to better uniformity. Let $Z(n)$ be the set of all possible sets $\{p_1, \ldots, p_n\}$ on D. A set $\mathcal{P}^* \in Z(n)$ is called a uniform design if it has the minimum value of M over $Z(n)$, i.e.,

$$M(\mathcal{P}^*) = \min_{\mathcal{P} \in Z(n)} M(\mathcal{P}). \tag{1}$$

The popular measures of uniformity are discrepancy (with various modified versions), dispersion, mean square error, and sample moments (Fang et al. 2000, Hickernell 1998, Fang and Wang 1994). However, most of these methods for UD are developed under the assumption that the experimental domain D can be reasonably transformed into a unit cube (e.g. rectangles). Motivated by the examples introduced in Section 1, in this study we propose a UD method that is suitable for experiments with any types of input domain.

A New Measure of Uniformity: Central Composite Discrepancy

We first introduce some notations that are necessary for constructing a new measure of uniformity called "central composite discrepancy". For any point $x \in \mathbb{R}$, denote the set

$$x^{(i)} = \{r \in \mathbb{R} : x + a_i < r \leq x + a_{i+1}\}, \quad i = 0, 1, \ldots, m-1, \tag{2}$$

where $a_0 = -\infty$, $a_m = \infty$, $a_1 < a_2 < \cdots < a_{m-1}$, and $a_j = 0$ for some

$1 \leq j \leq m-1$. Thus, the real line is divided into m parts at the point x. With the division on each coordinate of a given point $x = (x_1, \ldots, x_K) \in D \subset \mathbb{R}^K$, the input domain D is decomposed into (at most) m^K subregions, where the k-th subregion is denoted by $D_k(x) = \{x_1^{(i_1)} \times \cdots \times x_K^{(i_K)}\} \bigcap D$, and $(i_1, \ldots, i_K)$ is the base-m display of integer $k-1$. The examples of such a decomposition for a 2-dimensional convex polygon D are shown in Figure 2.

Consider a set of n experiment points $\mathcal{P} = \{p_1, \ldots, p_n\}$ on D and let

$$N(D_k(x), \mathcal{P}) = \sum_{i=1}^{n} I\{p_i \in D_k(x)\}, \tag{3}$$

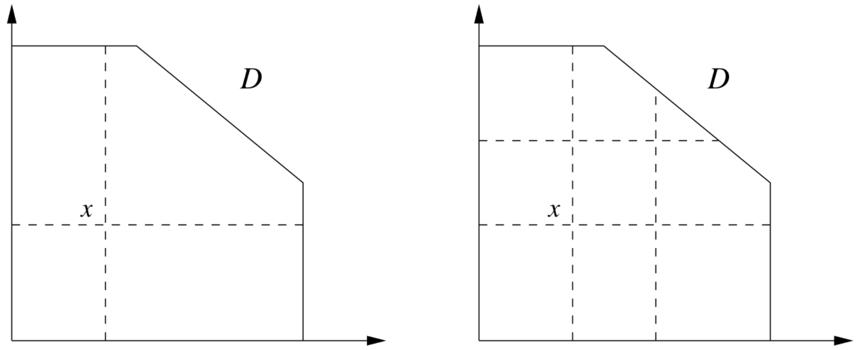

Fig. 2. (Left panel): The decomposition of D at x with $m = 2$ and $a_1 = 0$. (Right panel): The decomposition of D at x with $m = 3$ and $a_1 = 0$.

which represents the number of points allocated in the subregion $D_k(x)$ given by the decomposition of D at x, $x \in D$. The *central composite discrepancy* is defined as

$$CCD_p(n, \mathcal{P}) = \left\{ \frac{1}{v(D)} \int_D \frac{1}{m^K} \sum_{k=1}^{m^K} \left| \frac{N(D_k(x), \mathcal{P})}{n} - \frac{v(D_k(x))}{v(D)} \right|^p dx \right\}^{1/p}, \tag{4}$$

where $p > 0$, $v(D)$ and $v(D_k(x))$ denote the volume of D and $D_k(x)$, respectively. The optimal allocation of the n experiment points is the set that minimizes $CCD_p(n, \mathcal{P})$, that is,

$$\mathcal{P}^* = \arg \min_{\mathcal{P} \in Z(n)} CCD_p(n, \mathcal{P}). \tag{5}$$

Note that the goal of placing the quantity $1/v(D)$ in (4) is to rescale the input domain D so that it has volume one. However, this does not affect the optimal solution $\mathcal{P}^*$ for any given D.

The basic idea of the proposed central composite discrepancy is that each point x in D is treated as a "center", and uniformity is measured over all decomposed subregions around it. In the special case when D is a hyper-rectangle, $m = 2$, $a_1 = 0$, and $p = 1$, it is equivalent to the so-called "symmetrical L_1-discrepancy" (Ma 1997a). Note that the central composite discrepancy and the symmetrical discrepancy both share the same intuition that the optimal design is invariant under coordinate rotation. However, the former can be applied to the entire class of input domains while the latter can merely be applied to hyper-rectangles.

The Weighted Uniform Design

Let $f(x)$ be a continuous function defined on D, $f(x) > 0$ for all $x \in D$ and $\int_D f(x)dx = 1$. How do we find a set of n points $\mathcal{P} = \{p_1, \ldots, p_n\}$ on D so that they have a "good representation" for $f(x)$? By utilizing the measure of uniformity defined in (4), we next define the *weighted central composite discrepancy* by

$$WCCD_{f,p}(n, \mathcal{P}) = \left\{ \frac{1}{v(D)} \int_D \frac{1}{m^K} \sum_{k=1}^{m^K} \left| \frac{N(D_k(x), \mathcal{P})}{n} - F(D_k(x)) \right|^p dx \right\}^{1/p}, \tag{6}$$

where $F(D_k(x)) = \int_{D_k(x)} f(x)dx$ represents the proportion of points expected to be allocated on each subregion $D_k(x)$, $k = 1, \ldots, m^K$. Therefore, a good representation for $f(x)$ will be the set of points $\mathcal{P}^*$ that minimizes $WCCD_{f,p}(n, \mathcal{P})$.

Note that if $f(x)$ corresponds to a probability density function and p is chosen to be 1, then the quantity defined in (6) is a rotation-invariant version of the so-called "F-discrepancy" (Fang and Wang 1994). However, the interpretation of the function $f(x)$ is not restricted here. In general, it can represent the "weight" (or "importance") of each point x in D - the larger the value of $f(x)$ is, the more important the point x is considered. Some examples of how to choose the function $f(x)$ in correspondence with the prespecified targets of experiment are shown later in Section 3.

Construction of Nearly Uniform Designs

It is known that solving $\mathcal{P}^*$ is a NP hard problem as the number of allocated design points goes to infinity. In practice, a computationally more efficient way is to construct a so-called nearly uniform design (NUD) with a low measure of uniformity. Traditional techniques for constructing the NUDs are the good lattice point method and its modifications (Wang and Fang 1981, Fang and Li 1995, Ma 1997b), the method based on searching only a subset of U-type designs (Fang and Hickernell 1995), the construction methods based on Latin squares (Fang et al. 1999a) and orthogonal designs (Fang 1995), the threshold accepting method based on U-type designs (Winker and Fang 1998, Fang et al. 2001), the method by collapsing two uniform designs (Fang and Qin 2003), and the cutting method (Ma and Fang 2004). In order to deal with general types of input domain, here we utilize an efficient approach (called "switching algorithm") that have been widely used in design literature (Winker and Fang 1998, Fang et al. 2001) and cluster analysis (e.g. K-means clustering, Sharma 1996). The steps of the switching algorithm are summarized in the following.

The Switching Algorithm

Step 1: Superimpose N candidate grids $g_1, \ldots, g_N$ on the primary input domain and denote the new input domain by $D = \{g_1, \ldots, g_N\}$. Arbitrarily choose an initial design $\mathcal{P}^{(0)} = \{g_1, \ldots, g_n\}$ from D, set $i = 0$.

Step 2: Set $j = 1$ and $\mathcal{P}^{(i+1)} = \mathcal{P}^{(i)}$.

Step 3: Let $g^* = \arg\min_{g \in D \setminus \mathcal{P}^{(i+1)}} CCD_p(n, \{g\} \bigcup \mathcal{P}^{(i+1)} \setminus \{g_j\})$.
If $CCD_p(n, \{g^*\} \bigcup \mathcal{P}^{(i+1)} \setminus \{g_j\}) < CCD_p(n, \mathcal{P}^{(i+1)})$,
set $\mathcal{P}^{(i+1)} = \{g^*\} \bigcup \mathcal{P}^{(i+1)} \setminus \{g_j\}$.

Step 4: Set $j = j + 1$. If $j \leq n$, go to Step 3; otherwise go to Step 5.

Step 5: If $\mathcal{P}^{(i+1)} \neq \mathcal{P}^{(i)}$, set $i = i + 1$ and go to Step 2; otherwise return $\mathcal{P}^{(i)}$.

（五）結果與討論

We show that:
(i) solving $\mathcal{P}^{(i*)}$ requires at most $O(N^{2+p})$ computations of $CCD_p(n, \mathcal{P})$; and
(ii) the resulting $\mathcal{P}^{(i*)}$ approximates very well the optimal design $\mathcal{P}^*$.

Fact 1 *For any given $\mathcal{P} \subset D = \{g_1, \ldots, g_N\}$, $0 \leq [CCD_p(n, \mathcal{P})]^p \leq 1$.*

Fact 2 *$CCD_p(n, \mathcal{P}^{(i)})$ is a non-increasing function of i.*
Fact 3 *If p is a positive integer and $\mathcal{P}^{(i+1)} \neq \mathcal{P}^{(i)}$ in Step 5, then*

$$[CCD_p(n, \mathcal{P}^{(i)})]^p - [CCD_p(n, \mathcal{P}^{(i+1)})]^p \geq \frac{1}{n^p N^{1+p} m^K}.$$

Proof. Define

$$W(n, \mathcal{P}) = \sum_{g \in D} \sum_{k=1}^{m^K} |N \cdot N(D_k(g), \mathcal{P}) - n \cdot N(D_k(g), D)|^p,$$

it is clear that $W(n, \mathcal{P})$ is a positive integer and by definition

$$[CCD_p(n, \mathcal{P}^{(i)})]^p - [CCD_p(n, \mathcal{P}^{(i+1)})]^p = \frac{W(n, \mathcal{P}^{(i)}) - W(n, \mathcal{P}^{(i+1)})}{n^p N^{1+p} m^K}.$$

The result then follows since by Fact 2 we know that $W(n, \mathcal{P}^{(i)}) - W(n, \mathcal{P}^{(i+1)}) \geq 1$ when $\mathcal{P}^{(i+1)} \neq \mathcal{P}^{(i)}$.
Theorem 1 *For any positive integer p, the computation time of $CCD_p(n, \mathcal{P}^{(i)})$ in the switching algorithm is at most $O(N^{2+p})$.*

Proof. Note that to finish the update of each design $\mathcal{P}^{(i)}$, the required computation time of $CCD_p(n, \mathcal{P})$ is $n(N-n)$ (since there are n switchings needed to be checked and each switching requires $N-n$ computations of $CCD_p(n, \mathcal{P})$). In addition, from Fact 1–Fact 3 we know that $CCD_p(n, \mathcal{P}^{(i)})$ is a non-increasing function of i and $CCD_p(n, \mathcal{P}^{(0)})$ can be reduced at most $n^p N^{1+p} m^K$ times. These together imply the total computation time of $CCD_p(n, \mathcal{P})$ is bounded above by $n^p N^{1+p} m^K \cdot n(N-n)$, which can be represented as $O(N^{2+p})$.
The result of Theorem 1 is quite essential from the perspective of computational efficiency. To see this, note that the computation time of $CCD_p(n, \mathcal{P})$ for finding the optimal design (based on exhuastive search) is clearly $\binom{N}{n} = O(N^n)$. However, since p is often chosen to be 1 or 2, the computation time can be dramatically reduced to $O(N^{2+p})$ by using the switching algorithm. To investigate how the resulting NUD approximates the true optimal design $\mathcal{P}^*$, we consider the unit-square input domain on which 30 grids are superimposed (i.e., $N = 30$). The switching al-

gorithm is then carried out for 100 times (with each initial design chosen by simple random sampling) and the "average" central composite discrepancy (with $p = 2$) of all resulting NUDs is computed. The results for different sizes of experiment (say $n = 1, \ldots, 15$) are shown in Figure 3.

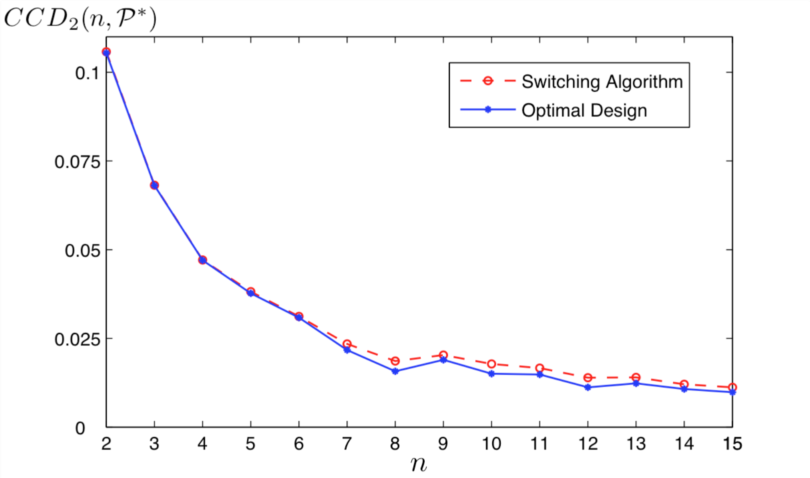

Fig. 3. The comparison of the optimal design and the NUD based on the switching algorithm for the input domain of unit square.

From Figure 3, we see that the NUDs obtained from the switching algorithm approximate very well (in average sense) the optimal design for various sizes of experiment. In addition, the numerical results show that the switching algorithm is quite stable since the standard deviation of the resulting $CCD_2(n, \mathcal{P}^{(i*)})$ (based on 100 NUDs) is less than 2×10^{-3} for all n. For practical purposes, the CPU time for finding the optimal design (based on exhaustive search) and the average CPU time for finding the NUDs (based on 100 iterations of the switching algorithm) are attached in Table 1. As can be seen from Table 1, as n becomes larger, the CPU time for finding the NUD based on the switching algorithm becomes significantly smaller than that of finding the optimal design.

Table 1

The CPU time (in seconds) for finding the optimal design (based on exhaustive search) and the average CPU time (in seconds) for finding the NUD.

n	Optimal Design	NUD	n	Optimal Design	NUD	n	Optimal Design	NUD
1	< 0.001	– –	6	34	0.289	11	4525	0.816
2	0.015	0.047	7	128	0.377	12	7592	0.878
3	0.172	0.100	8	384	0.446	13	11183	1.112
4	1.235	0.145	9	1048	0.586	14	14286	1.207
5	7	0.230	10	2342	0.686	15	19627	1.356

Discussion

Low-dimensional Uniformity. It is noted that the proposed measure of uniformity can be modified so that uniformity over low-dimensional spaces is also taken into account. Let ψ be a non-empty subset of $\{1, \ldots, K\}$ and $|\psi|$ be the number of elements in ψ. For any given weight function $f(x)$, define $f_\psi = \int f(x) dx_{-\psi}$, where $dx_{-\psi} = \prod_{i \in \{1,\ldots,K\} \setminus \psi} dx_i$. Thus, a more general measure of uniformity, called *projected central composite discrepancy*, can be defined as

$$PCCD_{\Psi,f,p}(n,\mathcal{P}) = \left\{ \frac{1}{|\Psi|} \sum_{\psi \in \Psi} [WCCD_{f_\psi,p}(n,\mathcal{P}_\psi)]^p \right\}^{1/p}, \quad (7)$$

where Ψ is the collection of all non-empty subsets of $\{1, \ldots, K\}$, and $\mathcal{P}_\psi$ is the projection of $\mathcal{P}$ to the subspace ψ. It is noted that similar ideas are used to construct the so-called "centered L_p-discrepancy" (Hickernell 1998). However, the consideration of projections to low-dimensional spaces will require a large amount of computation, as can be seen from (7). Therefore, here we restrict our attention on the uniformity in the K-dimensional space, i.e., we choose $\Psi = \{\{1, \ldots, K\}\}$.

Choice of N. It is noted that if N is too small (i.e., the number of candidate grids is too small), then the resulting NUD may not be good enough. On the other hand, if N is large, then the switching algorithm may not be efficient. Therefore, to choose the value of N, one must take into account the tradeoff between design optimality and computational cost. Here we provide a guideline for choosing N: if K is small, then choose $N = n^K$; if K is large, then choose $N = nK$. Based on the guideline, an n^K factorial design is used for choosing the candidate grids in a small space, while the number of candidate grids is chosen to be proportional to the size of design in a large space.

When K is Large. Another computational issue is how to evaluate the summations in (4) and (6) when K is large. In practice, this can be done by considering the summation over a random subset of all decomposed subregions. For example, equation (4) can be replaced by

$$\left\{ \frac{1}{v(D)} \int_D \frac{1}{L} \sum_{l=1}^{L} \left| \frac{N(D_{k_l}(x),\mathcal{P})}{n} - \frac{v(D_{k_l}(x))}{v(D)} \right|^p dx \right\}^{1/p},$$

where $\{k_1, \ldots, k_L\}$ is a random subset of $\{1, \ldots, m^K\}$, and similarly for equation (6).

參考文獻

Andre, J., Siarry, P., Dognon, T., 2000. An improvement of the standard genetic algorithm fighting premature convergence. Advances in Engineering Software 32 (1), 49–60.

Armony, M., Bambos, N., 2003. Queueing dynamics and maximal throughput scheduling in switched processing systems. Queueing Systems: Theory and Applications 44, 209–252.

Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers. In: The 5th Annual ACM Workshop on COLT. pp. 144–152.

Box, G.E.P., Drapper, D.R., 1987. Empirical Model Building and Response Surfaces. John Wiley & Sons, New York.

Cheng, C.S., Li, K.C., 1995. A study of the method of principal Hessian direction for analysis of data from design experiments. Statistica Sinica 5, 617–639.

Cortes, C., Vapnik, V.N., 1995. Support vector networks. Machine Learning 20, 273–297.

Fang, K.T., 1980. The uniform design: Application of number-theoretic methods in experimental design. Acta Mathematical Application Sinica 3, 363–372.

Fang, K.T., Hickernell, F.J., 1995. The uniform design and its applications. Bulletin of Institute of International Statistics 333–349. 50th Session, Book 1.

Fang, K.T., Li, J.K., 1995. Some new results on uniform design. Chinese Science Bulletin 40, 68–72.

Fang, K.T., Lin, D.K.J., 2003. Uniform experimental designs and their applications in industry. In: Handbook of Statistics, vol. 22. pp. 131–170.

Fang, K.T., Lin, D.K.J., Winker, P., Zhang, Y., 2000. Unifrom design: Theory and applications. Technometrics 42, 237–248.

Fang, K.T., Ma, C.X., Winker, P., 2001. Centered L_2-discrepancy of random sampling and Latin hypercube design, and construction of uniform design. Mathematical Computation 71, 275–296.

Fang, K.T., Qin, H., 2003. A note on construction of nearly uniform designs with large number of runs. Statistics & Probability Letters 61, 215–224.

Fang, K.T., Shiu, W.C., Pan, J.X., 1999a. Uniform designs based on Latin squares. Statistica Sinica 9, 905–912.

Fang, K.T., Tian, G.L., Xie, M.Y., 1999b. Uniform design over a convex polyhedron. Chinese Science Bulletin 44, 112–114.

Fang, K.T., Wang, Y., 1994. Number-theoretic Methods in Statistics. Chapman and Hall, London.

Fang, Y., 1995. Relationships between uniform design and orthogonal design. In: The 3rd International Chinese Statistical Association Conference, Beijing.

Guyon, I.M., Boser, B.E., Vapnik, V.N., 1993. Automatic capacity tuning of very large VC-dimension classifiers. Advances in Neural Information Processing Systems 5, 147–155.

Henderson, C.R., 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31, 423–447.

Hickernell, F.J., 1998. A generalized discrepancy and quadrature error bound. Math. Comput. 67, 299–322.

Hickernell, F.J., 1999. Goodness-of-fit statistics, discrepancies and robust dessigns. Statistics & Probability Letters 44, 73–78.

Hsu, C.W., Chang, C.C., Lin, C.J., 2003. A practical guide to support vector classification. Technical Report CWH03a, Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan.

Huang, C.M., Lee, Y.J., Lin, D.K.J., Huang, S.Y., 2007. Model selection for support vector machines via uniform design. Computational Statistics & Data Analysis 52, 335–346.

Hung, Y.C., Chang, C.C., 2008. Dynamic scheduling for switched processing systems with substantial service-mode switching times. Queueing Systems: Theory and Applications 60, 87–109.

Hung, Y.C., Michailidis, G., 2008. Modeling, scheduling, and simulation of switched processing systems. ACM Transactions on Modeling and Computer Simulation 18, Article 12.
Hung, Y.C., Michailidis, G., Bingham, D.R., 2003. Developing efficient simulation methodology for complex queueing networks. In: Proceedings of the Winter Simulation Conference, New Orleans. pp. 152–159.
Jones, D., Schonlau, M., Welch, W., 1998. Efficient global optimization of expensive Black–Box functions. Journal of Global Optimization 13, 455–492.
Keerthi, S.S., Lin, C.-J., 2003. Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Computation 15, 1667–1689.
Ma, C.X., 1997a. A new criterion of uniformity — Symmetrical discrepancy. Journal of Nankai University 30, 30–37.
Ma, C.X., 1997b. Construction of uniform designs using symmetrical discrepancy. Application of Statistics and Manegement 166–169.
Ma, C.X., Fang, K.T., 2004. A new approach to construction of nearly uniform designs. International Journal of Materials and Product Technology 20, 115–126.
Ranjan, R., Bingham, D., Michailidis, G., 2008. Sequential experiment design for contour estimation from complex computer codes. Technometrics 50, 527–541.
Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and analysis of computer experiments. Statistical Science 4, 409–423.
Schölkopf, B., 1997. Support Vector Learning. R. Oldenbourg Verlag, Munich.
Sharma, S., 1996. Applied Multivariate Techniques. Wiley.
Vapnik, V.N., 1998. Statistical Learning Theory. Wiley, New York.
Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer, New York.
Wang, Y., Fang, K.T., 1981. A note on uniform distribution and experimental design. KeXue TongBao 26, 485–489.
Winker, P., Fang, K.T., 1998. In: Niederreiter, H., Zinterhof, P., Hellekalek, P. (Eds.), Optimal U-type Design. Monte Carlo and Quasi-Monte Carlo Methods 1996. Springer, 436–448.
Wu, C.F.J., Hamada, M., 2000. Experiments: Planning, Analysis, and Parameter Design. Wiley, New York.

計畫成果自評

This research is mainly divided into two parts. In the first part, we propose a new UD method that is better suited to any convex types of design areas than the existing UD methods. The proposed UD method has an important feature that the optimal design is invariant under coordinate rotations. In order to reduce the computational cost of finding the optimal UD, we propose an efficient algorithm to construct a so-called nearly uniform design (NUD). The numerical results show that the constructed NUDs approximate very well the true optimal UD solution. In the second part, we develop an efficient methodology for estimating the target region of computer experiments. The methodology is sequential and comprised of two main components: (i) design; and (ii) fitting response surfaces. For component (i), the proposed UD method with sequentially updated weight functions is utilized; while for component (ii), a technique called on-line support vector regression (i.e., a sequentially updated SVR model) is employed to model the response surfaces of interest. It is noted that the proposed methodology can be viewed as an extension of the work done by Ranjan et al. (2008). However, it has a more general scope from the following viewpoints: (i) it is easy to implement; (ii) it is completely data-driven and does not require any model assumptions (such as the GASP model); (iii) it can easily handle the experiments with a large number of input factors; (iv) it can handle the target region comprised of multiple output measures (remember the work done by Ranjan et al. (2008) can handle merely one output measure). The numerical results also show that our proposed methodology outperforms other approaches in estimating the target region of various computer experiments. We are currently investigating the computational issues that arise with high dimensional input spaces and also how to best compare the performance of different approaches.

The result of this research has been published in:

Computational Statistics and Data Analysis 54 (2010) 219–232.

附件四

國科會補助專題研究計畫項下出席國際學術會議心得報告

日期: 100 年 1 月 5 日

計畫編號	NSC 98 - 2118 - M - 004 - 007 -		
計畫名稱	**凸面區域之均勻設計及其在電腦實驗上之應用**		
出國人員姓名	**洪英超**	服務機構及職稱	國立政治大學統計系
會議時間	99年 8 月26日至 99年 8 月28日	會議地點	City of Kitakyushu, Japan
會議名稱	(中文) 第七屆國際管理工程會議 (英文) The 7th International Symposium on Management Engineering		
發表論文題目	(中文) 吃角子老虎問題之最佳策略及其在資料路由網路之應用 (英文) An Optimal Bayesian Strategy for Bandit Problems with Applications to Data Routing Networks		

一、參加會議經過

第七屆國際管理工程會議由日本早稻田大學的 Management Engineering Research Group於北九州市之國際會議廳主辦，其內容包括： organization, corporate strategy, project management, management of technology, as well as intelligence, computation, operations research, probability and possibility theories, fuzzy sets, rough sets, approximate reasoning, linguistic information processing and automata, knowledge discovery, clustering and data analysis, fuzzy control and modeling, optimization under uncertainty and its applications, games and decision making, fuzzy analysis, fuzzy game等等。本人提出的方法引起許多與會人士的高度興趣，除此之外，我也得到許多與會人士的建議與回饋。

二、與會心得

本次與會的人士皆來自高科技的先進國家，所以我們也感到與世界強國競爭的壓力。主辦國日本在機器人(robot) 研究設計及fuzzy control、intelligent system方面的成就令人驚訝，其中更運用了許多統計的方法。這也啟發了我們未來在相關領域上的研究動機與方向。

三、考察參觀活動(無是項活動者略)

四、建議

我想臺灣若要在相關領域佔有一蓆之地，當務之急必須投下更多的人力與設備，並儘可能鼓勵年輕學者多參加大型國際會議。如此才能掌握研究的趨勢與動脈，做到真正的國際化與國際交流。

五、攜回資料名稱及內容

研討會論文集ＣＤ。

六、其他

國科會補助計畫衍生研發成果推廣資料表

日期:2011/01/04

國科會補助計畫	計畫名稱: 凸面區域之均勻設計及其在電腦實驗上之應用
	計畫主持人: 洪英超
	計畫編號: 98-2118-M-004-007- 學門領域: 其他應用統計

無研發成果推廣資料

98 年度專題研究計畫研究成果彙整表

計畫主持人：洪英超	計畫編號：98-2118-M-004-007-
計畫名稱：凸面區域之均勻設計及其在電腦實驗上之應用	

成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	2	2	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	4	4	100%	人次	
		博士生	1	1	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	1	1	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	2	2	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：■已發表 □未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

本研究內容主要分成二部份:第一部份是提出一個全新且適合一般所謂電腦實驗(Computer Experiments)的均勻設計(Uniform Design)方法。此方法解決了現有均勻設計不適用非矩形區域 (Non-rectangular Area)的缺點，且同時具有無旋轉性及低維度均勻等性質。第二部份是設計一套’估計電腦實驗目標區’的演算法，其想法是利用所提的均勻設計加入權重(weight)的概念，並透過逐步的模型配適(modelffitting)及更新權重函數的概念來配置實驗參數，進而估計出預設的電腦實驗目標區。此演算法的優點是：(1)適用任意形狀的實驗區域；(2)對高維度的輸入空間提供穩定且計算成本較低的解；(3)可同時處理多個輸出值(multiple response)的電腦實驗問題，這些優點都是現有文獻中所沒有的。透過不同的電腦模擬實驗，我們也驗證了所提出的方法比文獻中的所有方法更加的準確及有效率。