Design and Analysis of Computer Experiments

Hongquan Xu
UCLA Department of Statistics
November 2020

Outline

What are computer experiments?

- Computer experiments are increasingly being used to explore the behavior of complex physical systems.
- A computer model is a large computer code that implements a complex mathematical model of a physical process.
  e.g., simultaneous differential solver, finite element analysis computational fluid dynamics.

Computer Experiments

A typical engineering model (Page 1 of 3)

$$
\begin{aligned}
&\rho_b A_b \frac{\partial^2 w}{\partial t^2} + E_b I_b \frac{\partial^4 w}{\partial x^4} \\
&+\{(\rho_s A_s + \rho_c A_c)\frac{\partial^2 w}{\partial t^2} + \rho_s A_s\left(\frac{t_b+t_s}{2}\right)\left(\frac{\partial^3 u_b}{\partial x \partial t^2} - \frac{t_b+t_s}{2}\frac{\partial^4 w}{\partial x^2 \partial t^2} + \frac{t_s}{2}\frac{\partial^3 \beta}{\partial x \partial t^2}\right) \\
&+\rho_c A_c a\left(\frac{\partial^3 u_b}{\partial x \partial t^2} - a\frac{\partial^4 w}{\partial x^2 \partial t^2} + t_s \frac{\partial^3 \beta}{\partial x \partial t^2}\right) + C_{11}^{0} I_c \frac{\partial^4 w}{\partial x^4} - E_c A_c a\left(\frac{\partial^3 u_b}{\partial x^3} - a\frac{\partial^4 w}{\partial x^4} + t_s \frac{\partial^3 \beta}{\partial x^3}\right)\}[H(x-x_1) - H(x-x_2)] \\
&+\{\rho_s A_s\left(\frac{t_b+t_s}{2}\right)\left(\frac{\partial^2 u_b}{\partial t^2} - \frac{t_b+t_s}{2}\frac{\partial^3 w}{\partial x \partial t^2} + \frac{t_s}{2}\frac{\partial^2 \beta}{\partial t^2}\right) + \rho_c A_c a\left(\frac{\partial^2 u_b}{\partial t^2} - a\frac{\partial^3 w}{\partial x \partial t^2} + t_s \frac{\partial^2 \beta}{\partial t^2}\right) \\
&+2C_{11}^{0} I_c \frac{\partial^3 w}{\partial x^3} - 2E_c A_c a\left(\frac{\partial^2 u_b}{\partial x^2} - a\frac{\partial^3 w}{\partial x^3} + t_s \frac{\partial^2 \beta}{\partial x^2}\right)\}[\delta(x-x_1) - \delta(x-x_2)] \\
&+\{C_{11}^{0} I_c \frac{\partial^2 w}{\partial x^2} - E_c A_c a\left(\frac{\partial u_b}{\partial x} - a\frac{\partial^2 w}{\partial x^2} - t_s \frac{\partial \beta}{\partial x}\right) + b d_{31} E_c a V(t)\}[\delta'(x-x_1) - \delta'(x-x_2)] = f(x,t)
\end{aligned}
\quad (1)
$$

$$
\begin{aligned}
&\rho_b A_b \frac{\partial^2 u_b}{\partial t^2} - E_b A_b \frac{\partial^2 u_b}{\partial x^2} \\
&+\{\rho_s A_s\left(\frac{\partial^2 u_b}{\partial t^2} - \frac{t_b+t_s}{2}\frac{\partial^3 w}{\partial x \partial t^2} + \frac{t_s}{2}\frac{\partial^2 \beta}{\partial t^2}\right) + \rho_c A_c\left(\frac{\partial^2 u_b}{\partial t^2} - a\frac{\partial^3 w}{\partial x \partial t^2} + t_s \frac{\partial^2 \beta}{\partial t^2}\right) \\
&-E_c A_c\left(\frac{\partial^2 u_b}{\partial x^2} - a\frac{\partial^3 w}{\partial x^3} + t_s \frac{\partial^2 \beta}{\partial x^2}\right)\}[H(x-x_1) - H(x-x_2)] \\
&+\{-E_c A_c\left(\frac{\partial u_b}{\partial x} - a\frac{\partial^2 w}{\partial x^2} + t_s \frac{\partial \beta}{\partial x}\right) + b d_{31} E_c V(t)\}[\delta(x-x_1) - \delta(x-x_2)] = 0
\end{aligned}
\quad (2)
$$

$$
\begin{aligned}
&[\rho_s A_s \frac{t_s}{2}\left(\frac{\partial^2 u_b}{\partial t^2} - \frac{t_b+t_s}{2}\frac{\partial^3 w}{\partial x \partial t^2} + \frac{t_s}{2}\frac{\partial^2 \beta}{\partial t^2}\right) + \rho_c A_c t_s\left(\frac{\partial^2 u_b}{\partial t^2} - a\frac{\partial^3 w}{\partial x \partial t^2} + t_s \frac{\partial^2 \beta}{\partial t^2}\right) \\
&+A_s(G \circ \beta) - E_c A_c t_s\left(\frac{\partial^2 u_b}{\partial x^2} - a\frac{\partial^3 w}{\partial x^3} + t_s \frac{\partial^2 \beta}{\partial x^2}\right)][H(x-x_1) - H(x-x_2)]
\end{aligned}
\quad (3)
$$

Computer Experiments

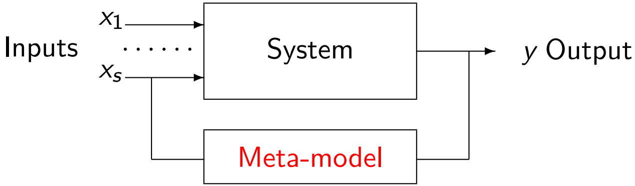


Figure 1: Computer experiment

Computer Experiments

Characteristics of computer experiments

- Mostly deterministic (lack of random error)
- May take hours or even days to produce a single output
- Many input variables
- The performance of the predictor depends upon the choice of the training data (design).

Principles in traditional DOE are irrelevant

- Replication
- Blocking
- Randomization

Modeling Computer Experiments: Kriging

For $x \in \mathcal{R}^m$, treat the deterministic response $y(x)$ as a realization of a stochastic process

$$Y(x) = \sum_{j=1}^{k} \beta_j f_j(x) + Z(x),$$

where $f_j(x)$ are known functions, β_j are unknown parameters and $Z(\cdot)$ is a **Gaussian process** with mean 0 and covariance

$$\text{cov}\,(Z(w), Z(x)) = \sigma^2 R(w, x).$$

- This is the Kriging model used in spatial statistics.
- Also called Gaussian process model in Machine Learning.

R packages: `DiceKriging`, `kergp`, etc.

Prediction

Given a design $S = \{s_1, \ldots, s_n\}$ and data $y_S = (y(s_1), \ldots, y(s_n))^T$. Consider the linear predictor

$$\hat{y}(x) = c(x)^T y_S.$$

Frequentists replace y_S by the random vector $Y_S = \{Y(s_1), \ldots, Y(s_n)\}^T$, and compute the MSE.

The **Best Linear Unbiased Predictor** (BLUP): choose $c(x)$ to minimize

$$\text{MSE}\,[\hat{y}(x)] = E[c(x)^T Y_S - Y(x)]^2$$

subject to

$$E[\hat{y}(x)] = E[c(x)^T Y_S] = E[Y(x)]$$

BLUP

Kriging model: $Y(x) = f(x)^T \beta + Z(x)$, where

$$f(x) = (f_1(x), \dots, f_k(x))^T, \beta = (\beta_1, \dots, \beta_k)^T$$

In matrix form:

$$Y_S = F\beta + Z, \ \text{cov}(Z) = \sigma^2 R$$

$$F = (f(s_1), \dots, f(s_n))^T = (f_j(s_i))_{n \times k}$$

$$R = (R(s_i, s_j))_{n \times n}$$

$$r(x) = (R(s_1, x), \dots, R(s_n, x))^T$$

The generalized LS estimate and BLUP are

$$\hat{\beta} = (F^T R^{-1} F)^{-1} F^T R^{-1} Y_S$$

$$\hat{y}(x) = f(x)^T \hat{\beta} + r(x)^T R^{-1}(Y_S - F\hat{\beta})$$

The GP interpolates the observed data: $\hat{y}(s_i) = y(s_i)$ for $s_i \in S$.

Correlation Functions

The correlation $R(w,x)$ has to be specified. Commonly used functions:

$$\begin{aligned} R(w,x) &= \prod \exp(-\theta_j |w_j - x_j|^{p_j}),\ 0 < p_j \leq 2, \\ R(w,x) &= \prod K(|w_j - x_j|; \theta_j) \end{aligned}$$

where $K()$ is Matérn correlation function with parameter $\nu = 5/2$.

$$K(h;\theta) = \left(1 + \frac{\sqrt{5}h}{\theta} + \frac{5h^2}{3\theta^2}\right) \exp\left(-\frac{\sqrt{5}h}{\theta}\right).$$

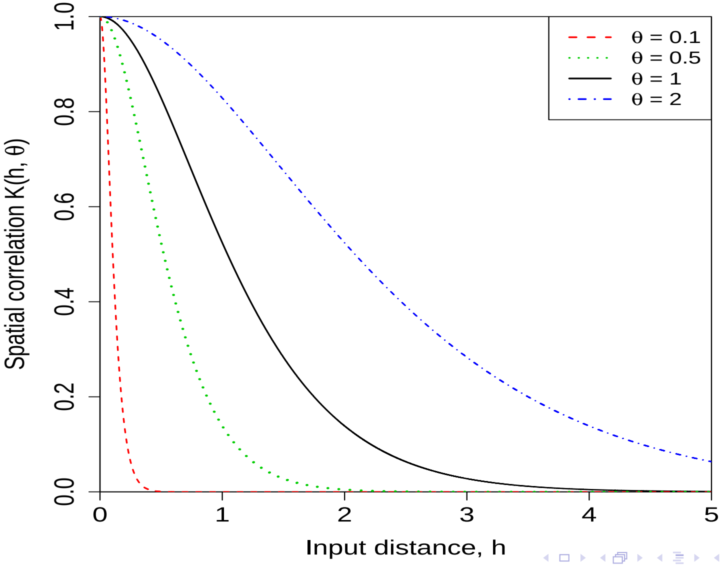
The correlation parameters (e.g., θ_j, p_j) need to be specified or estimated (by MLE or cross validation)

Given the correlation parameters, the MLEs are

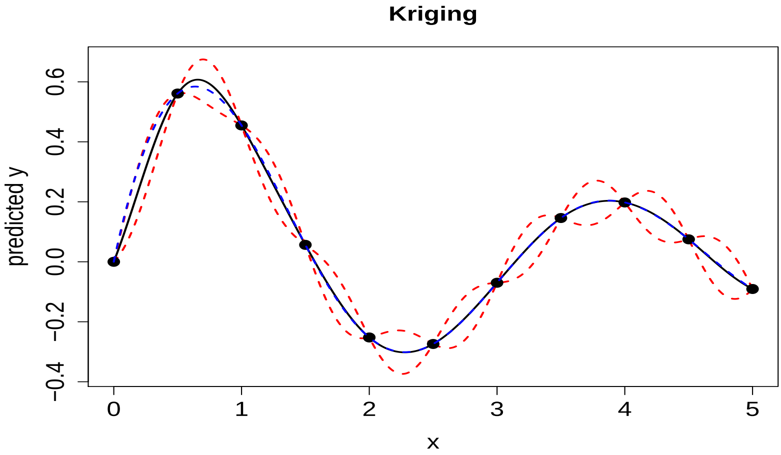
$$\hat{\beta} = \text{generalized l.s. estimate}$$

$$\hat{\sigma}^2 = \frac{1}{n}(Y_S - F\hat{\beta})' R^{-1} (Y_S - F\hat{\beta})$$

Examples of Matérn $\nu = 5/2$ correlation functions

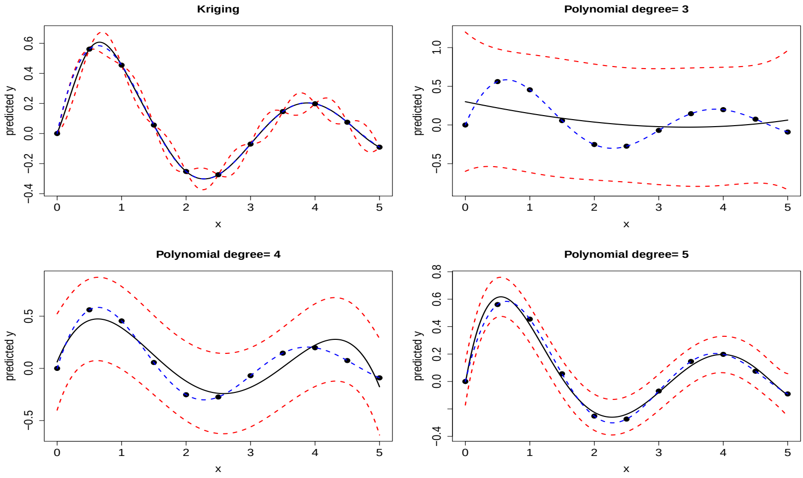


A toy example: Kriging



Data: $y = \sin(2x)/(1+x)$; Kriging: $Y = \mu + Z(x)$.

A toy example: Kriging vs Polynomial models



Data: $y = \sin(2x)/(1+x)$; Kriging: $Y = \mu + Z(x)$.

Designs for Computer Experiments

- Constructing a "good" design is crucial for the success of a computer experiment.
- A "good" design should be space-filling (i.e., cover as much space as possible), and have good projection properties.
  - Latin hypercube designs (LHD) [McKay et al. (1979)]
  - Maximin and minimax distance designs [Johnson et al. (1990)]
  - Orthogonal Array-based designs [Owen (1992), Tang (1993), He and Tang (2013, 2014)]
  - Uniform designs [Fang et al. (2006)]
  - Maximum projection designs [Joseph et al. (2015)]
  - Uniform projection designs [Sun et al. (2019)]
- Optimality criteria: maximin distance, minimax distance, column-orthogonality, uniformity (discrepancy) etc.
- R packages: `lhs`, `LHD`, `SLHD`, `UniDOE`, `MaxPro`, etc.

Design Criteria

Let $\hat{y}(x)$ be BLUE of $y(x)$ given a design $S = \{s_1, \ldots, s_n\}$.

- Integrated Mean Squared Error (IMSE)

$$\min_S : \int_{\mathcal{X}} \text{MSE}\,[\hat{y}(x)]\phi(x)dx,$$

  where $\phi(x)$ is a given weight function.

- Maximum Mean Squared Error (MMSE)

$$\min_S : \max_{x \in \mathcal{X}} \text{MSE}\,[\hat{y}(x)].$$

- Entropy (Gaussian process)

$$\max_S : \det(R) = \det(R(s_i, s_j)).$$

- Maximin distance criterion:

$$\max_S : \min_{i<j} d(s_i, s_j)$$

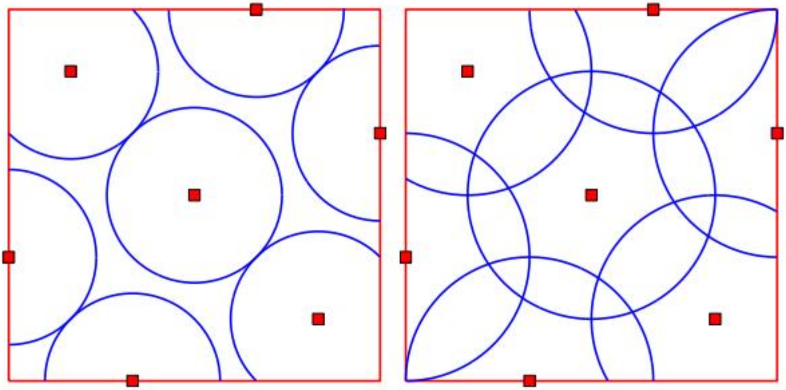


Figure 2: Maximin LHD (left) and Minimax LHD (right) with $n = 7$ and $m = 2$

Maximin distance designs

- For an (n, s^m) design $D = (x_{ik})_{n\times m}$,

$$d_p(x_i, x_j) = \sum_{k=1}^{m} |x_{ik} - x_{jk}|^p,$$

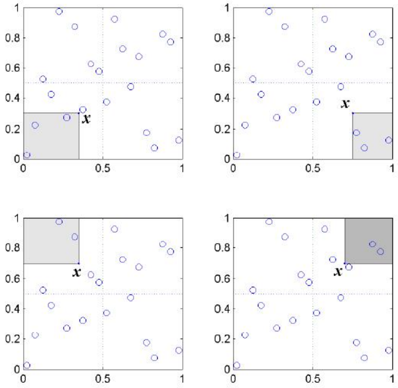
- Define the L_p-distance of D as

$$d_p(D) = \min\{d_p(x_i, x_j), 1 \le i < j \le n\}$$

- Maximin distance design: maximize $d_p(D)$ among all designs
- Most constructions are based on stochastic algorithms:
  - Morris and Mitchell (1995), Joseph and Hung (2008), Ba, Myers and Brenneman (2015, R package SLHD), etc.
  - Flexible but are not effective for large designs.
- Low-dimensional projections may not be space-filling.
  - Saturated $OA(n, 2^m)$'s are maximin distance designs when $m = n - 1$ (Xu 1999).

Uniform designs

Idea: choose design points from the design region with empirical distribution as "uniform" as possible (Fang et al, 2006).



Uniform Designs and Centered L_2-Discrepancy

For an $n \times m$ design D over $[0,1]^m$,

$$\mathrm{Disc}(\mathrm{D}) = \left\{ \int_{[0,1]^{\mathrm{m}}} \left| \mathrm{Vol}(\mathrm{J}(\mathrm{a_x}, \mathrm{x})) - \frac{\mathrm{N}(\mathrm{D} \cap \mathrm{J}(\mathrm{a_x}, \mathrm{x}))}{\mathrm{n}} \right|^2 \mathrm{dx} \right\}^{1/2}.$$

The (squared) centered L_2-discrepancy is defined by

$$\mathrm{CD}(\mathrm{D}) = \left\{ \sum_{\mathrm{u} \subseteq \{1:\mathrm{m}\}} |\mathrm{Disc}(\mathrm{D_u})|^2 \right\},$$

where $\mathrm{D_u}$ is the projected design of D onto dimensions indexed by the elements of u.

- Uniform designs may have poor projections in lower dimensional spaces.

Uniform Projection Designs: A New Class of Space-Filling Designs

- Focus on 2-dim projection uniformity
- *Uniform projection criterion* (Sun, Wang and Xu, 2019, Annals of Statistics)

$$\phi(D) = \frac{2}{m(m-1)} \sum_{|u|=2} CD(D_u), \tag{1}$$

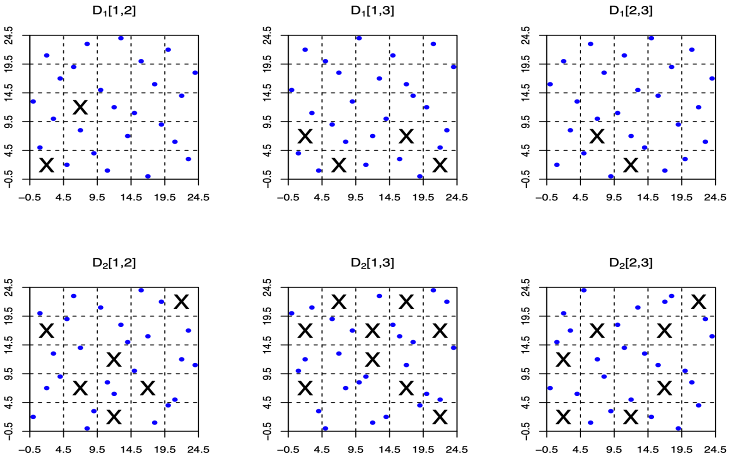
- A design achieving the minimum $\phi(D)$ value is a **uniform projection design** (UPD).
- The discrepancy has an analytical expression; for $D = (z_{ik})$ over $[0,1]^m$:

$$\begin{aligned} CD(D) &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}\left(1+\frac{1}{2}|z_{ik}|+\frac{1}{2}|z_{jk}|-\frac{1}{2}|z_{ik}-z_{jk}|\right) \\ &\quad -\frac{2}{n}\sum_{i=1}^{n}\prod_{k=1}^{m}\left(1+\frac{1}{2}|z_{ik}|-\frac{1}{2}|z_{ik}|^2\right)+\left(\frac{13}{12}\right)^m. \end{aligned}$$

Why we need a new criterion? Four 25 $\times$ 3 LHDs

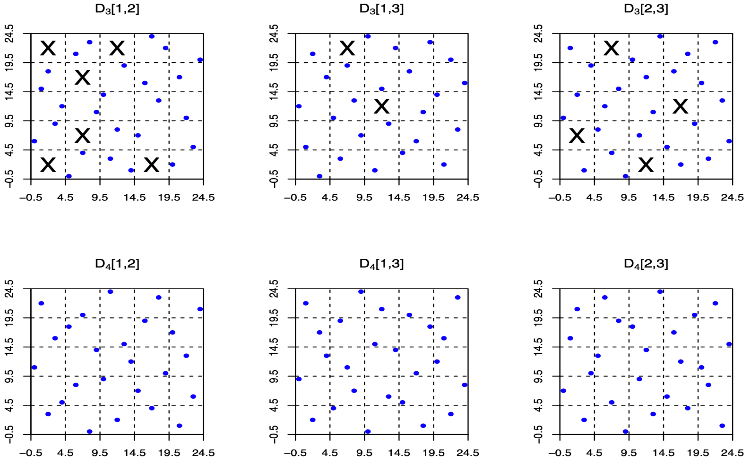
Uniform D_1			Maximin D_2			MaxPro D_3			UPD D_4		
18	16	14	0	2	20	0	6	12	2	3	2
19	9	0	1	20	10	1	15	5	4	5	13
11	1	2	2	7	12	2	18	21	0	11	9
16	20	3	3	13	21	3	9	0	3	16	17
20	22	12	4	9	3	4	12	17	1	22	22
14	7	10	5	19	0	5	0	10	8	0	7
4	17	1	6	23	19	6	21	3	6	8	19
12	12	7	7	14	13	7	4	19	9	14	24
10	15	24	8	0	7	8	23	13	5	18	4
22	14	5	9	3	17	9	11	7	7	20	11
2	21	22	10	21	8	10	14	24	12	2	21
15	11	21	11	8	9	11	3	1	10	9	0
1	5	4	12	6	1	12	8	15	14	12	14
3	10	11	13	18	23	13	19	9	13	15	6
23	3	8	14	15	2	14	1	22	11	24	15
0	13	15	15	10	18	15	7	4	17	4	10
8	23	6	16	24	16	16	16	18	15	7	5
7	8	18	17	16	11	17	24	6	19	10	18
9	4	13	18	1	15	18	13	11	16	19	20
6	19	9	19	22	4	19	22	23	18	23	1
24	18	19	20	4	6	20	2	14	21	1	16
21	6	23	21	5	24	21	17	2	23	6	23
13	24	17	22	12	5	22	10	20	22	13	3
17	0	16	23	17	22	23	5	8	20	17	12
5	2	20	24	11	14	24	20	16	24	21	8

Bivariate projections of Uniform D_1 and Maximin D_2



Note: 'X' means that there are no points in the grid.

Bivariate projections of MaxPro D_3 and UPD D_4



Note: 'X' means that there are no points in the grid.

Some Theoretical Results

Theorem 1

For a balanced (n, s^m) design D and any $2 \leq k \leq m$,

$$\frac{1}{\binom{m}{k}} \sum_{|u|=k} \phi(D_u) = \phi(D),$$

where D_u is the projected design onto k factors indexed by u.

- UPDs have good space-filling properties not only in two dimensions, but also in all dimensions.

Some Theoretical Results

Theorem 2

For a balanced (n, s^m) *design* $D = (x_{ik})$,

$$\phi(D) = \frac{g(D)}{4m(m-1)n^2s^2} + C(m,s), \qquad (2)$$

where

$$g(D) = \sum_{i=1}^{n}\sum_{j=1}^{n} d_1^2(x_i, x_j) - \frac{2}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{n} d_1(x_i, x_j)\right)^2 \qquad (3)$$

- $\phi(D)$ is a function of pairwise L_1-distances of the rows of D.
- An equidistant design under the L_1-distance is a UPD.

Application: Design and Modeling Comparison

- A 3-drug combination experiment on lung cancer (Al-Shyoukh et al. 2011; Xiao, Wang and Xu 2019).
- A 512-run and 8-level full factorial design to study 3 drugs.
- The response was the ATP level of the cells after the drug treatments.
- Kriging model with noise: $y(x) = \mu + Z(x) + \epsilon$

Table 1: Comparison of 1000×MSE for different models and designs

	Normal Cell				Cancer Cell			
	D_{512}	RD_{80}	MPD_{25}	UPD_{25}	D_{512}	RD_{80}	MPD_{25}	UPD_{25}
Kriging	0.002	0.21	0.62	0.22	0.003	0.37	1.87	0.21
NN	0.37	1.28	3.12	1.79	0.47	1.57	4.10	2.93
Polynomial	0.48	1.16	3.22	0.74	2.98	6.77	10.04	4.42

RD_{80}: Random 80-run design; MPD_{25}: MaxPro 25-run designs.

Comparison of projection properties

We compare four LHD(19, 18)'s:

1. The uniform design is from the uniform design website(UD)
2. The maximin distance design via R package SLHD (Ba, Myers and Brenneman, 2015, Technometrics).
3. The maximum projection (MaxPro) design were constructed via R package MaxPro (Joseph et al., 2015, Biometrika)
4. The uniform projection design (UPD): E_b.

We ran R commands maximinSLHD (with slice parameter $t = 1$) and MaxProLHD 100 times with default settings and chose the best designs.

Comparison of projection properties

Four criteria will be used in the comparison:

1. minimum Euclidean distance
2. maximum projection criterion (Joseph et al. 2015)

$$\psi(D) = \left\{ \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{\prod_{k=1}^{m} (x_{ik} - x_{jk})^2} \right\}^{1/m}$$

3. relative maximum centered L_2-discrepancy (CD)
4. maximum correlation ρ_{ave}.

For each k, we evaluate all $\binom{m}{k}$ projected designs and determine the worst projection with respect to four criteria.

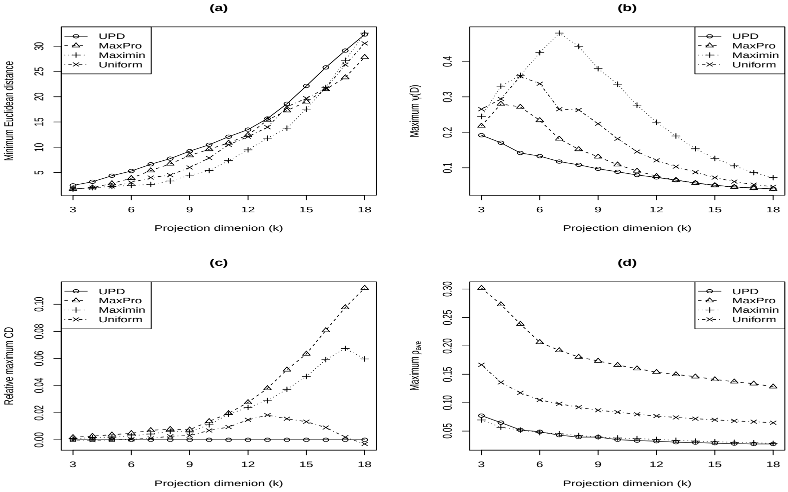


Figure 3: (a) minimum Euclidean distance (the larger the better), (b) maximum $\psi(D)$ (the smaller the better), (c) relative maximum CD (the smaller the better), and (d) maximum ρ_{ave} (the smaller the better).

- Good Lattice Point (GLP) designs are LHDs and often used to construct uniform designs (Fang and Wang, 1994).
- Let $h_1 < \ldots < h_p$ be p integers (from 1 to n) coprime to n

$$D = (x_{ij}) \text{ with } x_{ij} = i \times h_j \pmod{n}$$

An example $n = 7$:

$$D = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 1 & 3 & 5 \\ 3 & 6 & 2 & 5 & 1 & 4 \\ 4 & 1 & 5 & 2 & 6 & 3 \\ 5 & 3 & 1 & 6 & 4 & 2 \\ 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{array}{c} \text{with } d(D) = 12 \\ (\text{while } d_{upper} = 16) \end{array}$$

GLP designs

Results from Zhou and Xu (2015, Biometrika)

- An upper bound (for L_1-distance): For any $N \times n$ LHD D,

$$d(D) \leq d_{upper} = \lfloor (N+1)n/3 \rfloor,$$

  where $\lfloor x \rfloor$ is the integer part of x.
- Obtain the distances for four classes of GLP designs.
  - For an odd prime n, the $n \times (n-1)$ GLP design has $d(D) = (n+1)(n-1)/4$.
  - the upper bound is $d_{upper} = (n+1)(n-1)/3$
- The $d_{eff}(D) = d(D)/d_{upper}$ for a GLP design is 75%.
- A surprising result: any linear level permutation of any column does not decrease the distance $d(D)$.

GLP + Linear Permutation

Example: $n = 7$**:** Total $7^6 = 117,649$ linear permutations.

- consider only 7 simple permutations: $D_i = D + i \mod n$

$$D \quad \rightarrow \quad D_1 = D + 1 \mod n$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 1 & 3 & 5 \\ 3 & 6 & 2 & 5 & 1 & 4 \\ 4 & 1 & 5 & 2 & 6 & 3 \\ 5 & 3 & 1 & 6 & 4 & 2 \\ 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 0 \\ 3 & 5 & 0 & 2 & 4 & 6 \\ 4 & 0 & 3 & 6 & 2 & 5 \\ 5 & 2 & 6 & 3 & 0 & 4 \\ 6 & 4 & 2 & 0 & 5 & 3 \\ 0 & 6 & 5 & 4 & 3 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$d(D) = 12 \qquad d(D_1) = 13$$

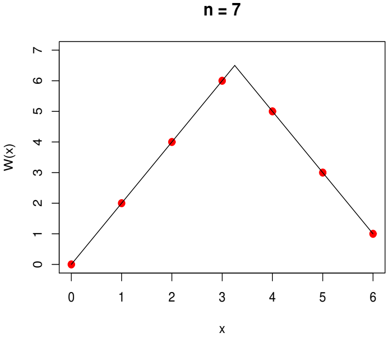
- After linear permutations, d_{eff} is about 90% for large n.

How About Nonlinear Permutations?

- Given an integer n, for $x = 0, \ldots, n-1$,

$$W(x) = \begin{cases} 2x, & \text{for } 0 \le x < n/2; \\ 2(n-x)-1, & \text{for } n/2 \le x < n. \end{cases}$$

- The W is a permutation of $\{0, \ldots, n-1\}$.



The W has been useful in

1. Latin squares, Williams (1949)
2. Orthogonal designs, Bailey (1982), Edmondson (1993)
3. Orthogonal LHDs under a second-order Fourier model, Butler (2001)

GLP + Williams Transformation

Algorithm (Wang, Xiao and Xu, 2018, Annals of Statistics)

Step 1. Generate an $n \times p$ GLP design D.

Step 2. For $b = 0, \dots, n-1$, generate $D_b = D + b \pmod{n}$.

Step 3. Let $E_b = W(D_b)$.

Step 4. Find the best E_b which maximizes $d(E_b)$.

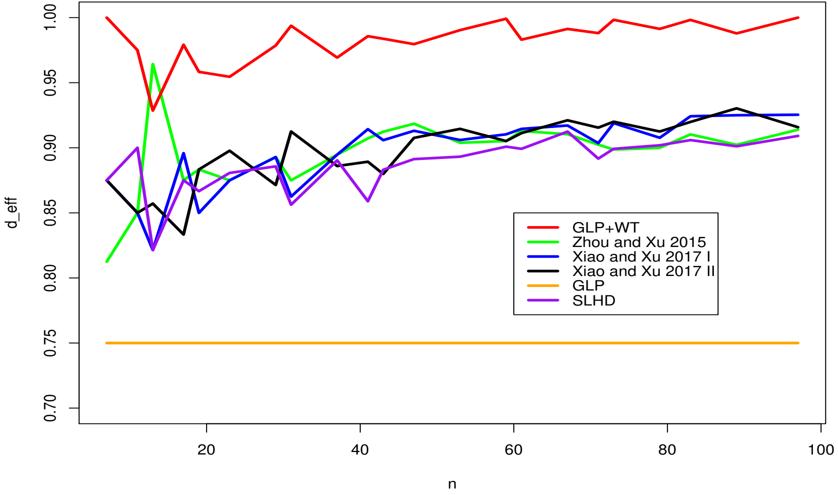
Example: $n = 7, b = 1$

$$D \quad \rightarrow \quad D_1 = D + 1 \pmod{n} \quad \rightarrow \quad E_1 = W(D_1)$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 1 & 3 & 5 \\ 3 & 6 & 2 & 5 & 1 & 4 \\ 4 & 1 & 5 & 2 & 6 & 3 \\ 5 & 3 & 1 & 6 & 4 & 2 \\ 6 & 5 & 4 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 0 \\ 3 & 5 & 0 & 2 & 4 & 6 \\ 4 & 0 & 3 & 6 & 2 & 5 \\ 5 & 2 & 6 & 3 & 0 & 4 \\ 6 & 4 & 2 & 0 & 5 & 3 \\ 0 & 6 & 5 & 4 & 3 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 4 & 6 & 5 & 3 & 1 & 0 \\ 6 & 3 & 0 & 4 & 5 & 1 \\ 5 & 0 & 6 & 1 & 4 & 3 \\ 3 & 4 & 1 & 6 & 0 & 5 \\ 1 & 5 & 4 & 0 & 3 & 6 \\ 0 & 1 & 3 & 5 & 6 & 4 \\ 2 & 2 & 2 & 2 & 2 & 2 \end{pmatrix}$$

$d(D) = 12$ $\quad$ $d(D_1) = 13$ $\quad$ $d(E_1) = 16 \ (= d_{upper})$

Comparison of Various $n \times (n-1)$ LHDs



Key Result

Let

$$b = W^{-1}\left(\frac{n-1}{2} \pm c\right)$$

where $c = \lfloor \sqrt{(n^2-1)/12} \rfloor$.

Theorem 3

Given a prime n and $p = n-1$, such defined b leads the best E_b, with

$$d_{eff}(E_b) \geq 1 - 2/\sqrt{3(n^2-1)}.$$

As $n \to \infty$, $d_{eff}(E_b) \to 1$.

- No need for computer search: $D \to D_b \to E_b = W(D_b)$
- Guaranteed efficiency
- Larger n, better design

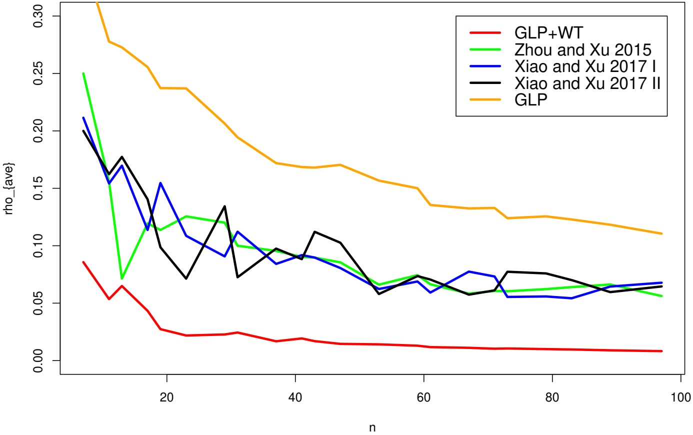
Correlations

For any $n \times p$ design $D = (x_{ij})$, define

$$\rho_{ave}(D) = \frac{\sum_{j \neq k} |\rho_{jk}|}{p(p-1)},$$

where ρ_{jk} is the correlation between columns j and k of D.

Comparison of ρ_{ave} for $n \times (n-1)$ Designs



Method I: $\rho_{ave}(E_b) < 2/(n-2)$

Summary

- Many available algorithms for constructing space-filling designs, but not efficient for constructing large designs
- A breakthrough — Wang, Xiao and Xu (2018) constructed maximin distance designs via good lattice points and a nonlinear transformation **without computer search**.
    - Large distance efficiencies: $d_{eff} = 1$ (or $\to 1$)
    - Low average correlation $\rho_{ave} \to 0$ as $n \to \infty$
    - Asymptotically optimal under the uniform projection criterion.
- A new class of space-filling designs — Uniform projection designs
    - suitable when only a subset of the input variables are active.
    - good space-filling not only in two dimensions, but also in all dimensions.
    - equivalent to maximin L_1-distance criterion if L_1-equidistant designs exist.
    - robust under other design criteria.
- There are still many open problems to be investigated.

Selected References

- Fang, K. T., Li, R. Z. and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*. Chapman and Hall/CRC, New York.
- Johnson, M. E., Moore, L. M. and Ylvisaker, D. (1990). Minimax and maximin distance designs. *J. Statist. Plan. Infer.* **26**, 131–48.
- Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical science*, **4**, 409–423.
- Santner, T. J., Williams, B. J., and Notz, W. (2003). *The design and analysis of computer experiments*. Springer.
- Sun, F., Wang, Y. and Xu, H. (2019), "Uniform Projection Designs," *Annals of Statistics*, 47, 641–661.
- Wang, L., Xiao, Q. and Xu, H. (2018), "Optimal Maximin L1-distance Latin Hypercube Designs Based on Good Lattice Point Designs," *Annals of Statistics*, 46, 3741-3766.
- Xiao, Q., Wang, L. and Xu, H. (2019). "Application of Kriging Models for a Drug Combination Experiment on Lung Cancer," *Statistics in Medicine*, 38, 236–246.
- Zhou, Y. D., and Xu, H. (2015). Space-filling properties of good lattice point sets. *Biometrika*, **102**, 959–966.