Design and Analysis for Computer Experiments with Qualitative and Quantitative Factors

Xinwei Deng

Department of Statistics, Virginia Tech

Outline

- Brief Introduction to Computer Experiments
- Marginally Coupled Design for Computer Experiments with Qualitative and Quantitative Factors
  - Review of designs for computer experiments
  - The proposed designs and their properties, constructions
- Additive Gaussian Process for Computer Experiments with Qualitative and Quantitative Factors
  - Review of the existing methods
  - The proposed model
  - Simulation and case study

Computer Experiments (CE)

- Many physical processes are difficult, expensive or impossible to observe
- Computer code exists to model the physical processes
- The computer code takes input settings **x** and produces a response $y(\mathbf{x})$

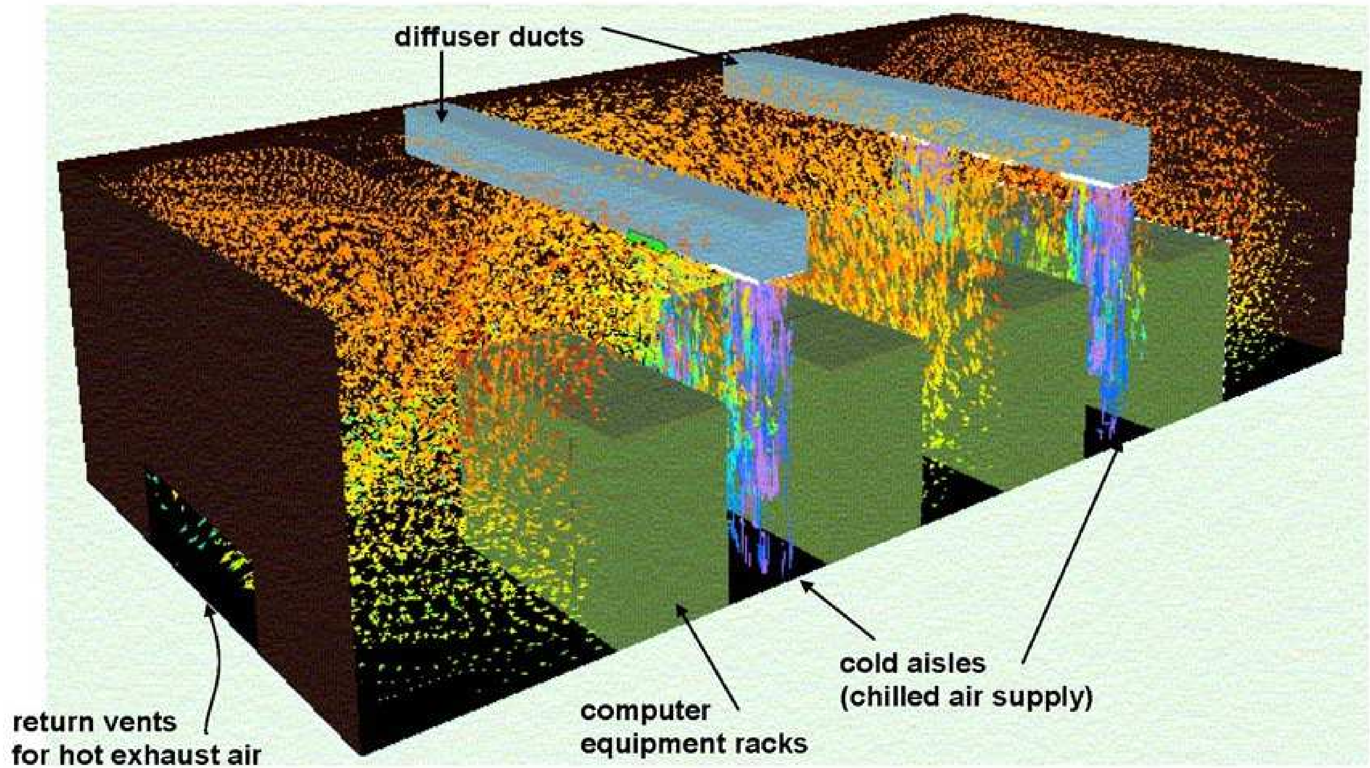
$$\mathbf{x} \longrightarrow \boxed{\textbf{Code}} \longrightarrow y(\mathbf{x})$$

- $y(\mathbf{x})$ is often deterministic; the computer code is often expensive to run

continuous
&
discrete

CE with Qualitative and Quantitative Factors

Computational Fluid Dynamics (CFD) Based
Computer Experiment



Data Center Thermal Management @2007 IBM Corporation

Configuration Variables

The configuration variables contain both qualitative and quantitative factors.

Variable	Description	Values
X_1	CRAC unit 1 flow rate (cfm)	0 7000 8500 10000 11500 13000
X_2	CRAC unit 2 flow rate (cfm)	0 7000 8500 10000 11500 13000
X_3	CRAC unit 3 flow rate (cfm)	0 2500 4000 5500
X_4	CRAC unit 4 flow rate (cfm)	0 2500 4000 5500
X_5	Room temperature (F)	65 67 69 71 73 75
X_6	Tile distribution (location)	Layout1 Layout2 Layout3
X_7	Tile percentage open area	$(0,1)$

Data Center Thermal Management @2007 IBM Corporation

Design for Computer Experiments

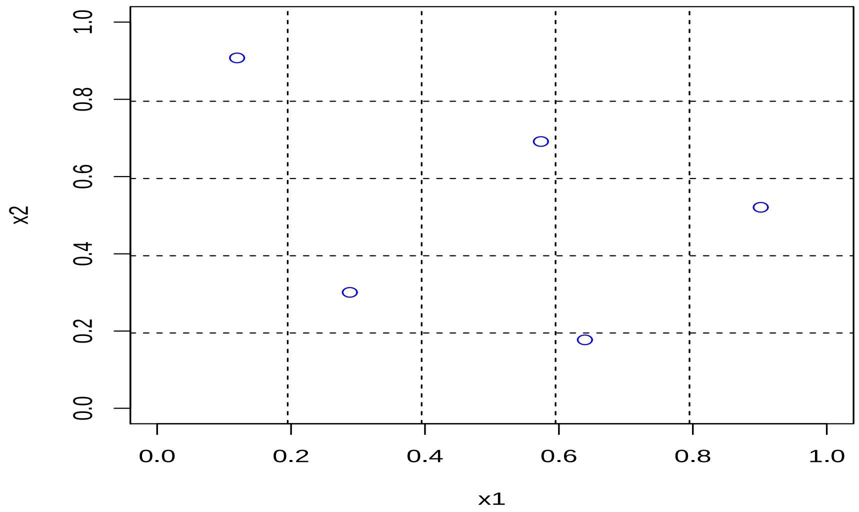
- Space-filling designs
  - Latin hypercubes and their generalizations
  - Designs based on distances between points (Maximin; Minimax)
  - Uniform designs, etc
- Designs with good low-dimensional projection properties
- Sequential designs (for optimization, sensitivity analysis, contour estimation, quantile estimation, global fitting)

Design for CE with Mixed Inputs

How to construct a good design for computer experiments with mixed inputs (both qualitative and quantitative factors)

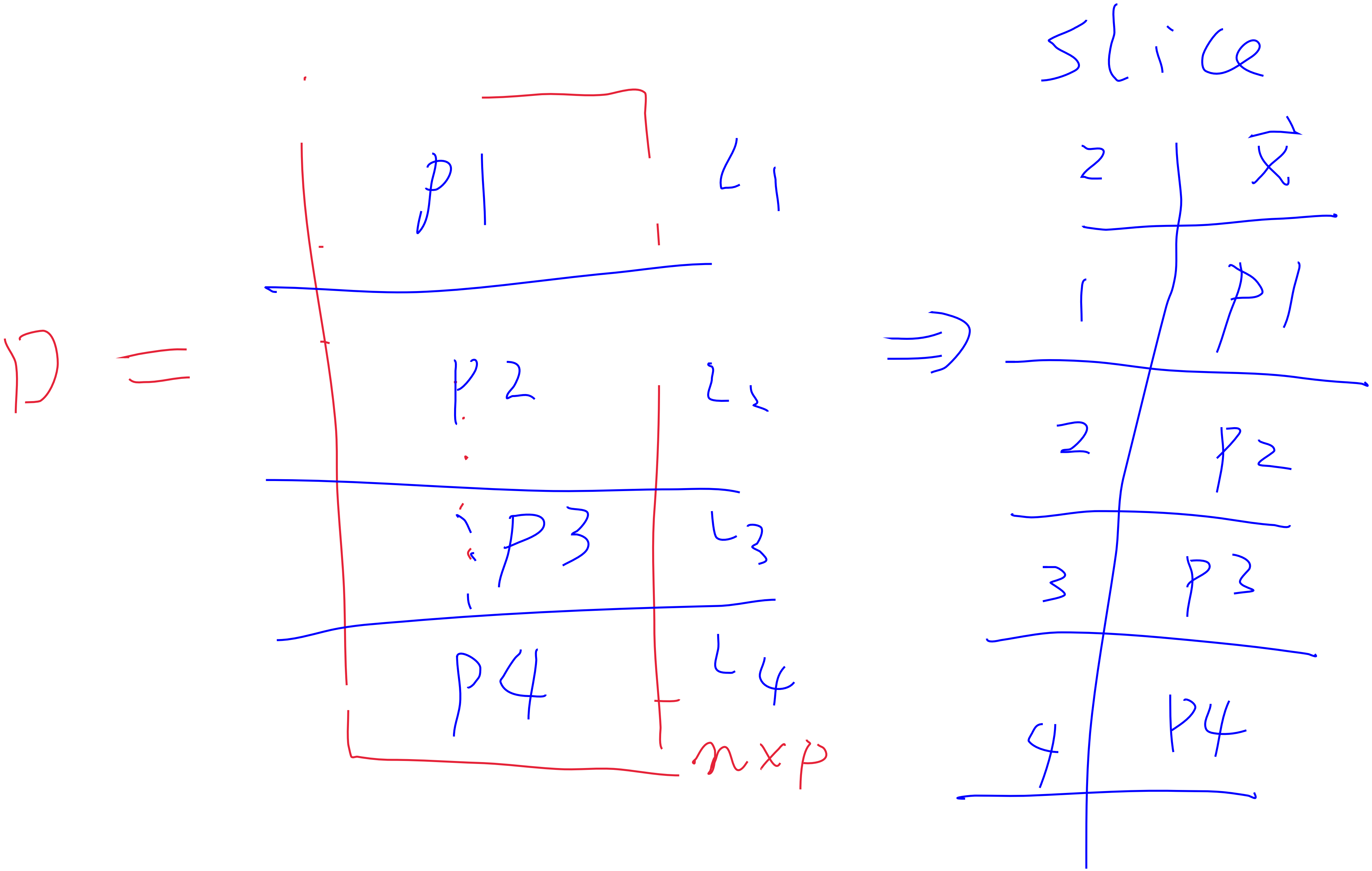
Latin Hypercube Design (McKay et al., 1979)

$$D = \begin{bmatrix} -1 & -1 \\ -2 & 2 \\ 0 & 1 \\ 2 & 0 \\ 1 & -2 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0.29 & 0.30 \\ 0.12 & 0.91 \\ 0.57 & 0.69 \\ 0.90 & 0.52 \\ 0.64 & 0.18 \end{bmatrix}$$



For convenience, we use $-(n-1)/2, -(n-3)/2, \ldots, (n-3)/2, (n-1)/2$ to represent the n levels in a Latin hypercube of n runs.

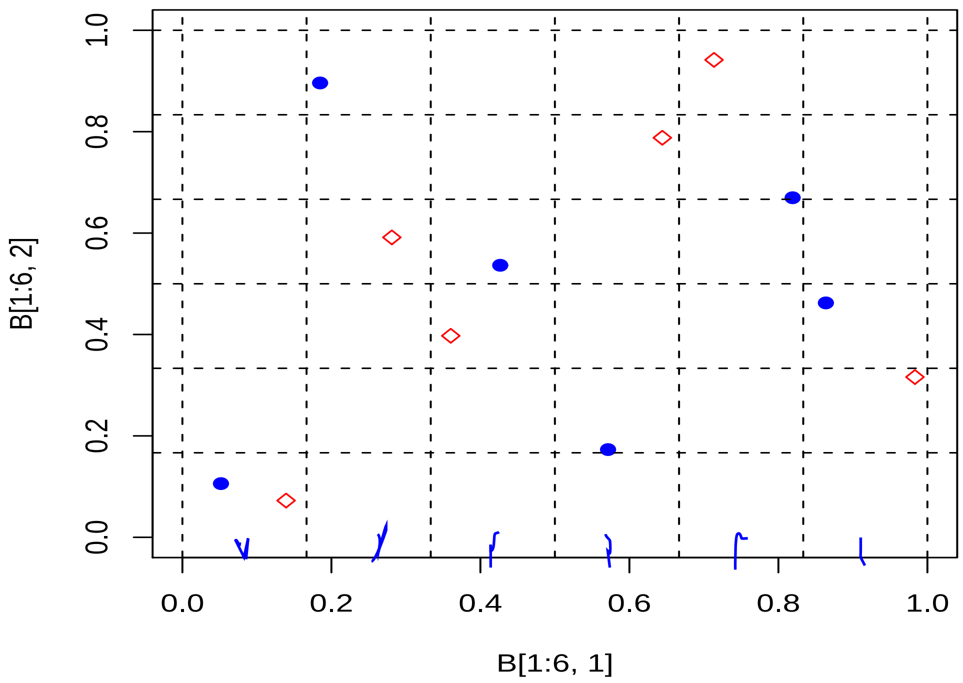
1d - projection property


D =
p1
p2
p3
p4
L1
L2
L3
L4
n×p
slice
z
x
1
2
3
4
p1
p2
p3
p4

Sliced Latin Hypercube Designs (Qian, 2012)

Sliced latin hypercube designs (SLHD): a special Latin hypercube design that can be partitioned into slices of smaller Latin hypercube designs.

$$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \left[\begin{array}{rr} 0 & 1 \\ -5 & -4 \\ 4 & 3 \\ 5 & 0 \\ 1 & -3 \\ -3 & 5 \\ \hdashline 3 & 6 \\ 2 & 4 \\ -2 & 2 \\ 6 & -2 \\ -4 & -5 \\ -1 & -1 \end{array}\right] \Longrightarrow \left[\begin{array}{cc} 0.42 & 0.53 \\ 0.05 & 0.10 \\ 0.81 & 0.66 \\ 0.86 & 0.46 \\ 0.57 & 0.17 \\ 0.18 & 0.89 \\ \hdashline 0.71 & 0.94 \\ 0.64 & 0.78 \\ 0.28 & 0.59 \\ 0.98 & 0.31 \\ 0.13 & 0.07 \\ 0.36 & 0.39 \end{array}\right]$$



Using SLHD for CE with Mixed Inputs

- A sliced Latin hypercube design is used for quantitative factors
- A factorial design is used for qualitative factors
- Each slice of a sliced Latin hypercube design corresponds to each level combination of qualitative variables.

Z_1	Z_2	
0	0	B_1
0	1	B_2
1	0	B_3
1	1	B_4

$$B = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{bmatrix}$$

  - It can be useful when the number of qualitative factors is small.
  - Such a design needs large run size when the number of qualitative factors is moderate or large.

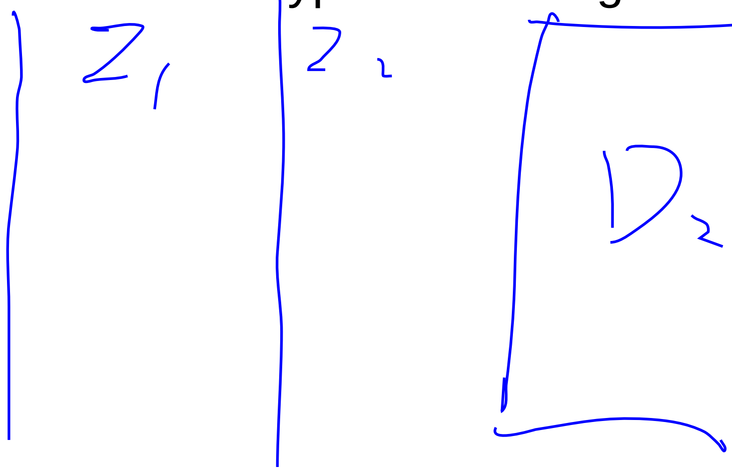
0	1	LHD
1	0	LHD
0	0	LHD
1	1	LHD

The Proposed Marginally Coupled Design

- Consider a computer experiment with q qualitative factors and p quantitative variables. Suppose that the ith qualitative factor has s_i levels, $1 \leq i \leq q$.

- Let D_1 and D_2 be the design matrices for qualitative factors and quantitative factors, respectively.

- A design $\boxed{D = (D_1, D_2)}$ is called a marginally coupled design if D_2 is a Latin hypercube design and the rows in D_2 corresponding to each level of any factor in D_1 form a small Latin hypercube design.



MCD

z_1	z_2	X
0	0	
0	1	
1	0	
1	1	

Example

discrete D_1 | D_2 continuous

z_1	z_2	x_1	x_2
0	0	-2	-2
0	1	-1	4
0	2	3	-1
1	0	0	1
1	1	2	-3
1	2	-3	2
2	0	4	3
2	1	-4	0
2	2	1	-4

Illustration

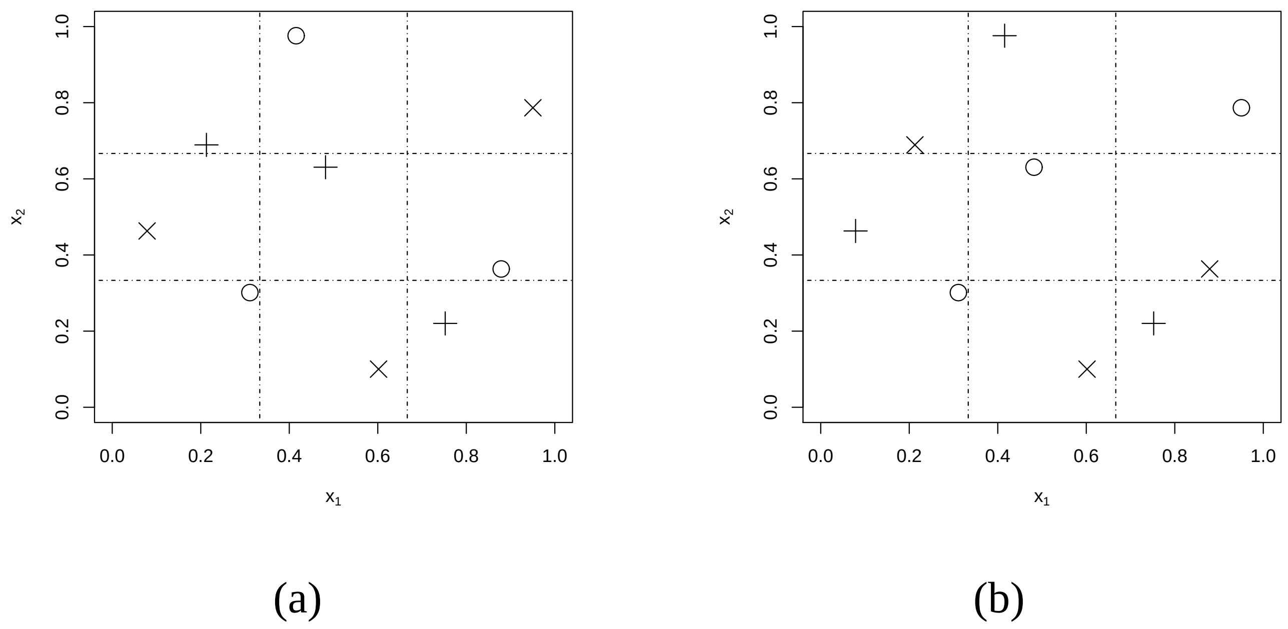


Figure 1: Scatter plots of x_1 versus x_2 where rows of D_2 corresponding to levels 0,1,2 of z_i are marked by $\times$, $\circ$, and $+$: (a) the levels of z_1; (b) the levels of z_2.

Orthogonal Array

An orthogonal array D of strength t, denoted by $\text{OA}(n, s_1 \cdots s_k, t)$, is an $n \times k$ matrix of which the ith column has s_i levels $0, \ldots, s_i - 1$ and for every $n \times t$ submatrix of D, each of all possible level combinations appears equally often. If not all s_i's are equal, an orthogonal array is mixed. Otherwise it is called symmetric. (Hedayat et al., 1999)

strength

$\text{OA}(9, 3^4, 2)$

0	0	0	0
0	1	1	2
0	2	2	1
1	0	1	1
1	1	2	0
1	2	0	2
2	0	2	2
2	1	0	1
2	2	1	0

$\text{OA}(8, 2^4 3^1, 2)$

0	0	0	0	0
1	1	1	1	0
0	0	1	1	1
1	1	0	0	1
0	1	0	1	2
1	0	1	0	2
0	1	1	0	3
1	0	0	1	3

Resolvable Orthogonal Array

An $OA(n, s_1^{q_1} \cdots s_k^{q_k}, 2)$ is said to be $(\alpha_1 \times \alpha_2 \times \cdots \times \alpha_k)$ resolvable if for $1 \leq j \leq k$, its rows can be partitioned into $n/(\alpha_j s_j)$ subarrays $A_1, \ldots, A_{n/(\alpha_j s_j)}$ of $\alpha_j s_j$ rows each such that each of $A_1, \ldots, A_{n/(\alpha_j s_j)}$ is an $OA(\alpha_j s_j, s_1^{q_1} \cdots s_k^{q_k}, 1)$. If $\alpha_1 = \cdots = \alpha_k = 1$, the orthogonal array is called completely resolvable.

Resolvable Orthogonal Array

$CROA(9, 3^3, 2)$

0	0	0
1	1	2
2	2	1
0	1	1
1	2	0
2	0	2
0	2	2
1	0	1
2	1	0

$\alpha = 1$

$CROA(16, 4^2 2^3, 2)$

0	2	1	1	1
3	1	0	0	1
2	0	1	0	0
1	3	0	1	0
3	0	0	1	0
0	3	1	0	0
1	2	0	0	1
2	1	1	1	1
0	0	0	0	1
3	3	1	1	1
1	1	1	0	0
2	2	0	1	0
0	1	0	1	0
1	0	1	1	1
3	2	1	0	0
2	3	0	0	1

$\alpha_1 = 1, \alpha_2 = 2$

Properties of Marginal Coupled Design

Proposition 1. *Given $D_1 = OA(n, s^q, 2)$, a marginally coupled design exists if and only if D_1 is a completely resolvable orthogonal array.*

Proposition 2. *Given $D_1 = OA(n, s_1^{q_1} s_2^{q_2}, 2)$ with $s_1 = \alpha_2 s_2$, a marginally coupled design exists if and only if D_1 is a $(1 \times \alpha_2)$-resolvable orthogonal array that can be expressed as*

$$\begin{pmatrix} A_{11} & A_{12} \\ \vdots & \vdots \\ A_{m1} & A_{m2} \end{pmatrix} \tag{1}$$

such that (A_{i1}, A_{i2}) is an $OA(s_1, s_1^{q_1} s_2^{q_2}, 1)$, where $m = n/s_1$, and for $1 \leq i \leq m$, the A_{i2} is completely resolvable.

Construction 1 (Tang, 1993)

A construction for s-level orthogonal arrays of s^2 runs

- Suppose an OA$(s_2, s_k, 2)$, say A, is available and D1 for qualitative factors is obtained by randomly taking q columns from A.

- Obtain a design, B, by randomly taking p columns from $A \backslash D_1$, where $q + p \leq k$. For each column of B, replace s positions with level $i-1$ by a random permutation of $\{(i-1)s+1\} - (s_2+1)/2, \ldots, \{(i-1)s+s\} - (s_2+1)/2$, for $1 \leq i \leq s$.

Example

$\text{OA}(9, 3^4, 2)$				D_1		D_2	
0	0	0	0	0	0	-2	-2
0	1	1	2	0	1	-1	4
0	2	2	1	0	2	4	0
1	0	1	1	1	0	0	1
1	1	2	0	1	1	2	-3
1	2	0	2	1	2	-3	3
2	0	2	2	2	0	3	2
2	1	0	1	2	1	-4	-1
2	2	1	0	2	2	1	-4

Construction 2

A construction for s-level orthogonal arrays of λs^2 runs, $\lambda > 1$

- Suppose an OA$\{\lambda s^2, s^k(\lambda s), 2\}$, say A, is available and D_1 for qualitative factors is obtained by randomly taking q columns from the first k columns of A.

- Denote the last column of A by a. For $1 \le j \le p$, let π_j be a random permutation of $\{0, \ldots, \lambda s - 1\}$ and $\pi_j(i)$ be the ith entry of π_j. Replace the s positions having level $\pi_j(i)$ in a by a random permutation of $\{(i-1)s+1\} - (\lambda s^2+1)/2, \ldots, \{(i-1)s+s\} - (\lambda s^2+1)/2$, for $1 \le i \le \lambda s$.

Construction 3

A construction for mixed orthogonal arrays

Let L be a marginal sliced Latin hypercube for $B = \text{OA}(n, s_1^{k_1} \cdots s_v^{k_v}, 2)$ and $C = (C_{ij})$ be an $u \times p$ matrix with $C_{ij} = \pm 1$. The following steps for construction are proposed.

I. Let $T = C \otimes L = (C_{ij} L)$ where $\otimes$ represents the usual Kronecker product.

II. Let $H = (H_{ij})$ be an $u \times k$ Latin hypercube, where $k = \sum_{i=1}^{v} k_i$. For $1 \leq j \leq pk, 1 \leq i \leq u$ and $1 \leq r \leq n$, obtain an $nu \times pk$ matrix M by letting the $[\{(i-1)n+r\}, j]$th entry $M_{\{(i-1)n+r\}j} = T_{\{(i-1)n+r\}j} + nH_{ij}$, where $T_{\{(i-1)n+r\}j}$ is the $[\{(i-1)n+r\}, j]$th entry of T.

Analysis of CE with Mixed Inputs

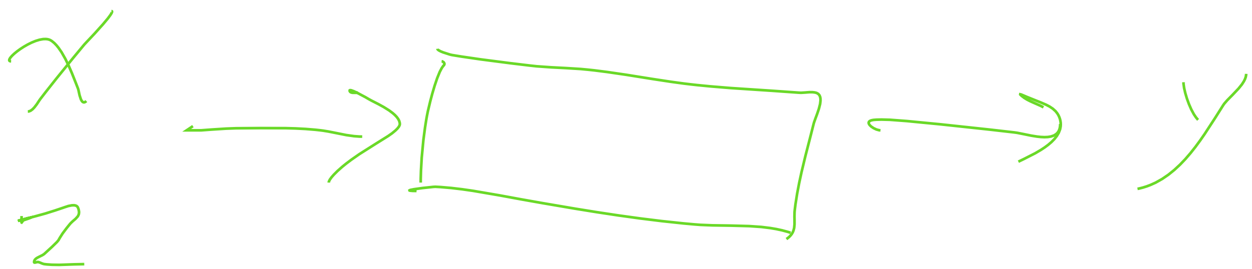
$$w_1 = \begin{pmatrix} x_1 \\ z_1 \end{pmatrix} \qquad w_2 = \begin{pmatrix} x_2 \\ z_2 \end{pmatrix}$$

$$\to d(x_1, x_2)$$

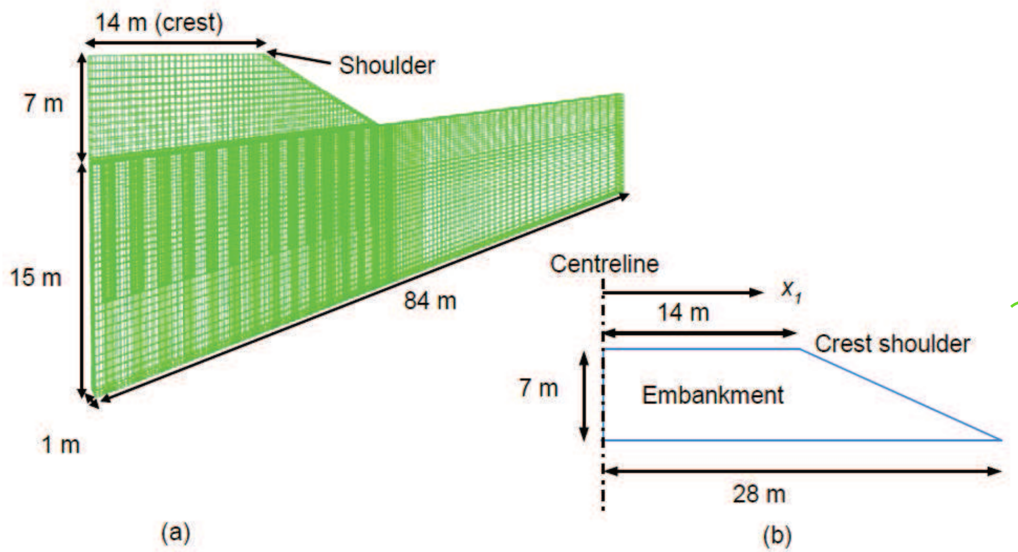
$$\to d(z_1, z_2)$$

$$\to d(w_1, w_2)$$

Additive Gaussian Process for Analyzing Computer Experiments with Qualitative and Quantitative Factors



A Motivating Example



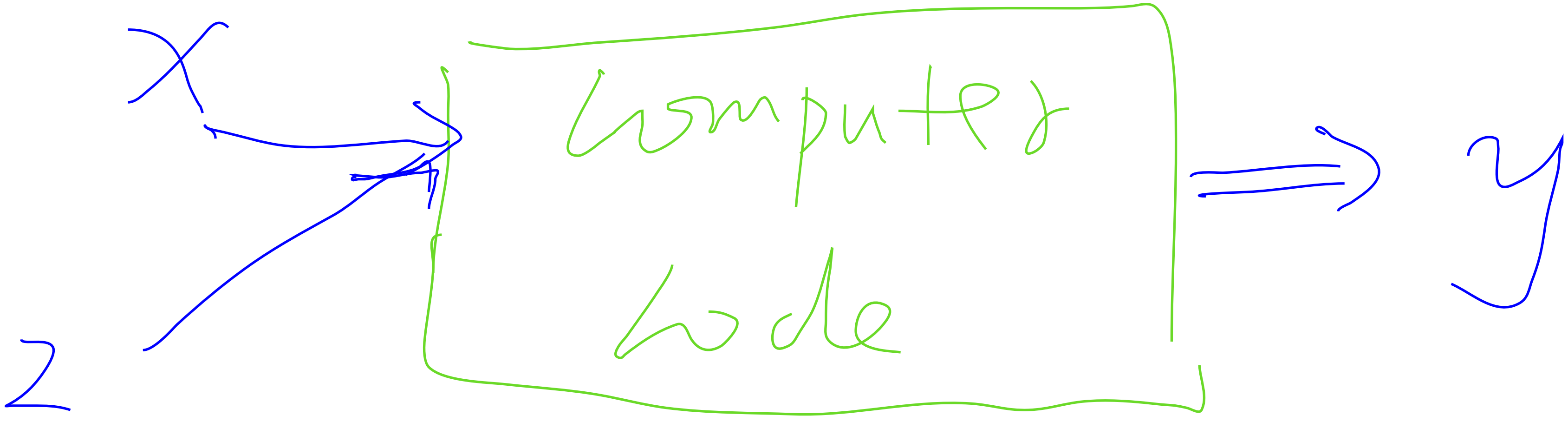
FEA

finite elements analysis

Figure 2: The embankment examined: (a) finite element mesh; (b) the schematic view of embankment constructed on foundation soil.

z_1	embankment construction rate	qualitative
z_2	Young's modulus of columns	qualitative
z_3	reinforcement stiffness	qualitative
x_1	the distance from the embankment centreline to the embankment shoulder	quantitative
Y	final embankment crest settlement	

continous



discrete

Analysis of CE with both Types of Inputs

- Qian, Wu and Wu (2008): use Gaussian process models with a restrictive correlation structure for the quantitative factors and employ optimization methods in the estimation to ensure that the correlation structure of the qualitative factors is positive-definite.
- Han et al. (2009): use hierarchical Bayesian Gaussian process models whose parameters at different levels to be i.i.d. draws from a prior distribution; perform well for cases where the responses have similar curvatures at different levels of the qualitative input variable.
- Zhou, Qian and Zhou (2011): use Gaussian process models with a correlation matrix being positive definite with unit diagonal elements.
- ...

Issue

- How to build an accurate surrogate model for computer experiments with qualitative and quantitative factors when
  - there can be a large number of qualitative factors.

$$x \longrightarrow \boxed{} \longrightarrow y$$

z

$$y = u + f(x, z_1) + f(x, z_2) + \cdots$$

effect

of z_1 . x

$$x \to \square \to y$$

$$y = u + G(x)$$

discrete

$$z \to \square \to y$$

$$y_{ij} = u + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon$$

ANOVA

Additive Gaussian Process

- Consider a computer experiment with the p quantitative factors $\mathbf{x} = (x_1, \ldots, x_p)$ and the q qualitative factors $\mathbf{z} = (z_1, \ldots, z_q)$.

z_1 . . . . z_q

- The proposed model is

$$Y(\mathbf{x}, z_1, \ldots, z_q) = \mu + G_1(z_1, \mathbf{x}) + \cdots + G_q(z_q, \mathbf{x}), \tag{2}$$

where μ is the overall mean, $G_j \sim \mathrm{GP}(0, \phi_j)$ and G_j's are independent.

similar idea to marginal couple design

Gaussian Process (GP) Models

$$\text{Model}: \quad Y(\mathbf{x}) = \sum_j \beta_j f_j(\mathbf{x}) + Z(\mathbf{x})$$

- $\mathbf{X} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix}$
- $Z(\mathbf{X}) = (Z(\mathbf{x}_1), \ldots, Z(\mathbf{x}_n))^T \sim MVN(0_n, \sigma_Z^2 \mathbf{R})$
- $R(x_i, x_j; \theta) = \prod_{k=1}^{p} \exp\{-\theta_k (x_{ik} - x_{jk})^2\}$ $= \exp\left\{-\sum \theta_k (x_{ik} - x_{jk})^2\right\}$

$+\infty$

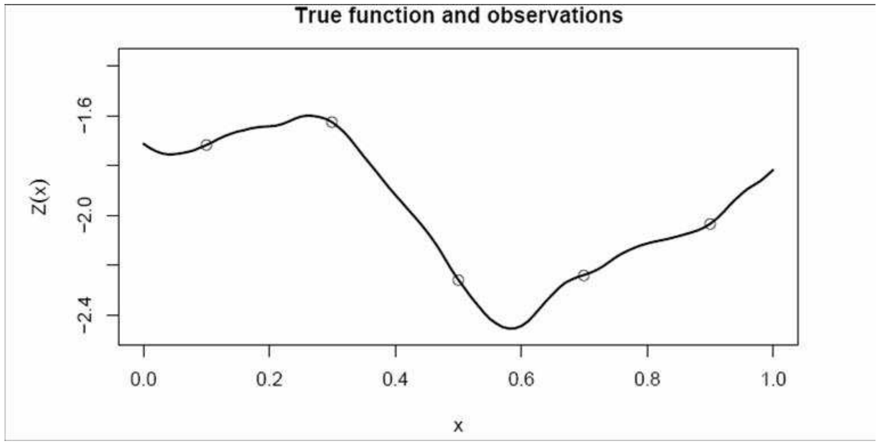
Gaussian Process (GP) Models

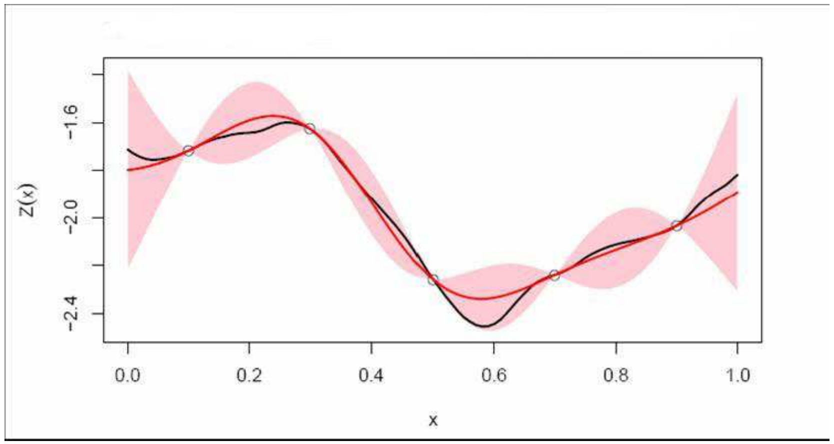
- Prediction:

$$\hat{y}(x_0) = \sum_j f_j(x_0)\hat{\beta}_j + r_0^T \hat{R}^{-1}(y - \sum_j f_j(\mathbf{x})\hat{\beta}_j)$$

$$s^2(x_0) = \sigma^2(1 - r_0^T \hat{R}^{-1} r_0 + \frac{1 - \mathbf{1}^T \hat{R}^{-1} r_0}{\mathbf{1}^T \hat{R}^{-1}\mathbf{1}})$$

where $r_0 = (\hat{R}(x_0 - x_1), \ldots, \hat{R}(x_0 - x_n))^T$





Additive Gaussian Process

- Consider a computer experiment with the p quantitative factors $\mathbf{x} = (x_1, \ldots, x_p)$ and the q qualitative factors $\mathbf{z} = (z_1, \ldots, z_q)$.

- The proposed model is

$$Y(\mathbf{x}, z_1, \ldots, z_q) = \mu + \boxed{G_1(z_1, \mathbf{x})} + \cdots + G_q(z_q, \mathbf{x}), \tag{3}$$

where μ is the overall mean, $G_j \sim \mathrm{GP}(0, \phi_j)$ and G_j's are independent.

Z_1	Z_2	
0	0	
0	1	
1	0	
1	1	

$\Rightarrow Y = \mu + G_1(z_1, x) + G_2(z_2, x$

z_1	z_2	z_3
0	0	0
0	1	1
1	0	1
1	1	0

2^{3-1}

$$y = u + f_1(z_1, x) + f_2(z_2, x) + f_3(z_3, x)$$

$$\Rightarrow z_1 = 1 \quad z_2 = 1 \quad z_3 = 1$$

Proposed Covariance Function

Proposed covariance:

$$\begin{aligned}\phi(Y(\mathbf{w}_1), Y(\mathbf{w}_2)) &= \mathrm{cov}(Y(\mathbf{x}_1, \mathbf{z}_1), Y(\mathbf{x}_2, \mathbf{z}_2)) \\ &= \sum_{j=1}^{q} \sigma_j^2 \mathrm{cor}(G_j(z_{1j}, \mathbf{x}_1), G_j(z_{2j}, \mathbf{x}_2)) \\ &= \sum_{j=1}^{q} \sigma_j^2 R(\mathbf{x}_1, \mathbf{x}_2 | \theta^{(j)}) \tau^{(j)}_{z_{1j}, z_{2j}}. \qquad (4)\end{aligned}$$

continuous

where

- $R(\mathbf{x}_1, \mathbf{x}_2 | \theta^{(j)})$ quantifies the correlation between inputs $\mathbf{x}_1$ and $\mathbf{x}_2$ associated with the jth qualitative factor.
- $\tau^{(j)}_{z_{1j}, z_{2j}}$ represents the correlation between level z_{1j} and level z_{2j} of the jth qualitative factor.

discrete

jth qualitative factor

z_j : 1, 2, 3.

$(1, 2) \longrightarrow \tau_{12}$

$(1, 3) \rightarrow \tau_{13}$

$(2, 3) \longrightarrow \tau_{23}$

$$T = \begin{bmatrix} 1 & \tau_{12} & \tau_{13} \\ \tau_{12} & 1 & \tau_{23} \\ \tau_{13} & \tau_{23} & 1 \end{bmatrix}$$ is a correlation matrix

Correlation Function for Qualitative Factors

- Suppose the jth qualitative factor having m_j levels, $j = 1, \ldots, q$.
- Let $\mathbf{T}_j = (\tau_{r,s}^{(j)})$ be a correlation matrix of the m_j levels of the jth qualitative factor, $j = 1, \ldots, q$.
- Model $\mathbf{T}_j$ such that $\mathbf{T}_j$ is a matrix with positive definiteness and unit diagonal element (PDUDE).

Key idea: cholesky decomposition

Step 1: find a lower triangular matrix with strictly positive diagonal entries $\mathbf{L}_j = (l_{r,s}^{(j)})$ through a Cholesky-type decomposition, that is, $\mathbf{T}_j = \mathbf{L}_j \mathbf{L}_j^T$ for $j = 1, \ldots, q$.

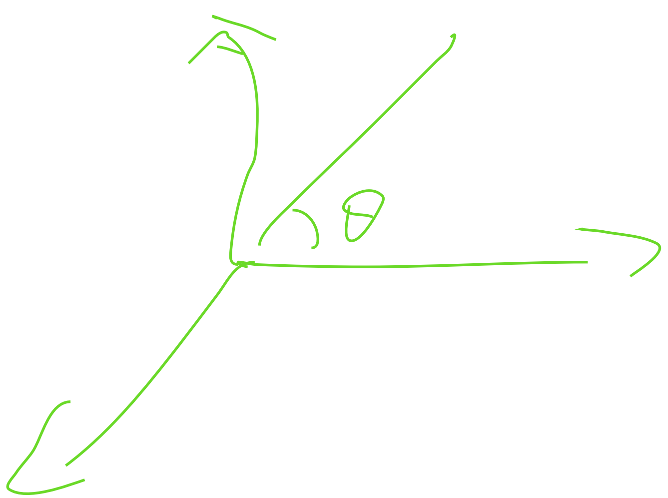
Step 2: each row vector $(l_{r,1}^{(j)}, \ldots, l_{r,r}^{(j)})$ in $\mathbf{L}_j$ is specified in the following way: for $r = 1$, $l_{j,1,1} = 1$ and for $r = 2, \ldots, m_j$,

$$\begin{cases} l_{r,1}^{(j)} = \cos(\varphi_{j,r,1}) \\ l_{r,s}^{(j)} = \sin(\varphi_{j,r,1}) \cdots \sin(\varphi_{j,r,s-1}) \cos(\varphi_{j,r,s}), \text{for } \ s = 2, \ldots, r-1 \\ l_{r,r}^{(j)} = \sin(\varphi_{j,r,1}) \cdots \sin(\varphi_{j,r,r-2}) \sin(\varphi_{j,r,r-1}), \end{cases}$$

where $\varphi_{j,r,s} \in (0, \pi)$ and $\tau_{r,r}^{(j)} = \sum_{s=1}^{r} (l_{r,s}^{(j)})^2 = 1$ for $r = 1, \ldots, m_j$.

$$T = L L^T$$

$$L = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$



Zhou et al. (2011)

$$Y(\mathbf{x}, z_1, \ldots, z_q) = \mu + G(\mathbf{x}, \mathbf{z})$$

where μ is the overall mean, $G \sim \mathrm{GP}(0, \phi)$.

- For any two inputs $\mathbf{w}_1 = (\mathbf{x}_1, \mathbf{z}_1)$ and $\mathbf{w}_2 = (\mathbf{x}_2, \mathbf{z}_2)$, Zhou et al. (2011) defined

$$\begin{aligned} \phi(Y(\mathbf{w}_1), Y(\mathbf{w}_2)) &= \sigma^2 \mathrm{cor}(Z(\mathbf{w}_1), Z(\mathbf{w}_2)) \\ &= \sigma^2 R(\mathbf{x}_1, \mathbf{x}_2 | \theta) \prod_{j=1}^{q} \tau^{(j)}_{z_{1j}, z_{2j}}, \end{aligned} \tag{5}$$

  where

  - $R(\mathbf{x}_1, \mathbf{x}_2 | \theta)$ quantifies the correlation between inputs $\mathbf{x}_1$ and $\mathbf{x}_2$,
  - $\tau^{(j)}_{z_{1j}, z_{2j}}$ represents the correlation between level z_{1j} and level z_{2j} of the jth qualitative factor.

Comparison of Covariance Functions

Proposed covariance function:

$$\begin{aligned}\phi(Y(\mathbf{w}_1), Y(\mathbf{w}_2)) &= \text{cov}(Y(\mathbf{x}_1, \mathbf{z}_1), Y(\mathbf{x}_2, \mathbf{z}_2)) \\ &= \sum_{j=1}^{q} \sigma_j^2 \text{cor}(G_j(z_{1j}, \mathbf{x}_1), G_j(z_{2j}, \mathbf{x}_2)) \\ &= \sum_{j=1}^{q} \sigma_j^2 R(\mathbf{x}_1, \mathbf{x}_2 | \theta^{(j)}) \tau_{z_{1j}, z_{2j}}^{(j)}. \qquad (6)\end{aligned}$$

Zhou, Qian and Zhou (2011):

$$\begin{aligned}\phi(Y(\mathbf{w}_1), Y(\mathbf{w}_2)) &= \sigma^2 \text{cor}(Z(\mathbf{w}_1), Z(\mathbf{w}_2)) \\ &= \sigma^2 R(\mathbf{x}_1, \mathbf{x}_2 | \theta) \prod_{j=1}^{q} \tau_{z_{1j}, z_{2j}}^{(j)}. \qquad (7)\end{aligned}$$

Parameter Estimation

- The log-likelihood function is

$$l(\mu, \theta, \sigma^2, \mathbf{T}) = -\frac{1}{2}\left[\log|\Phi(\theta, \sigma^2, \mathbf{T})| + (\mathbf{y} - \mu\mathbf{1})^T \boxed{\Phi^{-1}}(\theta, \sigma^2, \mathbf{T})(\mathbf{y} - \mu\mathbf{1})\right]. \quad (8)$$

Covariance matrix

- Estimation of μ given $(\theta, \sigma^2, \mathbf{T})$,

$$\hat{\mu} = (\mathbf{1}^T \Phi^{-1} \mathbf{1})^{-1} \mathbf{1}^T \Phi^{-1} \mathbf{y}. \quad (9)$$

- Estimation of $\theta, \sigma^2, \mathbf{T}$:

$$[\hat{\theta}, \hat{\sigma}^2, \hat{\mathbf{T}}] = \text{argmin} \frac{1}{2}\left[\log|\Phi(\theta, \sigma^2, \mathbf{T})| + (\mathbf{y}^T \Phi^{-1} \mathbf{y}) - (\mathbf{1}^T \Phi^{-1} \mathbf{1})^{-1} (\mathbf{1}^T \Phi^{-1} \mathbf{y})^2\right].$$

Prediction and Interpolation

- The prediction of y at a new location $\mathbf{w}_0$ is

$$\hat{y}(\mathbf{w}_0) = \hat{\mu} + \phi(\mathbf{w}_0)^T \Phi^{-1}(\hat{\theta}, \hat{\sigma}^2, \hat{\mathbf{T}})(\mathbf{y} - \hat{\mu}\mathbf{1}). \tag{10}$$

- When $\mathbf{w}_0 = \mathbf{w}_i$, the coefficient $\phi(\mathbf{w}_0)^T \Phi^{-1}(\hat{\theta}, \hat{\sigma}^2, \hat{\mathbf{T}})$ in (10) is an n-dimensional vector with the ith entry being 1 and otherwise 0. Thus, $\hat{y}(\mathbf{w}_0) = y_i$, achieving the property of interpolation.

Viewed as a Sequential Predictor

Sequential prediction

$$\hat{y}_{seq}(\mathbf{w}_0) = \sum_{j=1}^{q} \hat{y}_j(\mathbf{w}_0). \quad (11)$$

where

$$\hat{y}_1(\mathbf{w}_0) = \hat{\mu} + \phi_1^{\mathsf{T}}(\mathbf{w}_0)\Big(\sum_{j=1}^{q} \Phi_j\Big)^{-1}(\mathbf{y} - \hat{\mu}\mathbf{1}), \quad \mathbf{e}_1 = \mathbf{y} - \hat{\mathbf{y}}_1;$$

$$\hat{y}_2(\mathbf{w}_0) = \phi_2^{\mathsf{T}}(\mathbf{w}_0)\Big(\sum_{j=2}^{q} \Phi_j\Big)^{-1}\mathbf{e}_1, \quad \mathbf{e}_2 = \mathbf{y} - \hat{\mathbf{y}}_1 - \hat{\mathbf{y}}_2;$$

$$\vdots$$

$$\hat{y}_q(\mathbf{w}_0) = \phi_q^{\mathsf{T}}(\mathbf{w}_0)\Phi_q^{-1}\mathbf{e}_{q-1}.$$

Joint prediction:

$$\hat{y}(\mathbf{w}_0) = \hat{\mu} + (\phi_1(\mathbf{w}_0) + \cdots + \phi_q(\mathbf{w}_0))^T (\sum_{j=1}^{q} \Phi_j)^{-1} (\mathbf{y} - \mu\mathbf{1}), \tag{12}$$

where $\Phi_j = \sigma_j^2 \mathbf{R}_j \circ \mathbf{H}_j$, $\mathbf{R}_j = \left(R(\mathbf{x}_i, \mathbf{x}_{i'} | \theta^{(j)}) \right)_{n \times n}$ and $\mathbf{H}_j = \left(\tau^{(j)}_{z_{ij}, z_{i'j}} \right)_{n \times n}$.

Theorem 1. *With the same parameter values, the "joint prediction" in (12) and "sequential prediction" in (11) are equivalent.*

Simulation Study

Four correlation functions for qualitative factors are compared.

- **EC**: the exchangeable correlation function $\tau_{r,s} = c$ $(0 < c < 1)$ for $r \neq s$ (Joseph and Delaney, 2007; Qian et al., 2008);
- **MC**: the multiplicative correlation function $\tau_{r,s} = \exp\{-(\theta_r + \theta_s)\}$ $(\theta_r > 0, \theta_s > 0)$ for $r \neq s$ (McMillian et al. 1999; Qian et al., 2008);
- **UC**: the unrestrictive correlation function $\tau_{r,s}$ in Zhou, Qian and Zhou (2011);
- **AD**: the proposed correlation function.

$$T = \begin{pmatrix} 1 & c & c \\ c & 1 & c \\ c & c & 1 \end{pmatrix}$$

The four methods are evaluated using the root mean square error (RMSE) given by

$$\text{RMSE} = \sqrt{\frac{1}{|\mathcal{W}_{pred}|} \sum_{\mathbf{w} \in \mathcal{W}_{pred}} (\hat{y}(\mathbf{w}) - y(\mathbf{w}))^2}, \tag{13}$$

where $\hat{y}(\mathbf{w})$ and $y(\mathbf{w})$ are the predicted response and the true response at the new input $\mathbf{w}$ in the hold-out set $\mathcal{W}_{pred}$.

Training data

model estimation

Test data

model evaluation

Numerical Example

Consider the computer code with $p = 7$ quantitative factors ($x_1, \ldots, x_p$) and $q = 6$ qualitative factors ($x_{p+1}, \ldots, x_{p+q}$),

$$y = \sum_{i=1}^{p} \exp\{-x_i\} \cos(4x_{p+q-i}) \sin(4x_{i+2}), \tag{14}$$

where $0 < x_i < 1$ for $i = 1, \ldots, p$, $x_j = \{0.3, 0.8\}$ for $j = p+1, p+2$, $x_j = \{0.1, 0.5, 0.9\}$ for $j = p+3, p+4, p+5$ and $x_{p+6} = \{0.05, 0.35, 0.55, 0.95\}$.

Training data (for estimation):

- two replications of a 72-run mixed-level orthogonal array are used for qualitative factors
- a random Latin hypercube design of 144 runs for quantitative factors

Test data (for prediction):

- five replications of a full factorial design for qualitative factors
- a random Latin hypercube design of 2160 runs for quantitative factors

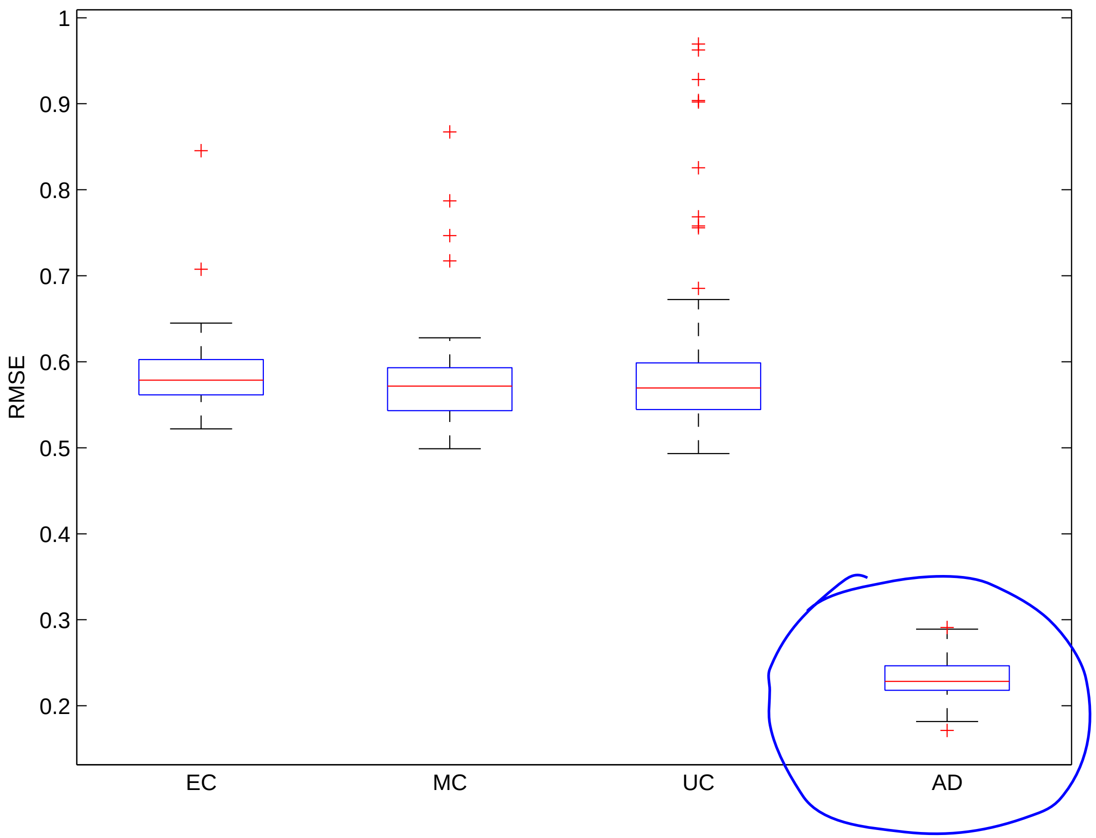


Figure 3: Boxplots of the RMSEs associated with 'EC', 'MC', 'UC' and 'AD' for the computer model in (14) over 100 simulations.

Real Application

Each of the three qualitative factors z_1, z_2, z_3 has three levels: the levels of z_1 are 1, 5, 10 m/month; the levels of z_2 are 50, 100, 200 MPa; and the levels of z_3 are 1578, 4800, 8000 kN/m.

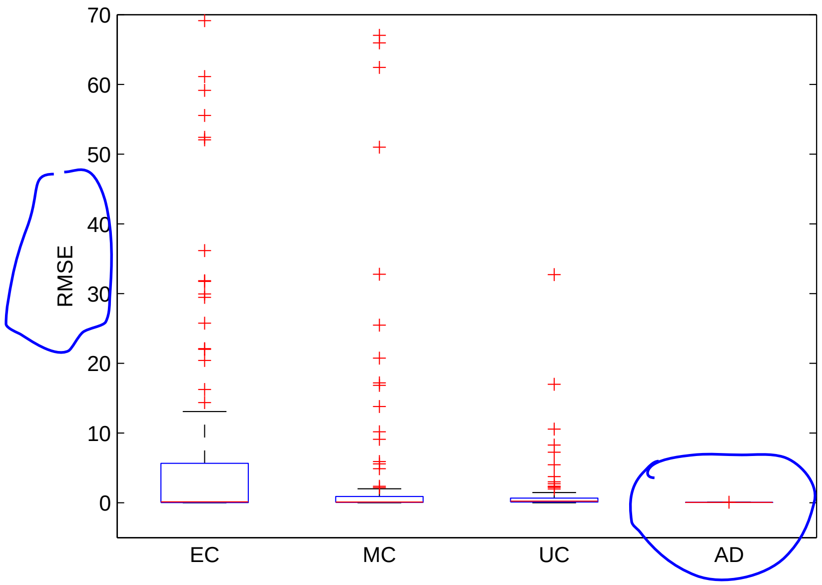
- Training data: The quantitative factor x_1 takes the 29 values uniformly from the interval $[0, 14]$.

  For each value of the quantitative factor, a three-level fractional factorial design of 9 runs is used for the qualitative factors.

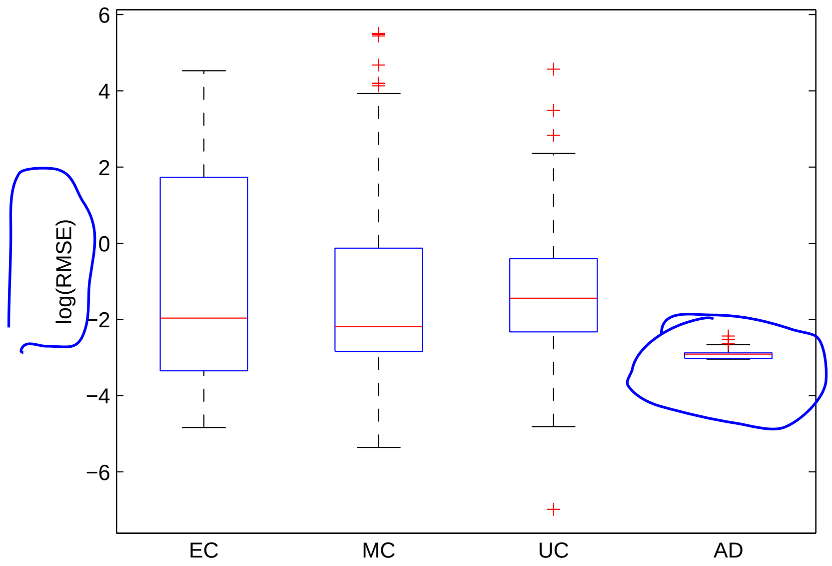
- Test data: The test data set contains 29 input settings in which the values of quantitative factor x_1 are taken uniformly from the interval $[0, 14]$, and the setting of the qualitative factors is $(z_1, z_2, z_3) = (5, 100, 4800)$.

- Make prediction at randomly chosen 20 input settings out of those 29 ones and compute the corresponding RMSE. Repeat this procedure 100 times.

Real Application



(a)



(b)

Figure 4: Boxplots of the RMSEs associated with ‘EC’, ‘MC’, ‘UC’ and ‘AD’ for the fully 3D coupled finite element model over 100 randomly chosen prediction sets of 20 input settings: (a) RMSE; (b) logarithm of RMSE.

Summary

- Design and analysis of computer experiments with both quantitative and qualitative factors are challenging.
- The proposed marginally coupled design enables that the design for quantitative factors is a sliced Latin hypercube design with respect to each qualitative factor.
- The proposed design is useful when there are a large number of qualitative factors.
- We propose a new model, additive Gaussian process model, for prediction with better accuracy.

Thank You!