Introduction to Computer Experiments

Po-Hsu Allen Chen

Computer Experiments Journal Club

The Ohio State University

09/22/2015

Outline

Introduction

What is computer experiments?

Introduction

What is computer experiments?

- Many physical experiments which cannot be implemented practically could only be studied by complex computer codes.
  - The output of a computer simulator relating x and $y(x)$ can be viewed as a black-box process

$$x \rightarrow \text{Simulators} \rightarrow y(x)$$

Introduction

What is computer experiments?

- Many physical experiments which cannot be implemented practically could only be studied by complex computer codes.
  - The output of a computer simulator relating x and $y(x)$ can be viewed as a black-box process

$$x \rightarrow \text{Simulators} \rightarrow y(x)$$

Is computer experiments popular?

- Personal opinion: computer experiments is not super popular in this big data generation but have become increasingly popular in the past two decades.
- References: Santner et al. (2003), and Jones and Johnson (2009)

Industrial Application

- The Langley glideback booster (Gramacy and Lee, 2008)

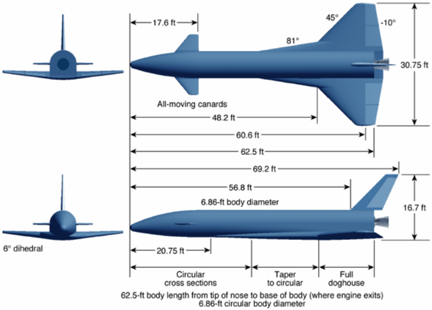


Figure 1: The glideback booster (NASA, Pamadi et al. 2014)

- The booster is being developed through the computer simulators.
- The primary goal is to model the lift force as a function of speed, the angle of attack, and the sideslip angle.

Environmental Application

- Dynamics of Ice Sheets and Glaicers (Greve 2004)

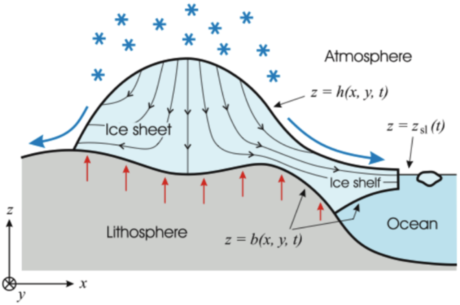


Figure 2: Ice Sheet Geometry (Course Note, Matthew Pratola 2013)

Biomechanical Application

- A finite element (FE) computational simulator of the dynamic tibial contact stress in gait (Leatherman et al. 2014)

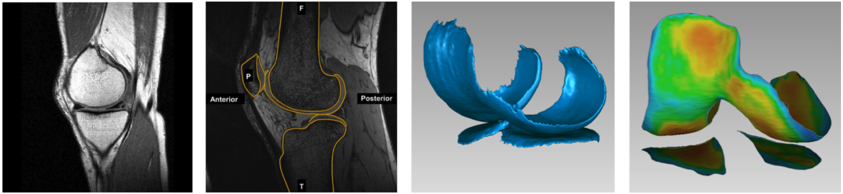


Figure 3: 3D FE simulator (HSS, Guo 2014)

- 60 points from the FE simulator
- 50 kinds of material properties
- Consider 10,000 different population properties in each kind of material properties
- Find the best material design!

Injection Molding Example

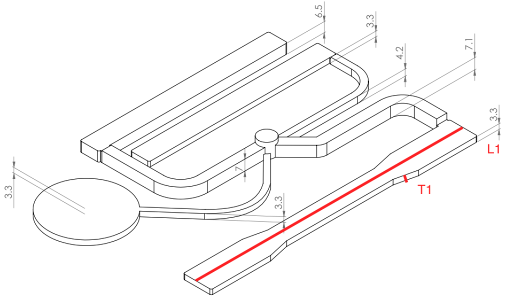


Figure 4: Injection Molding

Joint work with Jose Castro, Rachmat Mulyana, Maria Villarreal-Marroquin, Thomas Santner, and Angela Dean

Factors of Injection Molding

A computer simulator

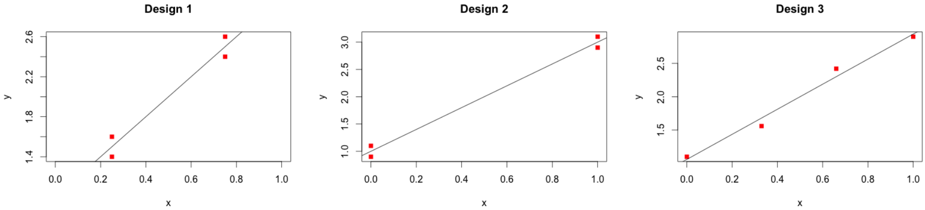
- 7 factors: Melting time, Packing time, Packing pressure, Cooling time, Heat transfer coefficient (HTC) flow, HTC pack, and HTC open
- 35 simulator data
- Some interested outputs: Length, Thickness, and Volume.

Outline

Design of Experiments

How to select design points?

- Physical experiments
  - George Box: Block what you can, randomize what you cannot
  - Reduce the variance: Suppose we want to fit a simple regression and we know the range of independent variable is $[0, 1]$. Which design will you pick?



Hint: $Var(\hat{\beta}) = (X^T X)^{-1}\sigma^2$

Design of Computer Experiments

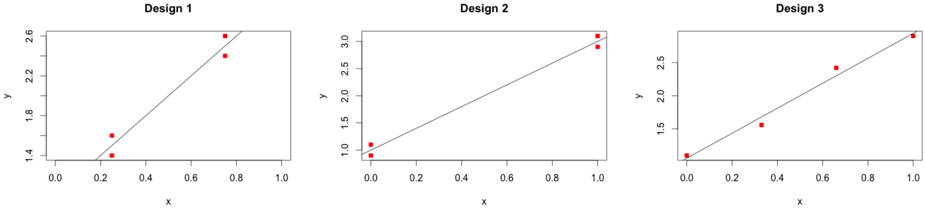
- The code is deterministic: No random error.
  - Designs should not take more than one observation at any set of inputs (no replication)

Design of Computer Experiments

- The code is deterministic: No random error.
  - Designs should not take more than one observation at any set of inputs (no replication)
- All input factors are known: No need for randomization and blocking.
  - Designs should provide information about all portions of the experimental region (space filling)

Design of Computer Experiments

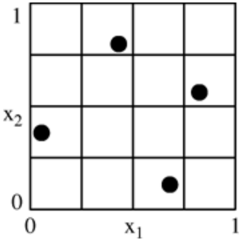
- The code is deterministic: No random error.
  - Designs should not take more than one observation at any set of inputs (no replication)
- All input factors are known: No need for randomization and blocking.
  - Designs should provide information about all portions of the experimental region (space filling)
- Suppose we want to select design points for a computer experiment. Which design will you pick?



Latin Hypercube Design

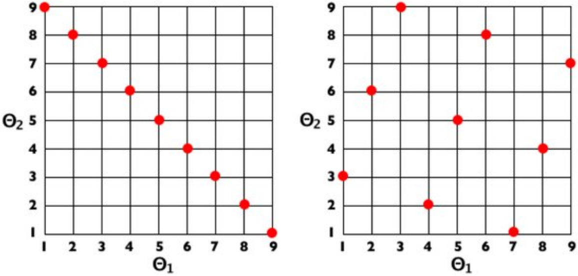
McKay et al. (1979) introduced the Latin Hypercube Design (LHD) which is the foundation of almost all design issues in computer experiments

- Experimental region is the unit square $[0,1]^2$ and a design consists 4 points: Divide each axis $[0,1]$ into the 4 equally spaced interval
- An arrangement in which each integer appears exactly once in each row and in each column
- Select a point at random



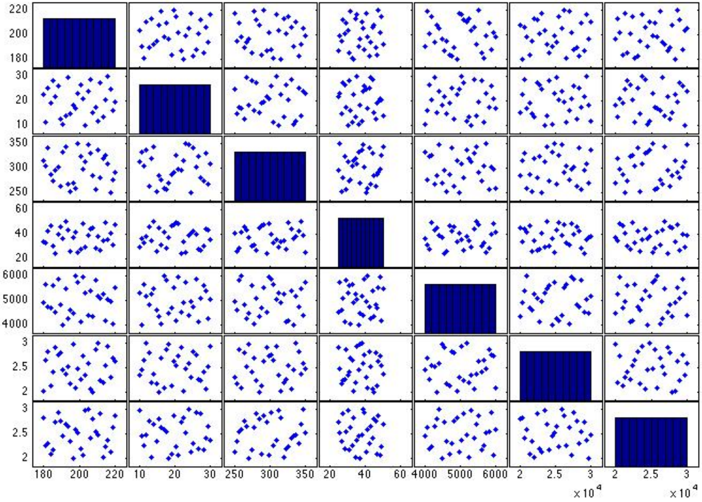
Maximin LHD

Not all LHDs are good!



Morris and Mitchell (1995) proposed to find the best LHD by maximizing the minimum distance between the points

Design for Injection Molding (2D)



Outline

Models for Computer Experiments

- Computer simulators: Time-consuming
  - Use the statistical model as a surrogate for the computer simulator

Models for Computer Experiments

- Computer simulators: Time-consuming
  - Use the statistical model as a surrogate for the computer simulator
- Gaussian Process (GP) Model (Kriging)
  View $\boldsymbol{y}(\boldsymbol{x})$ as a realization of the random function

  $$\boldsymbol{Y}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x})^T\beta + \boldsymbol{Z}(\boldsymbol{x}),$$

  where β is the unknown constant vector (regression parameter), the $\boldsymbol{F}$ is a known regression function, and $\boldsymbol{Z}(\boldsymbol{x})$ is a mean zero, second-order stationary Gaussian process, and

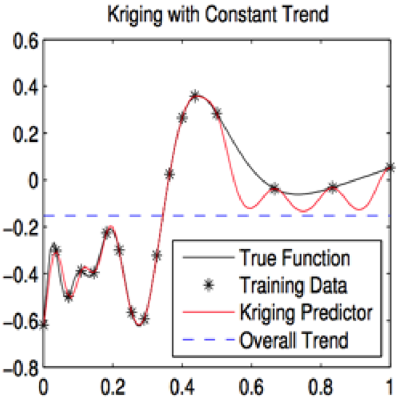
  $$Cov(Y(\boldsymbol{x}_2), Y(\boldsymbol{x}_2)) = \sigma^2 R(\boldsymbol{x}_1 - \boldsymbol{x}_2)$$

- Notice that if $\boldsymbol{Z}(\boldsymbol{x})$ was replaced by independent random errors, this would be the standard general linear model. However, we have allowed observations to be correlated.

Predictior

- Empirical best linear unbiased predictor (EBLUP)

$$\hat{y}(\boldsymbol{x}_0) = f^T(\boldsymbol{x}_0)\hat{\beta} + r_0^T \hat{\boldsymbol{R}}^{-1}(\boldsymbol{Y} - \boldsymbol{F}\hat{\beta}),$$



Prediction of Injection Molding

Here we randomly took 30 points as training data of the injection molding example and tried to predict the remaining 5 points using GP model.

Remaining Points	True Values	Pred. Values
r1	.0146	.0152
r2	.0149	.0157
r3	.0171	.0172
r4	.0149	.0149
r5	.0183	.0182
Training points	True Values	Pred. Values
t1	.0165	.0165
t2	.0169	.0169
t3	.0154	.0154
t4	.0173	.0173
t5	.0176	.0176

Outline

Sensitivity Analysis

- One property of computer experiment is a large number of input variables
- Sensitivity analysis tries to determine how variable the output is to change in the inputs.
- The most popular sensitivity analysis is based on an ANOVA-type Decompositions (Sobol´, 1990).

SA based on ANOVA Decomposition

- Overall mean: $y_0 = \int_{[0,1]^d} y(x_1, \ldots, x_d) dx_1 \cdots dx_d$

SA based on ANOVA Decomposition

- Overall mean: $y_0 = \int_{[0,1]^d} y(x_1, \ldots, x_d) dx_1 \cdots dx_d$
- Main effect: $y_i(x_i) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-i} - y_0$

SA based on ANOVA Decomposition

- Overall mean: $y_0 = \int_{[0,1]^d} y(x_1, \ldots, x_d) dx_1 \cdots dx_d$
- Main effect: $y_i(x_i) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-i} - y_0$
- Interaction effect:
  $y_{i,j}(x_i, x_j) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-ij} - y_0 - y_i(x_i) - y_j(x_j)$

SA based on ANOVA Decomposition

- Overall mean: $y_0 = \int_{[0,1]^d} y(x_1, \ldots, x_d) dx_1 \cdots dx_d$
- Main effect: $y_i(x_i) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-i} - y_0$
- Interaction effect:
  $y_{i,j}(x_i, x_j) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-ij} - y_0 - y_i(x_i) - y_j(x_j)$
- Total variance $V = \int_{[0,1]^d} y^2(x_1, \ldots, x_d) dx_1 \cdots dx_d - y_0^2$

SA based on ANOVA Decomposition

- Overall mean: $y_0 = \int_{[0,1]^d} y(x_1, \ldots, x_d) dx_1 \cdots dx_d$
- Main effect: $y_i(x_i) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-i} - y_0$
- Interaction effect:
  $y_{i,j}(x_i, x_j) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-ij} - y_0 - y_i(x_i) - y_j(x_j)$
- Total variance $V = \int_{[0,1]^d} y^2(x_1, \ldots, x_d) dx_1 \cdots dx_d - y_0^2$
- Partial variance $V_{i_1,\ldots,i_s} = \int_0^1 \cdots \int_0^1 y^2_{i_1,\ldots,i_s}(x_{i_1}, \ldots, x_{i_s}) dx_{i_1} \cdots dx_{i_s}$

SA based on ANOVA Decomposition

- Overall mean: $y_0 = \int_{[0,1]^d} y(x_1, \ldots, x_d) dx_1 \cdots dx_d$
- Main effect: $y_i(x_i) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-i} - y_0$
- Interaction effect:
  $y_{i,j}(x_i, x_j) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-ij} - y_0 - y_i(x_i) - y_j(x_j)$
- Total variance $V = \int_{[0,1]^d} y^2(x_1, \ldots, x_d) dx_1 \cdots dx_d - y_0^2$
- Partial variance $V_{i_1,\ldots,i_s} = \int_0^1 \cdots \int_0^1 y_{i_1,\ldots,i_s}^2 (x_{i_1}, \ldots, x_{i_s}) dx_{i_1} \cdots dx_{i_s}$
- First-order sensitivity index (Main effect index): $S_i = V_i / V$

SA based on ANOVA Decomposition

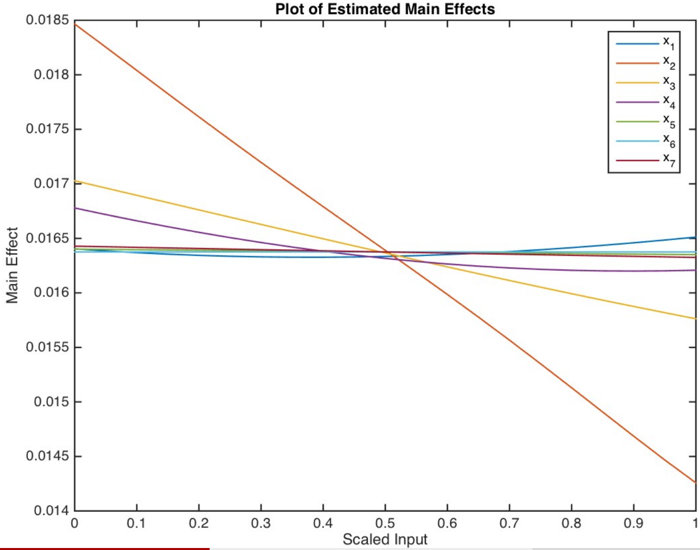
- Overall mean: $y_0 = \int_{[0,1]^d} y(x_1, \ldots, x_d) dx_1 \cdots dx_d$
- Main effect: $y_i(x_i) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-i} - y_0$
- Interaction effect:
  $y_{i,j}(x_i, x_j) = \int_0^1 \cdots \int_0^1 y(x_1, \ldots, x_d) dx_{-ij} - y_0 - y_i(x_i) - y_j(x_j)$
- Total variance $V = \int_{[0,1]^d} y^2(x_1, \ldots, x_d) dx_1 \cdots dx_d - y_0^2$
- Partial variance $V_{i_1,\ldots,i_s} = \int_0^1 \cdots \int_0^1 y^2_{i_1,\ldots,i_s}(x_{i_1}, \ldots, x_{i_s}) dx_{i_1} \cdots dx_{i_s}$
- First-order sensitivity index (Main effect index): $S_i = V_i / V$
- Total sensitivity index (Total effect index):
  $T_i = S_i + \sum_{j>i} S_{ij} + \sum_{j<i} S_{ji} + \cdots S_{1,2,\ldots,d}$

SA of Injection Molding

Factors	Main Effect Index	Total Effect Index
Melting time (x_1)	.0016	.0085
Packing time (x_2)	.8833	.8960
Packing pressure (x_3)	.0826	.0875
Cooling time (x_4)	.0187	.0208
HTC flow (x_5)	.0001	.0001
HTC pack (x_6)	.0001	.0001
HTC open (x_7)	.0001	.0001

Table 1: Sensitivity Indices of Injection Molding

Main Effect Plot of Injection Molding



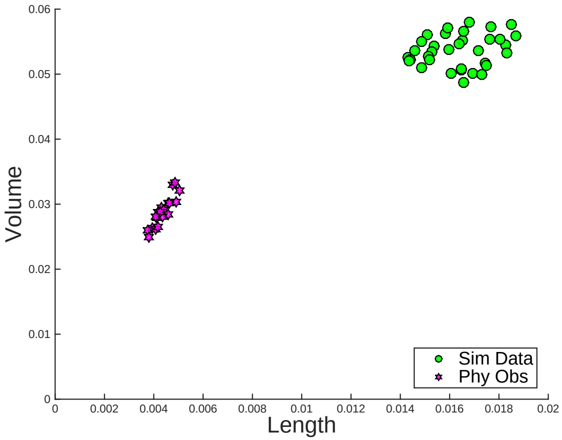
Outline

Calibration

- Does the computer model adequately represent reality?

Calibration

- Does the computer model adequately represent reality?
- 19 physical observations of the Injection Molding example



Calibration Parameters

- 4 factors in physical experiments:
  - Melting time, Packing time, Packing pressure, Cooling time
- 7 factors in computer experiments:
  - Melting time, Packing time, Packing pressure, Cooling time, HTC flow, HTC pack, and HTC open
- Calibration parameters:
  - Calibration parameters are those input variables which can be used to run computer simulators but are unknown in physical experiments.
  - HTC flow, HTC pack, and HTC open

Bayesian Calibration of Computer Experiments

- Physical system observations y_ℓ^p can be modeled as

$$Y_\ell^p(\boldsymbol{x}) = \mu_\ell(\boldsymbol{x}) + \epsilon_\ell(\boldsymbol{x}), \quad \ell = 1, \ldots, m, \tag{1}$$

where

1. $\mu_\ell(\cdot)$ is equal to the mean of a physical system.

Bayesian Calibration of Computer Experiments

- Physical system observations y_ℓ^p can be modeled as

$$Y_\ell^p(\boldsymbol{x}) = \mu_\ell(\boldsymbol{x}) + \epsilon_\ell(\boldsymbol{x}), \quad \ell = 1, \ldots, m, \tag{1}$$

where

1. $\mu_\ell(\cdot)$ is equal to the mean of a physical system.

- When $y_\ell^s(\boldsymbol{x}, \boldsymbol{t})$ is the data from a deterministic computer simulator, Kennedy and O'Hagan (2001) introduced

$$\delta_\ell(\boldsymbol{x}) \equiv \mu_\ell(\boldsymbol{x}) - y_\ell^s(\boldsymbol{x}, \phi). \tag{2}$$

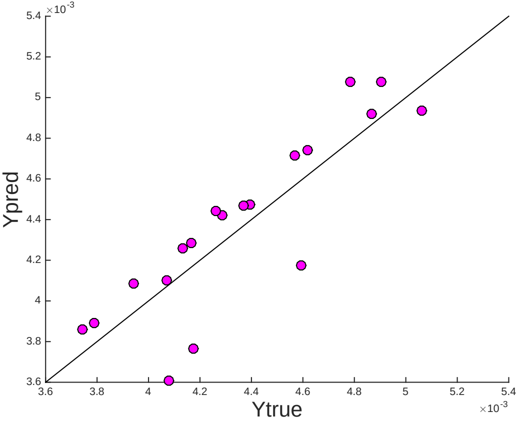
1. ϕ is the true value of calibration parameters,
2. We model $\delta_\ell(\cdot)$ and $y_\ell^s(\cdot)$ as realizations of Gaussian processes $U_\ell(\cdot)$ and $Y_\ell^s(\cdot)$, so

$$U_\ell(\boldsymbol{x}) = [Y_\ell^s(\boldsymbol{x}, \hat{\phi}) | y_\ell^p, y_\ell^s] + [\Delta_\ell(\boldsymbol{x}) | y_\ell^p, y_\ell^s], \tag{3}$$

3. $\delta_\ell(\cdot)$ is unidentifiable from frequentist viewpoint,
4. models $\delta_\ell(\cdot)$ based on Bayesian framework.

Calibrated Predictor

Using the calibrated predictor to predict 19 physical observations for output length



Thanks for your attention !

Outline

Organization

- Sep. 22nd: Allen, Introduction to Computer Experiments
- Oct. 6th
- Oct. 20th
- Nov. 3rd
- Nov. 11th
- Dec. 1st. Guest Speaker: Dr. Shan Ba, The Procter & Gamble Company.

References

Ba, Shan and Joseph, V. Roshan (2012), "Composite gaussian process models for emulating expensive functions", *Annals of Applied Statistics*, 6(4), 1838–1860.

Fricker, Thomas E., Oakley, Jeremy E., and Urban, Nathan M. (2013), "Multivariate gaussian process emulators with nonseparable covariance structures", *Technometrics*, 55(1), 47–56.

Gramacy, Robert B and Lee, Herbert K. H (2008), "Bayesian treed gaussian process models with an application to computer modeling", *Journal of the American Statistical Association*, 103(483), 1119–1130.

Higdon, D., Kennedy, M., Cavendish, J., Cafeo, J, and Ryne, R. (2004), "Combining field data and computer simulations for calibration and prediction", *SIAM Journal of Scientific Computing*, 26, 448–466.

Jones, Bradley and Johnson, Rachel T. (2009), "Design and Analysis for the Gaussian Process Model", *Quality and Reliability Engineering International*, 25, 515–524.

Kennedy, M. C. and O'Hagan, A. (2001), "Bayesian calibration of computer models (with discussion)", *Journal of the Royal Statistical Society B*, 63, 425–464.

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

Morris, M. D. and Mitchell, T. J. (1995), "Exploratory designs for computational experiments", *Journal of Statistical Planning and Inference*, 43, 381–402.

Ranjan, Pritam, Bingham, Derek, and Michailidis, George (2008), "Sequential experiment design for contour estimation from complex computer codes", *Technometrics*, 50(4), 527–541.

Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, New York: Springer Verlag.

Sobol´, I. M. (1990), "Sensitivity estimates for non-linear mathematical models", *Matematicheskoe Modelirovanie*, 2, 112–118.

Svenson, J. and Santner, T. (2011), "Multiobjective optimization of expensive black-box functions via expected maximin improvement", *submitted*.