A 2-Stage Design Procedure for Computer Experiments Comparative Study between Optimal Design and Space-Filling Design

N. Youssef
Supervisor: H. Wynn

Workpacage 3.1 Experimental Design

8 July 2009

Outline

Design of Computer Experiments

- Designs generated by elementary methods for selecting samples (Simple random sampling, stratified random sampling,etc...)
  - The design maybe evenly spread over the full experimental region but not onto all the factors
- Space filling designs,
  - latin hypercube design
  - designs based on measures of distance, e.g. maximin, minimax designs
  - Uniform designs or low discrepancy sequences borrowed from the Quasi-Monte Carlo Literature, e.g. Sobol's sequence, Halton sequence,...
  - Halton sequence construction is simple depending on the choice of a prime number
  - Sobol's sequence is well known space filling design, very useful for preforming sequential designs

Cont'd

- Model-Based optimal design
  - Selecting the design is based on a statistical criterion
  - The design is selected according to the model chosen for $Y(x)$
  - $D-$optimality is one example
  - Integrated mean squared error and Maximum Entropy Sampling criteria are common in computer experiments

Maximum Entropy Sampling Design (Shewry and Wynn (1987))

- Entropy is negative of information

$$\text{Ent}(Y_s) = \text{E}_{Y_s}[-\log p(Y_s)]$$

  - Y_s vector of observations
  - $p(.)$ density function of Y
  - $D_s = \{x_1, \ldots, x_n\}$
- Partition $Y_N = (Y_S, Y_{N \setminus S})$
- $\text{Ent}(Y_N) = \text{Ent}(Y_S) + \text{E}_{Y_S}\text{Ent}(Y_{N \setminus S}|Y_S)$
- $\text{Ent}(Y_N)$ is fixed
- Min $\text{Ent}(Y_{N \setminus S}|Y_S)$ is $\equiv$ Max $\text{Ent}(Y_S)$
- In the Gaussian case,

$$\text{Ent}(Y_S) = \frac{n}{2}[1 + \log 2\pi] + \frac{\log|\Sigma_s|}{2}$$

- The problem is finding the design that maximizes $\log|\Sigma_s|$

The Proposed Method for 2-stage Design

1. Initial data set (outputs and inputs)
2. MLE's are obtained for all model parameters
3. Given the initial data set another design is chosen
4. The output corresponding to the design points is obtained from the simulator
5. The new design points with their outputs are added to the initial data set
6. Another training data set is obtained used for validating our results
7. A comparison is made based on the accuracy of prediction

Kriging for Computer Experiments

- $$Y(x) = f^T(x)\beta + Z(x)$$
- $Y(x)$ is a Gaussian process with mean

  $$E(Y(x)) = X^T\beta$$

  where X is the design matrix and var-cov matrix

  $$\text{cov}(Y(x), Y(x)') = \sigma^2 R(x - x')$$

  R is the correlation matrix of process $Z(x)$, σ^2 constant.
- $R(x - x')$ has several forms, e.g. the exponential structure

  $$corr(Z(x), Z(x')) = \prod_{i=1}^{d} \exp(-\theta(x - x')^p)$$

Prediction with Kriging

- MLE's for β and $\sigma^2$are given by

$$\hat{\beta} = (X^T R^{-1} X)^{-1} X^T R^{-1} y_D$$
$$\hat{\sigma}^2 = \frac{1}{N}(y_D - X\hat{\beta})^T R^{-1}(y_D - X\hat{\beta})$$

- MLE's for covariance parameters are obtained using a software created by Ron Bates(LSE)
- The **BLUP** of the response $Y(x_*)$ is

$$\hat{y}(x_*) = f^T(x_*)\hat{\beta} + r^T R^{-1}(y_D - x\hat{\beta})$$

where

$$r = [R(x_1, x_*), \cdots, R(x_n, x_*)]$$

Measuring Accuracy of Prediction

- Cross validation EMPIRICAL ROOT MEAN SQUARED ERROR , (Welch et al, (1992))

$$\text{ERMSE}_{-1} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_{D_{-1}}(x_i) - y_D(x_i))^2}$$

  - $\hat{y}_{-1}$ data set without observation i, no validation data
  - Parameters are re-estimated except the covariance parameters

- EMPIRICAL ROOT MEAN SQUARED,

$$\text{ERMSE} = \sqrt{\frac{1}{N_v}\sum_{i=1}^{N_v}(\hat{y}_v(x_i) - y_v(x_i))^2}$$

- ROOT MEAN SQUARED ERROR

$$\text{RMSE} = \sqrt{\frac{1}{N_v}\sum_{i=1}^{N_v} MSE(\hat{y}_v(x_i))}$$

Other Measures of Prediction Accuracy

- Mahalanobis Distance,

$$MD(y_*) = (y(x_*) - \hat{y}(x_*))^T Var(\hat{y}(x_*))^{-1}(y(x_*) - \hat{y}(x_*))$$

borrowed from the literature of the regression diagnostics to detect leverages

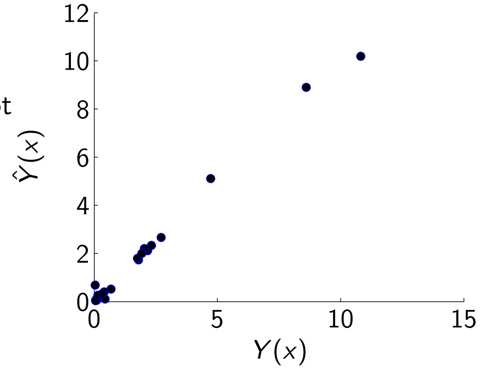
- Extremely large or small values indicates problems with the predictor (Emulator)

Saltelli Simulator (2000)

$$Y(x) = \frac{(x_2 + \frac{1}{2})^4}{(x_1 + \frac{1}{2})^2}$$

- Initial design of 20 points by simple random sampling
- $Y(x) = \mu + Z(x)$
- More complicated model did not show any improvement in prediction (Welch et al, 1989)
- MLEs for model parameters are obtained including the covariance parameters
- Validating the model using the initial design

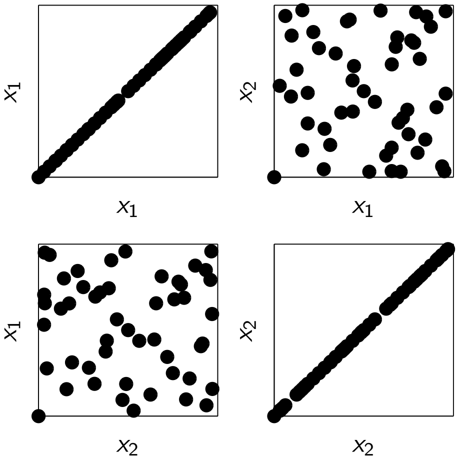
Figure: Cross-Validation:Predicted vs. Observed



Space filling design results for Saltelli's simulator(2000)

- Sobol's sequence is chosen as design
- We observe the $Y(x)$ at those set of points
- The results were augmented to the initial design
- No loss for the runs

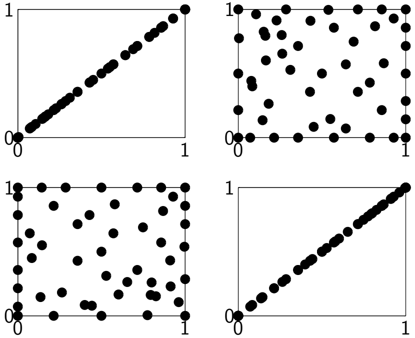
Figure: Sobol's sequence of 30 points



Entropy design for Saltelli simulator (2000)

- A design of size 30 obtained in addition to the initial design
- Based on the MLEs obtained from the initial design
- Exchange algorithm has been used

Figure: Entropy design with Initial design



Saltelli Simulator Results

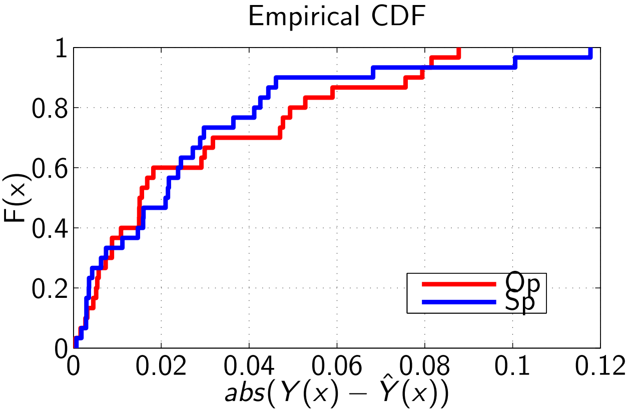


Figure: CDF for the absolute prediction error for Space filling and Optimal design

Saltelli Simulator Results

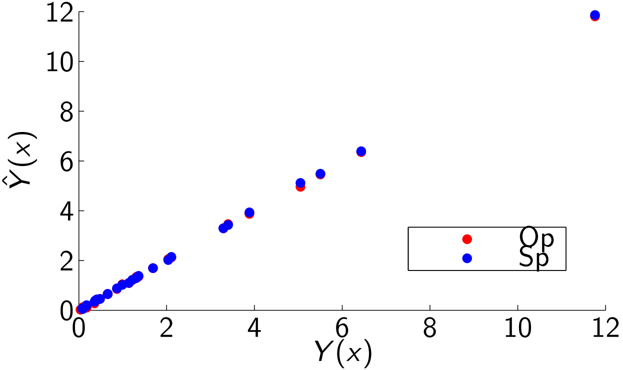


Figure: Predicted vs.Observed output for Space filling and Optimal design

Numerical Results for Saltelli Simulator

Parameter	Ini. des. est.	Sp. fill. est.	Opt. des. est.
$\hat{\mu}$	5.1276	11.2654	19.6862
$\hat{\sigma}^2$	28.1399	80.9928	357.4433
$\hat{\theta}_1$	2.7072	2.9366	2.3792
$\hat{\theta}_2$	0.88804	1.172	0.8761
$\hat{p}_1$	1.999	1.999	1.999
$\hat{p}_2$	1.999	1.999	1.999

Table: MLEs for the model parameters

- The estimates of the mean parameter are different
- Very close covariance estimates

Accuracy Measure Results for Saltelli Simulator

Measure	Ini. des. est.	Sp. fill. est.	Opt. des. est.
$ERMSE_{-1}$	0.2635	0.024743	0.2232
ERMSE	0.29005	0.038109	0.038426
RMSE	0.14023	0.047151	0.05117
Mah Dist	124.7476	32.8165	14.5268
Max abs. dev.	1.0727	0.11771	0.087713

Table: Accuracy measures of prediction

- Improvement in prediction than the initial design
- The optimal design is very similar to space filling one in terms of ERMSE

Application:Rolls Royce four Blade Fan Assembly Simulator

- The outputs are maximum amplitude for each blade $a_1, \cdots, a_4$ and frequency $f_1, \cdots, f_4$ at which that maximum resonant frequency occurs
- Because of symmetry, the analysis is for one output, the amplitude of the first blade
- The design variables $x_1, \cdots, x_4$ are the small amounts of 'mass' added to the blades.
- The amount of mass can take positive or negative values ($\pm 3 \times 10^{-6}$).
- To simulate we run ANSYS software, to preform modal and harmonic analysis
- The design points are transformed to be in $[0, 1]$ in the selecting process

Rolls Royce Results

Initial design is obtained by a combination of two fractional factorial design of ± 1 and ± 0.5.

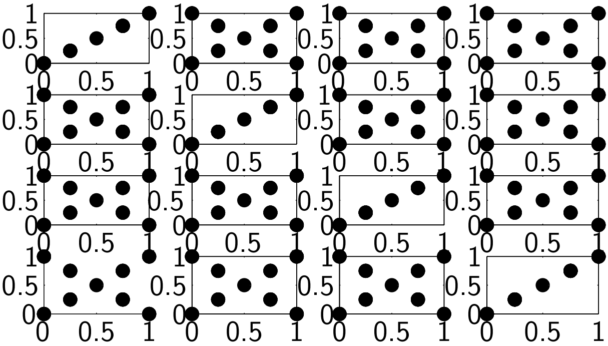


Figure: Initial Design of 33 Points

Rolls Royce Initial Results

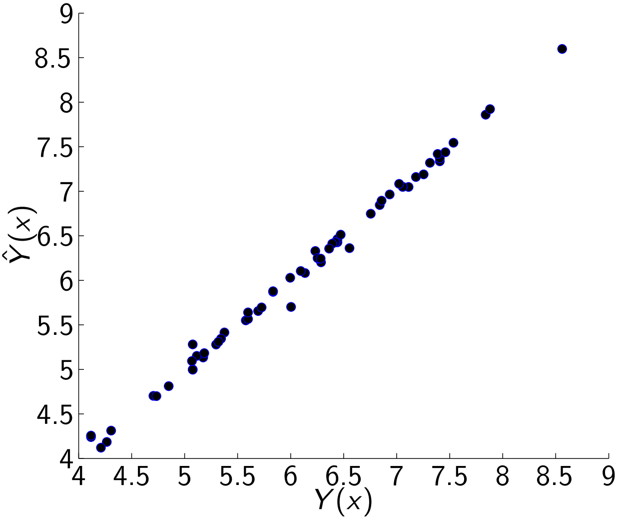
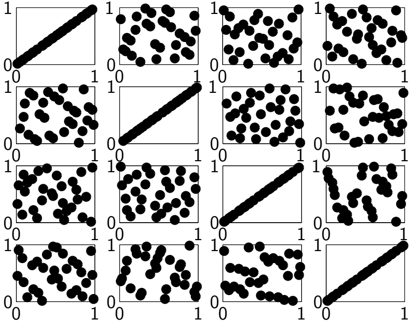


Figure: Cross Validation predction for intial design

Space Filling Design for Rolls Royce Simulator

- Sobol's sequence of 30 points is obtained
- The 30 points are added to the initial 33 points
- The MLE's are obtained for all model parameters including the covariance parameters

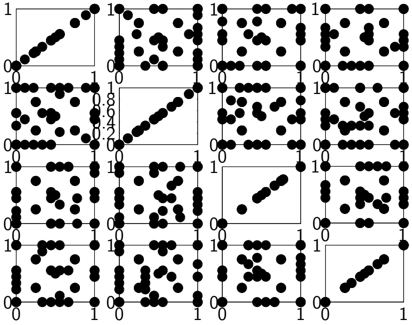
Figure: Sobol's sequence of 30 points



Maximum Entropy Sampling Design for Rolls Royce Simulator

- Maximum Entropy Sampling is obtained using the initial design points
- The design points are chosen over a grid, (10×10).
- Exchange algorithm is used to find the optimal design put of 10^4 points
- The MLE's are obtained for all model parameters

Figure: Entropy Design Including the Initial Design (63 points)



Rolls Royce Simulator Graphical Results

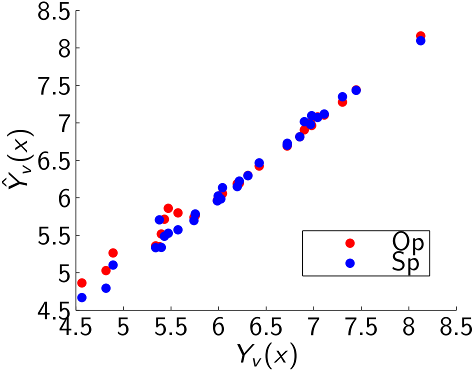


Figure: Predction for Space filling and Optimal design

Rolls Royce Results

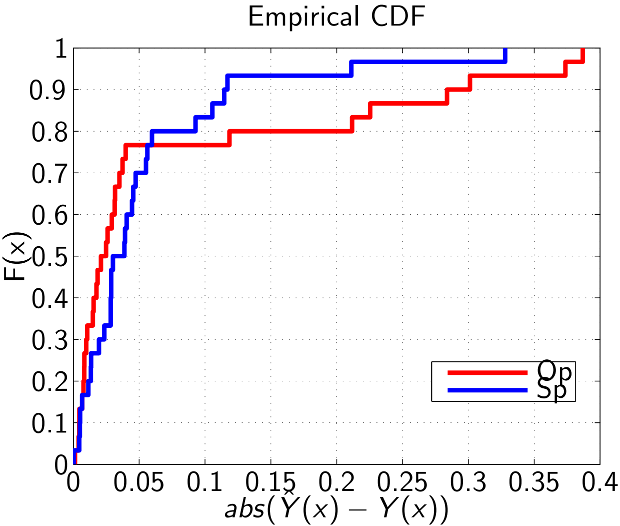


Figure: CDF for the absolute prediction error for Space filling and Optimal design

Numerical Results for Rolls Royce Simulator

Parameter	Ini. des. est.	Sp. fill. est.	Opt. des. est.
$\hat{\mu}$	5.5223	5.0722	4.8029
$\hat{\sigma}^2$	1.472	1.6301	1.6003
$\hat{\theta}_1$	0.52291	0.5004	0.45445
$\hat{\theta}_2$	0.17127	0.29412	0.26167
$\hat{\theta}_3$	0.4247	0.31322	0.2212
$\hat{\theta}_4$	0.35561	0.21175	0.21745
$\hat{p}_1$	1.9621	1.9796	1.4326
$\hat{p}_2$	1.7793	1.972	1.999
$\hat{p}_3$	1.999	1.991	1.999
$\hat{p}_4$	1.999	1.6136	1.999

Table: MLEs for the model parameters

- The estimates of the parameters are quite similar
- Some differences in the values of covariance parameters estimates

Accuracy Measure Results for Rolls Royce Simulator

Measure	Ini. des. est.	Sp. fill. est.	Opt. des. est.
$ERMSE_{-1}$	0.093422	0.086923	0.069252
ERMSE	0.066662	0.086636	0.13913
RMSE	0.10691	0.051009	0.088515
Mah Dist	4.4448	181.682	45.1828
Median abs.dev.	0.0395	0.0344	0.0229
Max abs. dev.	0.17144	0.32793	0.38686

Table: Accuracy measures of prediction

- Improvement in prediction
- The optimal design has a greater percentage of small absolute deviations than space filling
- Mahalanobis distance for optimal design is smaller than space filling design (cross correlation)

Some Interesting Results

Measure	Sp. fill. est.	Opt. des. est.
Entropy	-42.5889	-37.3700
RMSE1	0.0612	0.0707
RMSE2	0.0614	0.0529

Table: Entropy values and Sum of Mean Squared Error at the un-sampled points

- Using the same estimates for the covariance parameters, the RMSE1 favours the optimal design
- It is trivial because the Entropy design aims at minimizing the predictive variance
- Using the updated estimates of the covariance parameters, RMSE2 is in Space filling design side

Conclusions

- We cannot say which design is better
- They are very similar in the prediction accuracy results
- Space-filling shows the ability to learn about the covariance parameters
- We have to decide what we are aiming at from our experiment, then we decide which method we can choose

Main References

- Bastos and O'Hagan, Technometrics,(2009).
- Sacks J. et al, Technometrics,(1989).
- Saltelli A. et al, Statistical science, (2000).
- Shewry and Wynn , J. of Apply Statistics,(1977).