Designs for Response Surface Models with Random Block Effects

by
Dr. Sourish Saha
Research Statistics Unit, GlaxoSmithKline

Acknowledgement: **Prof. Andre I. Khuri**
Department of Statistics, University of Florida

2008 Quality and Productivity Research Conference
Madison, Wisconsin

Overview

- Model and Notation
- Motivation and Literature Review
- Design Dependence Problem
- Design criteria
- Quantile Dispersion Graphs (QDGs)
- Sequential Generation of D-optimal designs

Model and Notation

- Consider the following model of order d (≥ 1) in k input (control) variables $x_1, x_2, ..., x_k$

$$\boldsymbol{y} = \beta_0 \mathbf{1}_n + \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\epsilon}, \tag{1}$$

  where

  $\boldsymbol{y}$ is a vector of responses,
  β_0 and $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$ are unknown parameters,
  $\boldsymbol{\gamma} \sim N(\mathbf{0}, \sigma_\gamma^2 \boldsymbol{I}_b)$
  $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma_\epsilon^2 \boldsymbol{I}_n)$

- $\boldsymbol{X}_{n\times p}$ and $\boldsymbol{Z}_{n\times b}$ are known matrices with ranks p and b, respectively.
- $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{f}'(\boldsymbol{x}_1) \\ \boldsymbol{f}'(\boldsymbol{x}_2) \\ \vdots \\ \boldsymbol{f}'(\boldsymbol{x}_n) \end{pmatrix}$, $\boldsymbol{f}'(\boldsymbol{x}_u)(u = 1, 2, \ldots, n)$ are vectors whose elements are powers and crossproducts of powers of $x_1, x_2, \ldots, x_k$ of degree d evaluated at $\boldsymbol{x}_u$.
- $\boldsymbol{Z} = diag(1_{n_1}, 1_{n_2}, ..., 1_{n_b})$

- **Analysis of Response Surface Models (RSMs)** - Khuri (1992, 1996, 2006)

- **Designs for RSMs with *fixed* block effects**
  - Atkinson and Donev (1989)
  - Trinca and Gilmour (2000)
  - Cook and Nachtsheim (1989)

- **Designs for RSMs with *random* block effects**
  - Cheng (1995)
  - Atkins and Cheng (1999)
  - Goos and Vandebroek (2001)

- **Exchange-type algorithms**
  - Fedorov (1972)
  - DETMAX algorithm (Mitchell, 1974)
  - Modified Fedorov algorithm (Cook and Nachtsheim, 1989)
  - KL-exchange and BLKL algorithm (Atkinson and Donev, 1989)
  - Coordinate exchange algorithm (Meyer and Nachtsheim, 1995)
  - First-Order algorithm (Fedorov, Gagnon, Leonov and Wu, 2007)
- An excellent treatment of the construction of D-optimal designs is given in Atkinson and Donev (1992), Goos (2002) and Atkinson, Donev and Tobias (2007) .

Design Dependence Problem

- **What is the "Design Dependence Problem" ?**
  - The prediction variance ***depends*** on the unknown parameters.
    - $\Rightarrow$ Assessment of the design effect on the prediction variance depends on the value of those unknown parameters.

- Methods to deal with this *dependence* problem
  1. locally optimal designs
  2. **sequential method**
  3. Bayesian approach
  4. **quantile dispersion graphs**

Design Criteria

Model (1) can be written as

$$\boldsymbol{y} = \boldsymbol{W\tau} + \boldsymbol{Z}\gamma + \boldsymbol{\epsilon}, \tag{2}$$

where $\boldsymbol{\tau} = (\beta_0, \boldsymbol{\beta}')'$ and $\boldsymbol{W} = [\mathbf{1}_n : \boldsymbol{X}]$. Hence,

$$\begin{aligned} E(\boldsymbol{y}) &= \boldsymbol{\mu} \\ &= \beta_0 \mathbf{1}_n + \boldsymbol{X\beta} \\ &= \boldsymbol{W\tau} \end{aligned}$$

$$\begin{aligned} V(\boldsymbol{y}) &= \boldsymbol{\Sigma} \\ &= \sigma_\epsilon^2 \boldsymbol{I}_n + \sigma_\gamma^2 \boldsymbol{ZZ}' \\ &= \sigma_\epsilon^2 \boldsymbol{A} \end{aligned}$$

where

$$\boldsymbol{A} = diag(\boldsymbol{A}_1, \boldsymbol{A}_2, ..., \boldsymbol{A}_b), \tag{3}$$

$\boldsymbol{A}_l = \boldsymbol{I}_{n_l} + \eta \boldsymbol{J}_{n_l}$ $(l = 1, 2, ..., b)$, and $\eta = \sigma_\gamma^2 / \sigma_\epsilon^2$.

Design Criteria contd...

- If $\eta = \sigma_\gamma^2/\sigma_\epsilon^2$ is known, then

$$\hat{\boldsymbol{\tau}} = (\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{y}. \tag{4}$$

  and the variance-covariance matrix is

$$Var(\hat{\boldsymbol{\tau}}) = (\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{W})^{-1}\sigma_\epsilon^2. \tag{5}$$

Let $\hat{y}(\boldsymbol{x}) = \boldsymbol{g}'(\boldsymbol{x})\hat{\boldsymbol{\tau}}$ be the ***predicted response*** at a point $\boldsymbol{x}$ in the experimental region R, where $\boldsymbol{g}'(\boldsymbol{x}) = [1, \boldsymbol{f}'(\boldsymbol{x})]$.

Design Criteria contd...

- $\hat{y}(\boldsymbol{x}) = \boldsymbol{g}'(\boldsymbol{x})\hat{\boldsymbol{\tau}}$
- The prediction variance for $\hat{y}(\boldsymbol{x})$, $Var[\hat{y}(\boldsymbol{x})]$, is of the form

$$Var[\hat{y}(\boldsymbol{x})] = \boldsymbol{g}'(\boldsymbol{x})(\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{W})^{-1}\boldsymbol{g}(\boldsymbol{x})\sigma_{\epsilon}^{2}. \tag{6}$$

- The ***scaled prediction variance*** is defined as

$$\frac{n}{\sigma_{\epsilon}^{2}} Var[\hat{y}(\boldsymbol{x})] = n[\boldsymbol{g}'(\boldsymbol{x})(\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{W})^{-1}\boldsymbol{g}(\boldsymbol{x})] \tag{7}$$

- **A criterion for the choice of a design is the minimization of** $\frac{n}{\sigma_{\epsilon}^{2}} Var[\hat{y}(\boldsymbol{x})]$.

Another Design Criteria - Power

(Saha and Khuri, 2008, Journal of Applied Statistics, under review)

- Let us rewrite model (2) as

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{\tau} + \boldsymbol{\epsilon}^* \tag{8}$$

  where $\boldsymbol{\tau} = (\beta_0, \boldsymbol{\beta}')'$, $\boldsymbol{W} = [\mathbf{1}_n : \boldsymbol{X}]$ and $\boldsymbol{\epsilon}^* = \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$.
- Consider the hypothesis test
  $H_0 : \boldsymbol{L}\boldsymbol{\tau} - \boldsymbol{d} = \mathbf{0}$ vs $H_a : \boldsymbol{L}\boldsymbol{\tau} - \boldsymbol{d} = \boldsymbol{\delta}$
  where
  - $\boldsymbol{L}_{q\times(p+1)}$ is a known matrix of rank $q(< p+1)$,
  - $\boldsymbol{d}$ is a known $q \times 1$ vector, and
  - $\boldsymbol{\delta}$ is an alternative value of $\boldsymbol{L}\boldsymbol{\tau} - \boldsymbol{d}$ ($\neq \mathbf{0}$).

Another Design Criteria contd...

- For known η, the generalized least squares (GLS) F test is based on

$$F_{GLS}(\eta) = \frac{(\boldsymbol{L}\hat{\boldsymbol{\tau}} - \boldsymbol{d})'[\boldsymbol{L}(\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{W})^{-1}\boldsymbol{L}']^{-1}(\boldsymbol{L}\hat{\boldsymbol{\tau}} - \boldsymbol{d})}{qMS_E}$$

  where MS_E denotes the error mean square for model (8), and is given by

$$MS_E = \frac{\boldsymbol{y}'[\boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{A}^{-1}]\boldsymbol{y}}{(n-p-1)}$$

- The test statistic $F_{GLS}(\eta)$ has an exact F distribution under H_0 with q and $n-p-1$ degrees of freedom

Another Design Criteria contd...

- Under H_a, $F_{GLS}(\eta)$ is distributed as a noncentral F with q and $n-p-1$ degrees of freedom and a noncentrality parameter, ξ, given by (see for example, Searle (1971, Section 3.6))

$$\xi = \boldsymbol{\delta}'[\boldsymbol{L}(\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{W})^{-1}\boldsymbol{L}']^{-1}\boldsymbol{\delta}/\sigma_\epsilon^2$$

- The power of $F_{GLS}(\eta)$ is a monotonically increasing function of the noncentrality parameter, ξ. Let us denote the power by ψ, which is given by

$$\psi = P[F_{GLS}(\eta) \geq F_{\alpha;q,n-p-1} | \xi \neq 0]$$

- Hence, **the noncentrality parameter and the power function of the F-test can be used as design criteria for comparing designs**.

Quantile Dispersion Graphs (QDGs)

(Khuri, 1997, Journal of Applied Statistics)

- Study the distribution of a design criterion throughout the experimental region, R.
- Distribution is determined in terms of its ***quantiles***.
- Displays the ***dependence*** of the design on the ***model's parameters***.
- Key Advantage:
  Performance of a design evaluated ***throughout the region*** R, not based on a single measure, such as D-efficiency.

Construction of QDGs

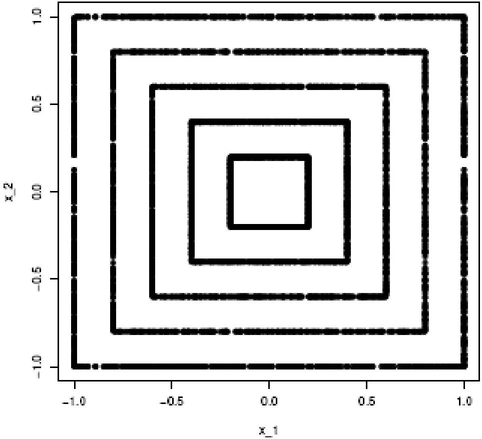
Using the ***scaled prediction variance*** as a design criterion:

$$\frac{n}{\sigma_\epsilon^2} Var[\hat{y}(\boldsymbol{x})] = n[\boldsymbol{g}'(\boldsymbol{x})(\boldsymbol{W}'\boldsymbol{A}^{-1}\boldsymbol{W})^{-1}\boldsymbol{g}(\boldsymbol{x})] \tag{9}$$

- Denote the scaled prediction variance of design D, at $\boldsymbol{x}$ in R, by $\Delta_D(\boldsymbol{x}, \eta)$.
- Distribution of $\Delta_D(\boldsymbol{x}, \eta)$ within R is determined in terms of its quantiles.
- R_λ represents the surface of the region obtained by shrinking R by a factor λ.
- Parameter space, C is specified.
- For a given D and $\eta \in C$, $Q_D(p, \eta, \lambda)$ denotes the pth quantile on concentric surfaces R_λ.

Construction of QDGs contd...

- Compute $Q_D(p, \eta, \lambda)$ for several values of η.
- $Q_D^{min}(p, \lambda) = \min_{\eta \in C} Q_D(p, \eta, \lambda)$
  $Q_D^{max}(p, \lambda) = \max_{\eta \in C} Q_D(p, \eta, \lambda)$
- Plotting these values against p results in the QDGs of the scaled prediction variance over R_λ.
- Understanding the QDGs:
  Small values of Q_D^{max} are desirable $\Rightarrow$ ***small prediction variance*** within the experimental region, R.
  Closeness of the Q_D^{max} and Q_D^{min} $\Rightarrow$ ***robustness*** of the design to changes in η.


1.0
0.5
0.0
-0.5
-1.0
x_2
-1.0
-0.5
0.0
0.5
1.0
x_1

Parameter Space

(Saha and Khuri, 2008, Journal of Quality Technology Quantitative Management, accepted)

- $H_0 : \eta = 0$. If H_0 is true, then $S_{\boldsymbol{\gamma}}/(\sigma^2_\epsilon + m\sigma^2_\gamma) \sim \chi^2_r$.
- However, if H_0 is false, then the distribution of $S_{\boldsymbol{\gamma}}/(\sigma^2_\epsilon + m\sigma^2_\gamma)$ is ***not*** χ^2_r, but is a linear combination of two or more independently distributed chi-squared random variables.
- $G(\eta; \boldsymbol{y}) \sim F(r, f)$ (Harville and Fenech, 1985).
- The exact two-sided $(1-\alpha)100\%$ confidence interval on η is given by $[l^*, u^*]$, where l^*, u^* are, respectively, the roots of the following equations (Burdick and Graybill, 1992)

$$
\begin{aligned}
G(\eta; \boldsymbol{y}) &= F_{1-\alpha/2; r, f} \\
G(\eta; \boldsymbol{y}) &= F_{\alpha/2; r, f}
\end{aligned}
$$

Example: Using the scaled prediction variance as design criterion

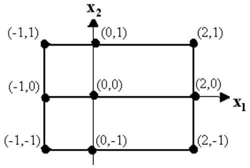
- Experiment: effects of temperature and curing time on the shear strength of the bonding of galvanized steel bars with a certain adhesive.
- Batches of steel were selected at random over a period of 12 days (blocks). The same 3^2 design was run on each of the 12 dates. The variance components are $\sigma_\gamma^2, \sigma_\epsilon^2$.
- Design: 3×3 factorial. The mean response is

$$\begin{aligned}\eta(x) &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 \\ &\quad + \beta_{12} x_1 x_2\end{aligned}$$

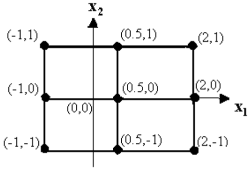
$$\begin{aligned}x_1 &= \frac{\text{temperature} - 400}{25} \\ x_2 &= \frac{\text{time} - 35}{5}\end{aligned}$$

- The experimental region, R, is rectangular with $-1 \leq x_1 \leq 2$, $-1 \leq x_2 \leq 1$.

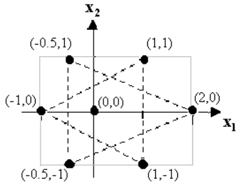
- D_1 is the combination of all 3^2 factorial designs (in those 12 blocks). We compare D_1 with four other second-order designs.
- For each design, we consider the distribution of $\Delta_D(\boldsymbol{x}, \eta)$ on each of several concentric rectangles, R_λ $(0.5 < \lambda \leq 1)$.
- In order to investigate the dependence of $\Delta_D(\boldsymbol{x}, \eta)$ on η, we obtained the 95% confidence interval, C, namely $(l^*, u^*) = (0.076, 0.967)$.
- For each design and a selected value of η in C, quantiles of the distribution of $\Delta_D(\boldsymbol{x}, \eta)$ are obtained for $\boldsymbol{x} \in R_\lambda$.
- The number of points chosen on each R_λ was 2000 consisting of 500 on each side.
- The quantiles are calculated for $p = 0(0.05)1$. The procedure is repeated for other values of η in C. Then $Q_D^{max}(p, \lambda)$ and $Q_D^{min}(p, \lambda)$ are calculated.


x2
x1
(-1,1)
(0,1)
(2,1)
(-1,0)
(0,0)
(2,0)
(-1,-1)
(0,-1)
(2,-1)

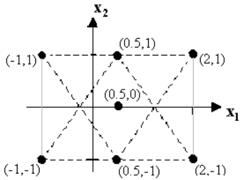
D1


x2
x1
(-1,1)
(0.5,1)
(2,1)
(-1,0)
(0.5,0)
(2,0)
(0,0)
(-1,-1)
(0.5,-1)
(2,-1)

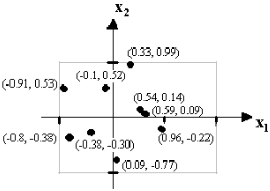
D2


x2
x1
(-0.5,1)
(1,1)
(-1,0)
(0,0)
(2,0)
(-0.5,-1)
(1,-1)

D3


x2
x1
(0.5,1)
(-1,1)
(2,1)
(0.5,0)
(-1,-1)
(0.5,-1)
(2,-1)

D4


x2
x1
(0.33, 0.99)
(-0.1, 0.52)
(-0.91, 0.53)
(0.54, 0.14)
(0.59, 0.09)
(-0.8, -0.38)
(-0.38, -0.30)
(0.96, -0.22)
(0.09, -0.77)

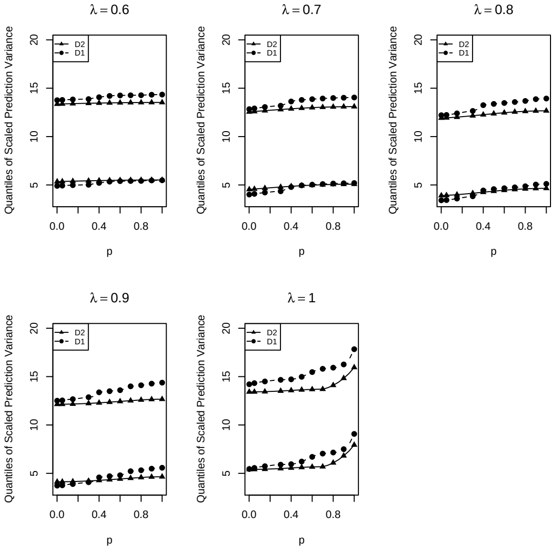


Figure: Quantile dispersion graphs for D_1 and D_2 [Example]

Conclusion

- It is evident that for small values of λ (less than 0.9), D_2 performs slightly better than D_1. For $\lambda = 0.9$ and 1, the difference in prediction capability is distinctly noticed in favor of D_2.
- The semi uniform shell designs - D_3 and D_4 perform better than D_1 for $\lambda = 0.6, 0.7$ and 0.8. However, for $\lambda = 0.9$ and 1 and high values of p, D_1 has better prediction capability than D_3 and D_4.
- The QDGs for the comparison of design D_1 with D_5 clearly depict better prediction capability with design D_1. Also, design D_4 is better than D_3 and D_5.

Sequential/Iterative/Adaptive D-optimal designs

- Normality of random effect and errors, and linearity of the fitted model $\rightarrow$ a confidence region for $\boldsymbol{\tau}$.
- A precise estimate of $\boldsymbol{\tau}$ can be obtained by making the volume of this region small. Since volume is proportional to $|\boldsymbol{M}(\zeta, \boldsymbol{V})|^{-1/2}$, one can achieve this objective by choosing a design measure which maximizes $|\boldsymbol{M}(\zeta, \boldsymbol{V})|$ (D-optimality).
- Besides improving the precision of $\hat{\boldsymbol{\tau}}$, D-optimal design measures minimize the maximum of the variance of the predicted response over the region of interest χ (G-optimality).
- Kiefer and Wolfowitz (1960): Single response
- Federov (1969, 1972): Multi-response (known variance-covariance matrix)
- Wijesinha and Khuri (1987) proposed a sequential algorithm to generate a *D*-optimal multiresponse design (unknown variance covariance matrix - Zellner (1962)).
- Atkins and Cheng (1999)
- Fedorov, Gagnon and Leonov (2002), Fedorov and Leonov (2004) and Dragalin, Fedorov and Wu (2006)