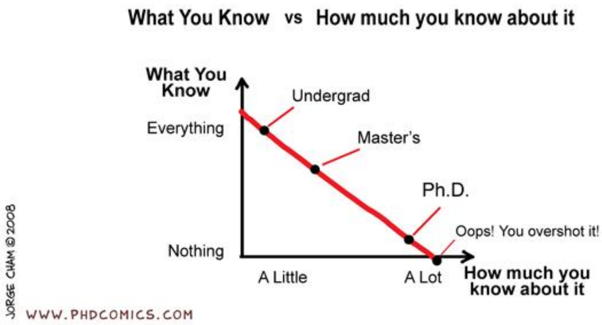
Multiresponse Exchange Algorithms for Model-Robust Experimental Design

Byran Smucker, Enrique del Castillo, and James Rosenberger

The Pennsylvania State University

October 8, 2009

PhD Comics


What You Know vs How much you know about it
What You Know
Everything
Nothing
Undergrad
Master's
Ph.D.
Oops! You overshot it!
A Little
A Lot
How much you know about it
JORGE CHAM © 2008
WWW.PHDCOMICS.COM

PhD Comics


DID YOU TAKE A LOOK AT THOSE SLIDES I SENT?
UH-HUH, SURE.
JORGE CHAM © 2007
SIGN #137 OF A DYSFUNCTIONAL STUDENT-FACULTY RELATIONSHIP:
HAVE YOU LOOKED AT THE SECOND ORDER CORRELATIONS?
ALSO, YOUR HYPOTHESIS IS A COMPLETE JOKE.
YOUR ADVISOR ASKS MORE QUESTIONS THAN ANYBODY ELSE AT YOUR OWN SEMINAR.
WWW.PHDCOMICS.COM

Outline

Introduction

Plastic formulation (Snee 1985)

- Suppose you must study the effect of five mixture factors on the hardness of a plastic product using 25 runs
- Mixture constraint, other single-factor and multi-factor constraints
- The form of the regression model is unknown; main effects? quadratic? cubic terms?

Figure: Manufactured Plastics[a]

[a] www.plastipak.com.pe/empresa_ingles.htm

Model-robustness problem

- Find design that is "good" for all models of interest
- We accomplish this by generalizing extant exchange algorithms and appealing to multiresponse optimal design theory

$\mathcal{D}$-optimal design:

$$\textbf{Assume } f \rightarrow \textbf{Expand to } \mathbf{X} \rightarrow \textbf{Choose } \xi_n \textbf{ to maximize } |\mathbf{X}'\mathbf{X}|$$

Model-robust design:

$$\textbf{Assume } \mathcal{F} \rightarrow \textbf{Expand each to } \mathbf{X}_i \rightarrow \textbf{Choose } \xi_n \textbf{ to maximize } \prod_{\mathcal{F}} |\mathbf{X}_i'\mathbf{X}_i|$$

Setting

Consider the univariate regression model of response y on factors $\mathbf{x}$:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \epsilon \tag{1}$$

- $\mathbf{X}$ is the expanded design matrix
- OLS estimators are $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$
- Assuming $E(\epsilon_i) = 0$ and $Var(\epsilon_i) = \sigma^2$, $Var(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$
- Information matrix: $\frac{\mathbf{X}'\mathbf{X}}{\sigma^2} \propto \mathbf{X}'\mathbf{X} = \mathbf{M}$

Experimental Design

There are many standard experimental designs, e.g. central composite

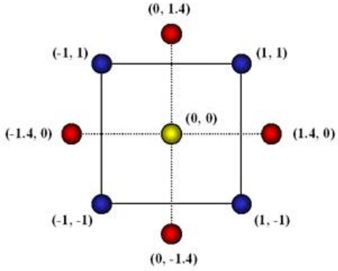


Figure: Central composite design for two factors[1]

[1] http://www.smallprecisiontools.com/Image/products/bonding/capillaries/Capillary_DOE_Matrix_1_320x240_col.jpg

Inadequacy of Standard Designs

Standard designs are usually sufficient, but not always

- Design space constrained (e.g. mixture experiments)
- Categorical factors
- Nonstandard sample sizes required

In such cases, it is natural to choose a design based on a criterion of "design goodness."

Optimal Design

Optimize some function of the information matrix or prediction variance

- Many criteria exist: $\mathcal{A}$, $\mathcal{D}$, $\mathcal{E}$, $\mathcal{G}$, $\mathcal{IV}$; which is best?
- $\mathcal{D}$-optimality: $\phi(\mathbf{M}) = |\mathbf{M}|$
- $\mathcal{D}$-optimal designs minimize the volume of confidence ellipsoid of parameters
- Computationally simple to work with

Criticism: Requires complete specification of the model-form, which is usually unknown at design stage. Ramifications?

Model-Robust Designs

Goal: choose a design which is robust to possible misspecification of model

- Large literature, but most use continuous design theory and/or have other unrealistic assumptions
- Our approach: Allow experimenters to specify a class of models, $\mathcal{F}$, and use exchange algorithms to find a design which optimizes some function of $\mathbf{M}_{\mathcal{F}} = (\mathbf{M}_1, \ldots, \mathbf{M}_r)$
- Few exact methods (Welch 1983; DuMouchel and Jones 1994; Heredia-Langner et al. 2004); none use $\mathcal{F}$ in tandem with exchange algorithms

Univariate Exchange Algorithms

- The machinery upon which our model-robust algorithms are based
- Used in most common commercial software (e.g. SAS, Minitab) to find $\mathcal{D}$-optimal designs
- Original algorithm by Fedorov (1972); many improvements and refinements
- I will describe the modified Fedorov algorithm (Cook and Nachtsheim 1980), but model-robust versions of other algorithms could be executed

Modified Federov Exchange Algorithm

1. Begin with a nonsingular design and construct grid of points, $\mathcal{G}$, over design space
2. For a given design point, consider exchanges with all $\mathbf{x} \in \mathcal{G}$.
3. For each potential exchange, evaluate change in determinant via updating formula
4. Choose the exchange which increases the determinant the most
5. Repeat this process for each design point
6. Continue until convergence

Computationally Cheap Determinant Updating Formula

$$|\mathbf{M}(\tilde{\xi}_n)| = |\mathbf{M}(\xi_n)|\,(1 + \Delta(\mathbf{x}_j, \mathbf{x}, \xi_n)) \qquad (2)$$

where

$$\Delta(\mathbf{x}_j, \mathbf{x}, \xi_n) = \mathbf{V}(\mathbf{x}, \xi_n) - \mathbf{V}(\mathbf{x}, \xi_n)\mathbf{V}(\mathbf{x_i}, \xi_n) + \mathbf{V}^2(\mathbf{x}, \mathbf{x}_j, \xi_n) - \mathbf{V}(\mathbf{x}_j, \xi_n) \qquad (3)$$

under the assumption that $\sigma^2 = 1$, with
$\mathbf{V}(\mathbf{x}, \xi_n) = f'(\mathbf{x})\mathbf{M}^{-1}(\xi_n)f(\mathbf{x})$ and
$\mathbf{V}(\mathbf{x}, \mathbf{x}_j, \xi_n) = f'(\mathbf{x})\mathbf{M}^{-1}(\xi_n)f(\mathbf{x}_j)$.

Outline

Motivation: Multiresponse regression

- $\mathbf{M}_m = \mathbf{Z}'(\boldsymbol{\Sigma} \otimes \mathbf{I})^{-1}\mathbf{Z}$ where

$$\mathbf{Z} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{X}_r \end{pmatrix} \tag{4}$$

- Multiresponse $\mathcal{D}$-optimal design seeks to maximize $|\mathbf{M}_m|$.

Motivation

- Multiresponse optimal design seeks a design good for all responses, which is the same goal as for single response model-robust design when using $\mathcal{F}$.
  - Note: Several months ago, we found a technical report by Emmett, Goos, and Stillman[2] which makes the same point. This work is independent of theirs.
- Problem: No good exact multiresponse optimal design algorithms in the statistical literature
- We can generalize determinant updating formula to the multiresponse case, similar to what was done by Huizenga et al. (2002) in pharmacological literature
- Multiresponse/model-robust not a perfect analog because of $\mathbf{\Sigma}$

[2] www.shef.ac.uk/pas/ResearchReports/Paper_Emmett_Nov07.pdf

Model-robust Criterion

However, if we then assume that $\mathbf{\Sigma} = \mathbf{I}$, we can show that:

$$|\mathbf{M}_m(\tilde{\xi}_n)| = \prod_{i=1}^{r} |\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i| = \prod_{i=1}^{r} \left|\mathbf{X}_i'\mathbf{X}_i\right| (1 + \Delta_i(\mathbf{x}_j, \mathbf{x}))$$

To prevent the algorithm from choosing a really bad exchange, take as the criterion:

$$\prod_{i=1}^{r}(1 + \Delta_i(\mathbf{x}_j, \mathbf{x}))\mathbb{I}(1 + \Delta_i(\mathbf{x}_j, \mathbf{x}) > 0) \quad (5)$$

Model-robust Modified Fedorov Exchange Algorithm

Then, use the modified Fedorov algorithm, except with the model-robust criterion as standard for choosing exchanges.

- For $\mathcal{F}$ composed of nested models, by theory of Bischoff (1993), MRMF designs are multiresponse $\mathcal{D}$-optimal
- Interpretation: The estimates of parameters for all models simultaneously give the minimum volume of the confidence ellipsoid of the parameters

$\mathcal{D}$-efficiencies and Methods to Compare

- For this talk, comparisons are made using $\mathcal{D}$-efficiency: $\mathcal{D}_{eff} = \left(\frac{|M_{f,\xi}|}{|M_f^*|}\right)^{(1/p_f)}$
- [DuMouchel and Jones, 1994]: Bayesian method which divides terms into *primary* and *potential*, using a prior distribution on the potential terms to regulate their impact on design
- [Heredia-Langner et al., 2004]: Genetic algorithm which uses as a model-robust criterion a desirability function involving, essentially, $\mathcal{D}_{eff}$

Example 1: Constrained Response Surface

- $n = 6$
- Design region:

$$\chi = \{\mathbf{x} = (x_1, x_2) : -1 \leq x_1, x_2 \leq 1$$
$$x_1 + x_2 \leq 1$$
$$-0.5 \leq x_1 + x_2\}$$

- $\mathcal{F} = \{f_i^{'}(\mathbf{x})\beta_i, 1 \leq i \leq 3, \mathbf{x} \in \chi\}$ where

$$f_1^{'}(\mathbf{x}) = (1, x_1, x_2) \quad (6)$$
$$f_2^{'}(\mathbf{x}) = (1, x_1, x_2, x_1x_2) \quad (7)$$
$$f_3^{'}(\mathbf{x}) = (1, x_1, x_2, x_1x_2, x_1^2, x_2^2) \quad (8)$$

Example 1

	Model			
Design	(6)	(7)	(8)	Product
MRMF	.810	.907	.995	.731
Genetic Algorithm	.849	.862	.945	.692
Bayes ($\frac{1}{\tau} = 1$)	.810	.907	.995	.731
Optimal Design for (8)	.853	.737	1	.629

Table: $\mathcal{D}$-efficiencies for constrained design with $n = 6$, protecting against three models.

Example 1

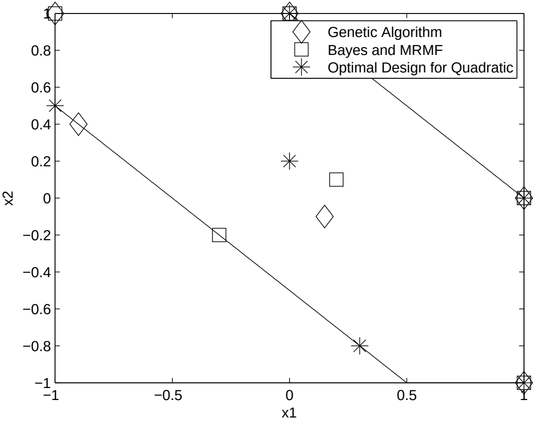


Figure: Model-robust Designs for Example 1

Example 2: Another Constrained Example

- $n = 20$ with design region

$$\begin{aligned}\chi = \{\mathbf{x} = (x_1, x_2, x_3) : -1 \leq & x_1, x_2, x_3 && \leq 1 \\ -1 \leq & x_1 + x_2 + x_3 && \leq 1 \\ -1 \leq & x_1 + x_2 && \leq 1 \\ -1 \leq & x_1 + x_3 && \leq 1 \\ -1 \leq & x_2 + x_3 && \leq 1\}\end{aligned}$$

- $\mathcal{F} = \{f_i^{'}(\mathbf{x})\beta_i, 1 \leq i \leq 5, \mathbf{x} \in \chi\}$ with

$$f_1^{'}(\mathbf{x}) = (1, x_1, x_2, x_3) \quad (9)$$

$$f_2^{'}(\mathbf{x}) = (f_1^{'}, x_1x_2, x_1x_3, x_2x_3) \quad (10)$$

$$f_3^{'}(\mathbf{x}) = (f_2^{'}, x_1^2, x_2^2, x_3^2) \quad (11)$$

$$f_4^{'}(\mathbf{x}) = (f_3^{'}, x_1^2x_2, x_1^2x_3, x_1x_2^2, x_2^2x_3, x_1x_3^2, x_2x_3^2, x_1x_2x_3) \quad (12)$$

$$f_5^{'}(\mathbf{x}) = (f_4^{'}, x_1^3, x_2^3, x_3^3) \quad (13)$$

Example 2

	Model					
Design	(9)	(10)	(11)	(12)	(13)	Product
MRMF	.864	.756	.870	.955	.979	.531
Bayes ($\frac{1}{\tau} = 1$)	.867	.749	.860	.892	.999	.498
Bayes ($\frac{1}{\tau} = 16$)	.843	.731	.872	.852	.970	.444
Optimal for (13)	.860	.744	.859	.878	1	.483

Table: $\mathcal{D}$-efficiencies for fictitious 3-factor constrained design with $n = 20$, protecting against five models.

Example 3: Plastic Formulation Example

- Mixture experiment with $n = 25$
- Response is hardness of plastic
- Five factors were a binder (x_1), cobinder (x_2), plasticizer (x_3), and two monomers (x_4 and x_5)
- Design region

$$\chi = \{\mathbf{x} = (x_1, \ldots, x_5) : \begin{array}{l} \sum_{i=1}^{5} x_i = 1, \\ 0.50 \le x_1 \le 0.70, \\ 0.05 \le x_2 \le 0.15, \\ 0.05 \le x_3 \le 0.15, \\ 0.10 \le x_4 \le 0.25, \\ 0.00 \le x_5 \le 0.15, \\ 0.18 \le x_4 + x_5 \le 0.26, \\ 0.00 \le x_3 + x_4 + x_5 \le 0.35\} \end{array}$$

- $\mathcal{F} = \{f_i'(\mathbf{x})\beta_i, 1 \leq i \leq 4, \mathbf{x} \in \chi\}$ with

$$f_1'(\mathbf{x}) = (\{x_i, i = 1, \ldots, 5\}) \tag{14}$$

$$f_2'(\mathbf{x}) = (f_1', \{x_i x_j, i < j \leq 5\}) \tag{15}$$

$$f_3'(\mathbf{x}) = (f_2', \{x_i x_j x_k, i < j < k \leq 5\}) \tag{16}$$

- Used as candidate list the extreme vertices and approximate centroids found using code by Piepel (1988)

Example 3

	Model			
Design	(14)	(15)	(16)	Product
MRMF (50 tries)[3]	.822	.955	.950	.746
Bayesian ($\frac{1}{\tau} = 1$, 50 tries)	.805	.940	.965	.730
Optimal for (16)[4]	.800	.924	1	.739

Table: $\mathcal{D}$-efficiencies for constrained mixture example with $n = 25$, protecting against three models.

[3] Product of determinants is less than the Bayesian design; more algorithm tries found a similar design.

[4] larger candidate list

Example 4: Mixture Experiment with Disparate Possible Models

Pharmaceutical experiment (Frisbee and McGinity 1994)

- Response was glass transition temperature of a film obtained from a pseudolatex
- Three nonionic surfactant mixture factors
- $n = 11$, and models of interest may be:

$$f_1'(\mathbf{x}) = (\{x_i, i = 1, 2, 3\}) \tag{17}$$

$$f_2'(\mathbf{x}) = (f_1', \{x_i x_j, i < j \leq 3\}) \tag{18}$$

$$f_3'(\mathbf{x}) = (f_1', \{min(x_i, x_j), i < j \leq 3\}) \tag{19}$$

- Though Frisbee and McGinity fit a model like (18), Rajagopal and Castillo showed that a Becker model like (19) also fit well.
- These disparate models cannot easily be fit into the Bayesian framework.

Outline

Discussion

- Presented a practical, exchange algorithm-based design method that relaxes optimal design model assumption
- Simply require experimenters to specify possible models of interest
- Explicitly considers all models of interest and has multiresponse $\mathcal{D}$-optimal interpretation for nested models
- Can handle models of disparate types

References

DuMouchel, W. and Jones, B. (1994).
A simple bayesian modification of d-optimal designs to reduce dependence on an assumed model.
Technometrics, 36(1):37–47.

Heredia-Langner, A., Montgomery, D. C., Carlyle, W. M., and Borror, C. M. (2004).
Model-robust optimal designs: A genetic algorithm approach.
Journal of Quality Technology, 36(3):263–279.