Bayesian Factor Screening and Response Surface Designs

Abhyuday Mandal, Derek Bingham, Hugh Chipman
University of Michigan, University of Michigan, University of Waterloo

Introduction
Standard practice in RSM is to first perform factor screening, followed by response surface exploration.

Cheng and Wu (2001) proposed methodology for RSM which performs the two objectives with the same design.

Approach assumes that each interaction effect can be active ONLY if both parents are significant. Assumption is restrictive and can lead to model misspecification.

We propose a Bayesian approach that attempts to more directly incorporate the standard assumptions of industrial experiments into both the design and analysis.

Potential Problems
Approach only considers interactions at the 2nd stage.
Thus, significant interactions are only identified if both parents are significant (strong heredity).

If the true model contains interactions with only one significant parent (weak heredity), then:
1. spurious effects may be identified,
2. the effect estimates can be biased,
3. the weak heredity interactions will be missed.

IDEA: A design which can distinguish between models is preferable.

New criterion:
\[ HD = \sum_{i,j} P(M_i) P(M_j) H(f_i, f_j). \]

where, \( H(f_i, f_j) = \left( I(f_i - f_j)^2 \right)^{1/2} \) is the Hellinger distance between the predictive densities \( f_i \) and \( f_j \) for models \( M_i \) and \( M_j \).
The design enters the criterion through the predictive densities.
The distance measure aims to reduce model aliasing.
The pair-wise distances are weighted by the model priors determined in (1).
Criterion is bounded: \( 0 \leq HD \leq 1 \).

Design Approach

Example 1
Using model priors as in (1), the optimal design is not an Orthogonal Array.

Each column has 3 zeros and 7 or 8 entries having values ±1.

Example 2
Optimal design according to HD criterion is an 18 run OA when the following priors are used.

Simulation:
Models were randomly generated from prior (1).
Prior was calibrated so \( E(#\text{effects}) = 5 \), giving \( p = 0.344 \).

Two-Stage Analysis
Situation: Want to perform screening and surface exploration in the same region. To save time and resources, will do both steps with the same design.

Assumptions: Effect sparsity, effect hierarchy, and strong heredity.

Analysis Approach
We use Bayesian variable selection for designs with complex aliasing (Chipman, Hamada and Wu, 1997).

Will use usual set-up for priors in a normal linear model. Unique feature is hierarchical priors for the probability that an effect is active:

\[
\begin{align*}
 p(\delta_A=0) &= p \\
 p(\delta_{ij}=0) &= \begin{cases} p & \text{if } \delta_i = 1 \\
 0.1p & \text{if } \delta_i = 0 \\
 \end{cases} \\
 p(\delta_{ij}=1) &= \begin{cases} 0.3p & \text{if } \delta_i = \delta_j = 0 \\
 0.5p & \text{if } \delta_i + \delta_j = 1 \\
 1 & \text{if } \delta_i + \delta_j = 2 \\
 \end{cases} \\
 p(\delta_{ij}=2) &= \begin{cases} 1 & \text{if } \delta_i + \delta_j = 2 \\
 0.3p & \text{if } \delta_i + \delta_j = 1 \\
 0.1p & \text{if } \delta_i + \delta_j = 0 \\
 \end{cases}
\end{align*}
\]

The Gibbs sampler is used to find the probability an effect is active, given the data.

Conclusions
New design/analysis framework for RSM.
Simulation suggests with this approach identifies the correct model more often under usual assumptions.
Using the HD criterion, designs can be obtained for any run size.
Can directly incorporate sparsity and heredity assumptions into criterion.

Acknowledgements
This research was supported by NSF grant DMS-0103886.
The authors are grateful to the DAE 2003 organizers for their support.