A New Approach to the Construction and Analysis of Supersaturated Designs

Bradley Jones
Principal Research Fellow
JMP Division of SAS

Dibyen Majumdar
University Illinois at Chicago

Statistical Discovery. From SAS®

Outline

- Introduction to Supersaturated Design (SSD)
- Bayesian D-Optimal Supersaturated designs
- D-Optimal Supersaturated Designs
- Ideas for Construction
- Analysis
- Summary

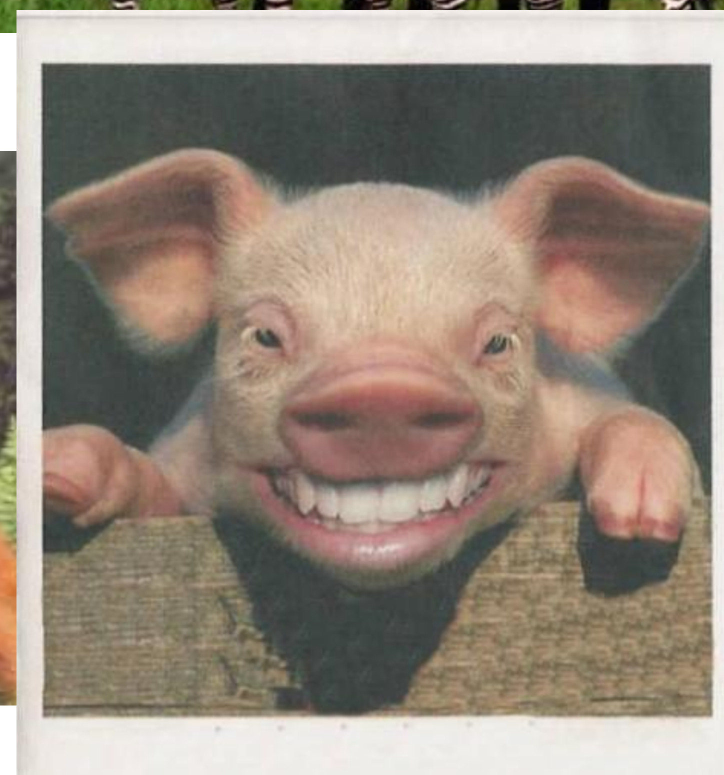
What is a supersaturated design?

Supersaturated designs have more factors than runs.

That is you might be interested in the possible effects of 24 factors but only have the budget for 20 runs.

At first this may seem laughable…


DASH村



Why the laughter?

Are supersaturated designs a bad idea?

“Supersaturated designs are evil.” Randall Tobias

Why did my colleague say this?

1. Design matrix is singular so multiple regression fails.
2. Factor aliasing is complex.
3. “You can’t get something for nothing.”

Early Literature

Booth and Cox (1962) $E(s^2)$ criterion

$$E(s^2) = \frac{\sum_{i>j} (x_i^T x_j)^2}{\frac{k(k-1)}{2}}$$

Booth, K.H.V. and Cox, D.R. (1962), "Some Systematic Supersaturated Designs," *Technometrics*, **4**, 489-495.

Re-introduction in the Literature

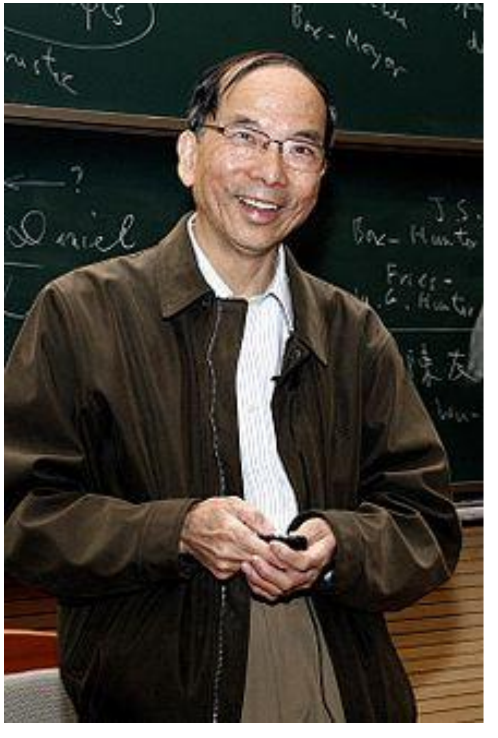
Lin (1993) construction using Hadamard matrix

Lin, D. K. J. (1993), "A New Class of Supersaturated Designs," *Technometrics*, **35** 28-31.

Re-introduction in the Literature

Wu (1993) construction using partially aliased interactions

Wu, C. F. J. (1993) Construction of supersaturated designs through partially aliased interactions. Biometrika, 80, 661-669.



Why would you consider using such a design?

1. Runs are expensive.
2. Brainstorming often yields dozens of possible factors.
3. You don’t want to eliminate factors in absence of data.

Outline

- Introduction to Supersaturated Design
- Bayesian D-Optimal Supersaturated designs
- D-Optimal Supersaturated Designs
- Ideas for construction
- Analysis

D-Optimal Design Definition

Given the usual linear regression model

$$y = X\beta + \varepsilon$$

find a design matrix, *X*, to maximize the determinant of the information matrix, $\left| X^T X \right|$

But, if there are more factors than runs, this determinant is always zero no matter what *X* is.

Bayesian D-Optimal designs

Find a design matrix, *X*, to maximize

$$D_{Bayes} = \left| X^T X + I / \gamma \right|$$

where γ is a tuning parameter.

This determinant is never zero and you can improve it with a clever choice of X.

Benefits of Bayesian D-Optimal Supersaturated Design

1. Easy and fast to compute
2. Flexible formulation (sample size, factor type, etc.)

References:

DuMouchel and Jones, Technometrics (1994) vol.36 #1 pp. 37-47.

Jones, B., Lin, D., and Nachtsheim, C. (2008) “Bayesian D-Optimal Supersaturated Designs.” *Journal of Statistical Planning and Inference*, 138, 86-92.

Drawbacks of Bayesian D-Optimal Design

1. Bayesian *D*-Optimality requires you to be a Bayesian…
2. What about that tuning parameter, γ?
3. You need an optimal design algorithm to generate these.
4. Optimization code can take a while to run and you are rarely sure that you have the global optimum.

Outline

- Introduction to Supersaturated Design
- Bayesian D-Optimal Supersaturated designs
- D-Optimal Supersaturated Designs
- Ideas for Construction
- Analysis

D-Optimal Definition – Supersaturated Designs

Find a design matrix, *X*, to maximize

$$\left| XX^T \right|$$

XX^T is n x n so you can always find an X so that the determinant above is nonzero.

n is the number of experimental runs.

Why use this new criterion?

1. Does not require a Bayesian framework.
2. No tuning parameter.
3. Results in good designs – minimum bias estimates
4. Is a limiting case of Bayesian D-Optimality.

   Prior variance goes to infinity
5. Is also a limiting case of Es^2 optimality. *

   Prior variance goes to zero.
6. Fast (and sometimes no) computation.

Estimating effects for supersaturated designs

Consider the standard linear model.

$$Y = X\beta + \varepsilon$$

Suppose there are n runs and p parameters and $n < p$.

$$Y = X\beta + \varepsilon$$

Re-write the model as below.

$$X\beta = XX'(XX')^{-1}X\beta$$

$$X\beta = XX'\gamma$$

Note that γ is n x 1. So, equivalently,

$$Y = XX'\gamma + \varepsilon$$

We then estimate γ as,

$$\widehat{\gamma} = (XX')^{-1}Y$$

This leads to the minimum bias estimator for β,

$$\widehat{\beta}^* = X'\widehat{\gamma} = X'(XX')^{-1}Y$$

So, given our minimum bias estimator,

$$\widehat{\beta}^* = X'\widehat{\gamma} = X'(XX')^{-1}Y$$

the variance of this estimator is

$$V(\widehat{\beta}^*) = \sigma^2 X'(XX')^{-2}X$$

This leads us to the previously stated optimality criterion.

What about our D-optimality criterion?

$V(\widehat{\beta}^{*})$ is singular so its determinant is zero.

But we can minimize the product of its non-zero eigenvalues. It turns out that minimizing that product is the same as maximizing

$$|XX'|$$

But that is the D-optimality criterion we suggest.

Outline

- Introduction to Supersaturated Design
- Bayesian D-Optimal Supersaturated designs
- D-Optimal Supersaturated Designs
- Ideas for Construction
- Analysis

Question:

Can we generate globally “D-optimal” supersaturated designs?

Answer: Yes!

Note: I will show how to do this where $p \bmod 4 = 0$ but there are also constructions for the other 3 cases.

Procedure:

1. Choose n rows of a pxp Hadamard matrix where $n < p$
2. Make sure that the first column of the Hadamard matrix is all +1.
3. Do this by multiplying rows by -1 if necessary.

Call the result, X

The $p - 1$ non-constant columns of X are the factor settings.

X is a globally D-optimal supersaturated design XX^T equals nI where n is the number of runs and I is the identity matrix.

Here is a Hadamard matrix with 8 rows and 8 columns.

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1, \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1, \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1, \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1, \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1, \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1, \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1, \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

$$H^T H = HH^T = 8I$$

where I is the identity matrix

Choose *any* 6 rows.

I	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
1	-1	1	-1	1	-1	1	-1
1	1	-1	-1	1	1	-1	-1
1	-1	-1	1	1	-1	-1	1
1	1	1	1	-1	-1	-1	-1
1	-1	1	-1	-1	1	-1	1

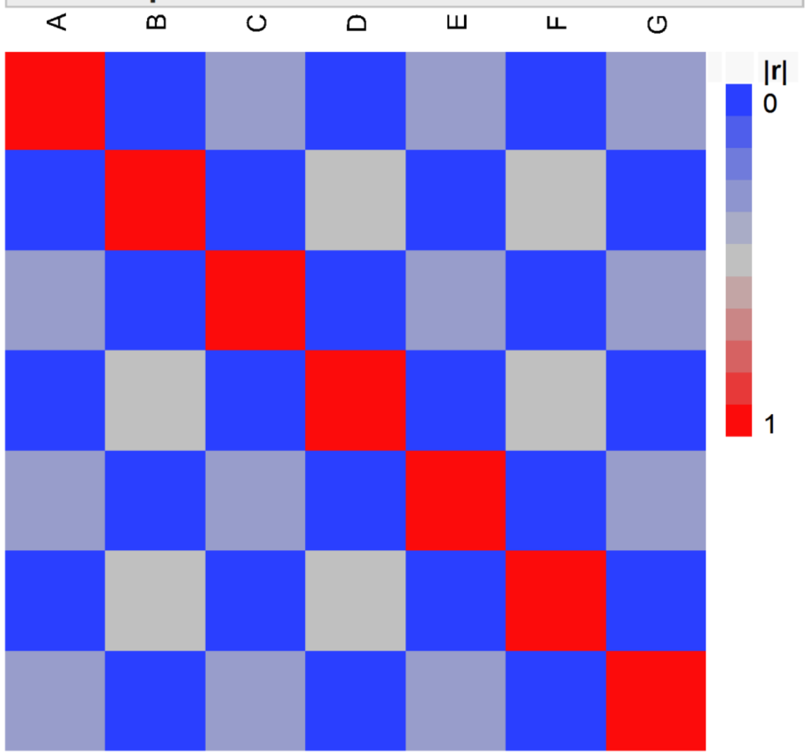
Call the above X. Then $XX^T = 8\ I_6$

Globally D-optimal Supersaturated Design

Design

Run	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1
2	-1	1	-1	1	-1	1	-1
3	1	-1	-1	1	1	-1	-1
4	-1	-1	1	1	-1	-1	1
5	1	1	1	-1	-1	-1	-1
6	-1	1	-1	-1	1	-1	1

Color Map On Correlations



But in our example, there are 28 such designs.

So, we need a secondary criterion to help us choose which n rows of the Hadamard matrix to use.

Ideas for a secondary criterion:

1. Minimize the maximum squared off-diagonal element of X^TX.

2. Minimize the sum of squared off-diagonal elements of X^TX.

3. Use “extra” columns in X in a clever way.

Let’s think about option 3.

Note: Many constructions require the column sums to be zero.

The usual definition of Es^2 assumes this restriction. We have lower bounds for an unrestricted Es^2 (UEs^2) that allows for unbalanced columns.

Hadamard matrices exist with numbers of rows that are multiples of 4. Therefore, the number of columns in X will also be a multiple of 4.

Generally, the number of factors is not a multiple of 4. What can we do with the “extra” columns?

By *extra* I mean the columns we do not assign to factors.

Answer:

If we are clever, we can use them to aid in model selection.

Let us partition X into W and Z where the columns of W are our factors and the columns of Z are our extra columns.

Suppose we can make $W^T Z = 0$.

That means that the extra columns, Z, are orthogonal to the factor columns. The rank of W+Z is the rank of X, *n*.

Usually we make W have rank n. But here we make Z have rank, *q* so W has rank *n*-*q*.

That is, we make W even more supersaturated. In exchange we get an estimator for the error variance that we can use in model selection to screen the factors.

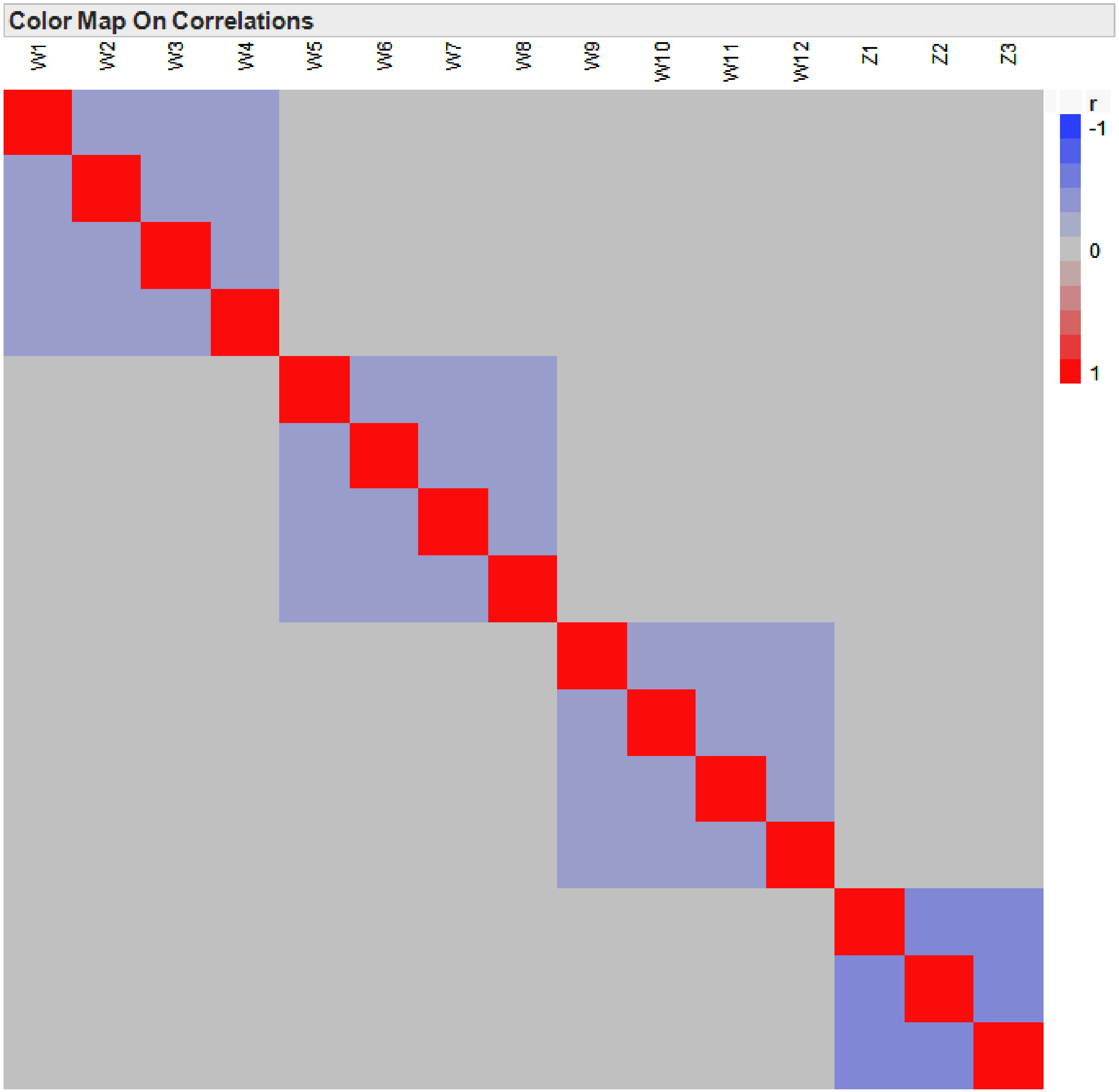
Example 1

Suppose there are 12 factors and we can do12 runs.

Take the 16x16 Hadamard matrix and remove 4 rows to make $W^T Z = 0$

Intercept	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	Z1	Z2	Z3
1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1	1	-1	-1
1	1	-1	1	-1	-1	1	-1	1	-1	1	1	-1	-1	-1	1
1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	-1	1	-1
1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	-1	-1
1	-1	-1	1	1	-1	-1	1	1	1	1	-1	-1	1	-1	-1
1	1	-1	-1	1	1	-1	-1	1	-1	1	-1	1	-1	1	-1
1	-1	1	-1	1	-1	1	1	-1	-1	1	1	-1	-1	1	-1
1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	-1	1
1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	1
1	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	-1	1
1	-1	1	1	-1	-1	1	-1	1	1	-1	-1	1	-1	1	-1
1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1

Correlation Color Plot



Properties of Example 1 Design

1. 3 independent groups of 4 factors.
2. Each group has rank 3.
3. Extra columns have rank 2 so you can estimate σ^2 with 2 df.

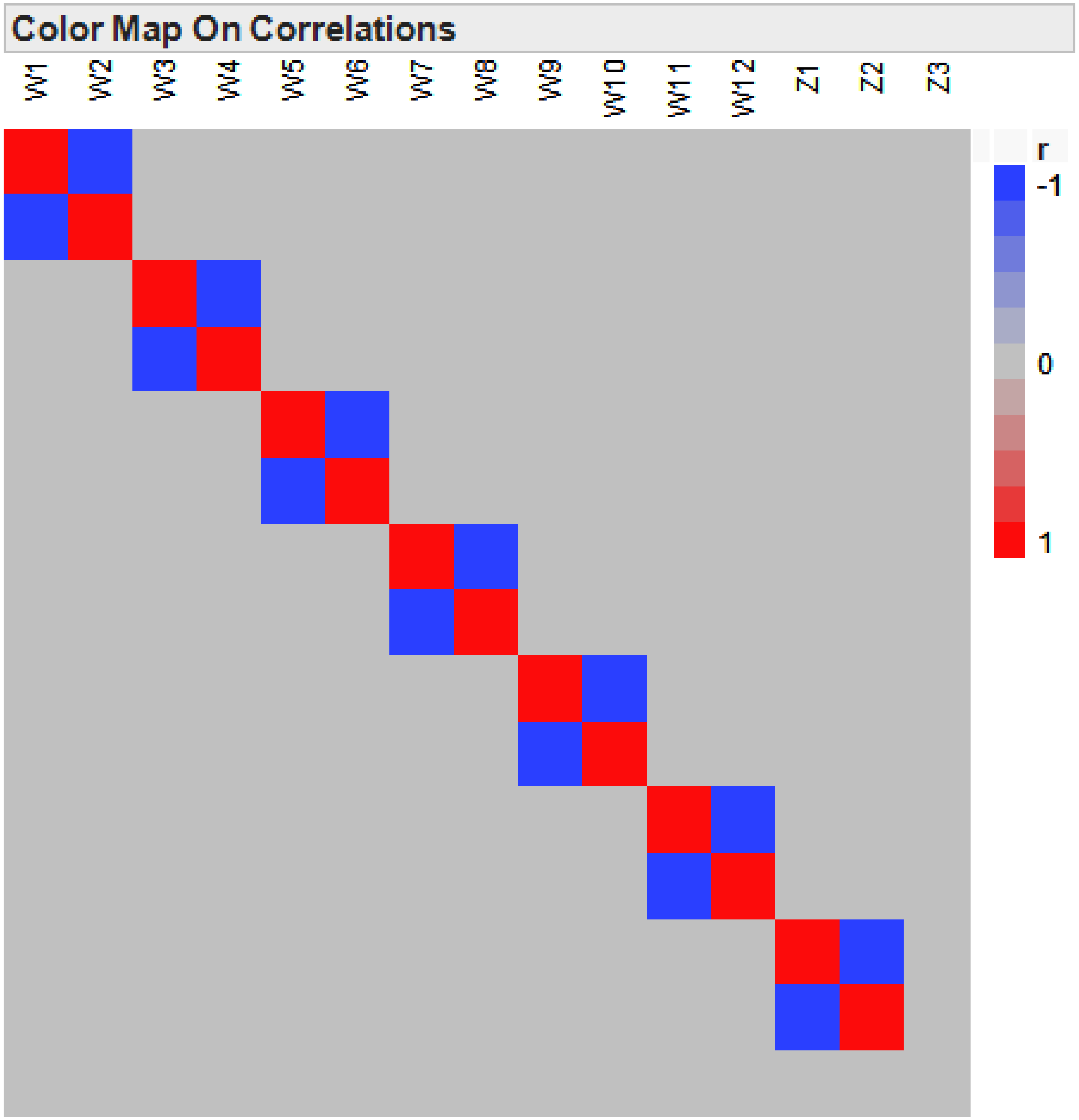
Example 2

Suppose there are 12 factors and we can do 8 runs.

Use the same procedure as before.

I	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	Z1	Z2	Z3
1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1
1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1
1	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1
1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1
1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	1	1	-1	-1
1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1
1	-1	1	-1	1	1	-1	1	-1	1	-1	1	-1	-1	1	-1
1	1	-1	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1

Correlation Plot



Properties of Example 2 Design

1. 6 independent groups of 2 factors.
2. Each group has rank 1. Main effects are confounded in pairs.
3. Extra columns have rank 1 so you can estimate σ^2 with 1 df.

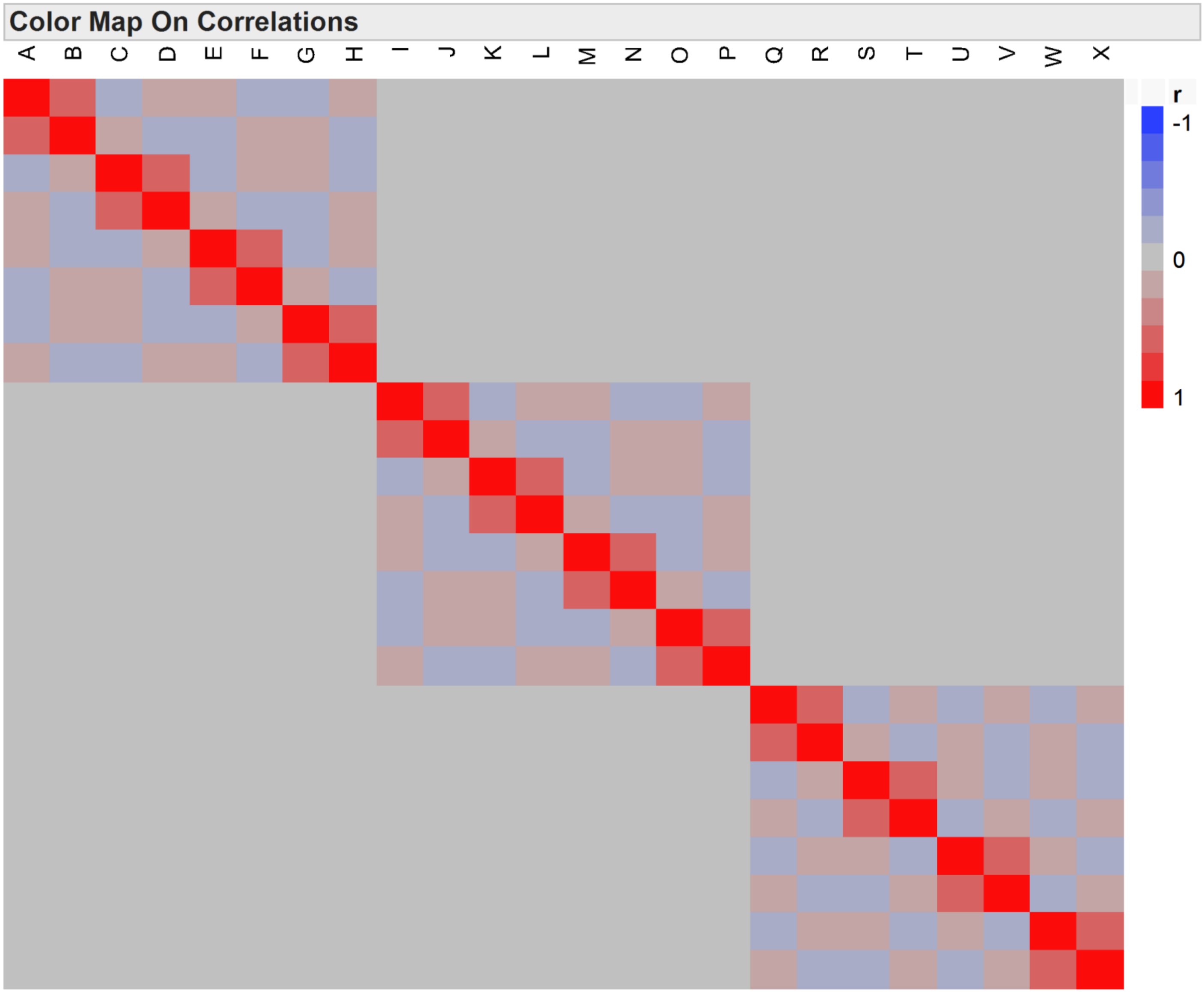
This may seem odd, but this is a group screening design.
These have been advocated by Vine, et. al.

A. E Vine, S. M Lewis, A. M Dean, D Brunson. A Critical Assessment of Two-Stage Group Screening Through Industrial Experimentation Technometrics. February 1, 2008, 50(1): 15-25.

Example 3 – 24 Factors in 20 Runs

+	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+	+	−	−	+	−	+	+	−
−	−	+	+	+	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−
+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+
+	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+
+	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−	−	−	+	+	+	+	−	−
−	−	+	+	−	−	+	+	+	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−
+	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+
−	+	−	+	−	+	−	+	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−
+	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+	−	−	−	−
−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+
+	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−	+	+	−	−
−	−	−	−	+	+	+	+	+	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−
−	+	+	−	+	−	−	+	+	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−
+	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+
−	−	+	+	+	+	−	−	−	−	+	+	+	+	−	−	+	+	−	−	−	−	+	+
+	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+	+	−	+	−	+	−	+	−
−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
+	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+
−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	+
−	+	+	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−	+	−	−	+

Example 3 – 24 Factors in 20 Runs



Properties of Example 3 Design

1. 3 independent groups of 8 factors.
2. Each group has rank 5.
3. Extra columns have rank 4 so you can estimate σ^2 with 4 df.

How did we generate these designs again?

We want two things:

1) Make the factor column correlations small – min $E\,s^2$

2) Make the $W^TZ = 0$

We use a combined optimization criterion by minimizing the sum of $E\,s^2$ and the sum of squares of W^TZ

We use a row exchange algorithm with the Hadamard matrix as the candidate set.

Outline

- Introduction to Supersaturated Design
- Supersaturated Bayesian D-Optimal designs
- D-Optimal Supersaturated Designs
- Ideas for Construction
- Analysis

Procedure

1. Do a principal components analysis for each of the 3 groups of 8 factors.

   Note that you can explain each group with 5 PCs.

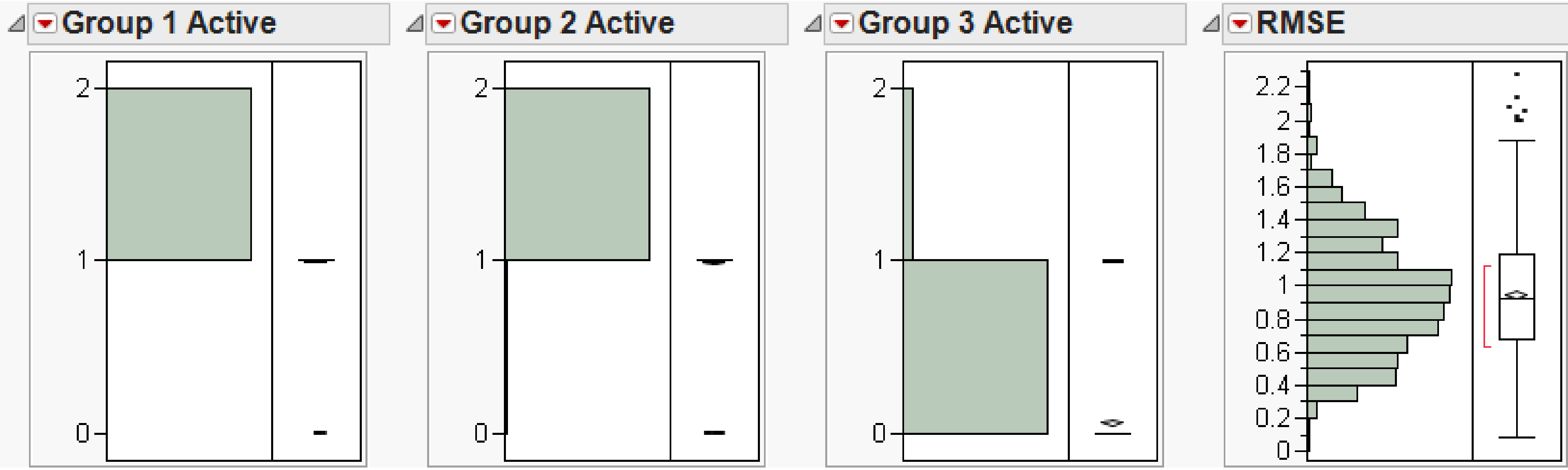
2. Compute the sum of squares for each factor group by regressing the response on each set of PCs.
3. Compute the error sum of squares from the extra columns.
4. Do a separate F-test for each factor group.
   1. F = (Group SS/5)/(SSE/4)
   2. p = 1 – F CDF (F_{α},num df,denom df)

Demonstration

$y = 100 + 1.5(A + B + I + P) + \varepsilon$

where $\varepsilon \sim N(0,1)$

So, Group 1 and Group 2 are active but Group 3 is not.



Summary

1. Introduced a new optimality criterion for SSDs.
2. Gave some ideas for design construction.
3. Showed examples of designs using the "orthogonal extra columns" construction idea.
4. Demonstrated an analytical approach for the new supersaturated design type.

Statistical Discovery.™ From SAS.

SEEING IS BELIEVING