The Isle of Misfit Designs

A Guided Tour of Optimal Designs That Break the Mold

Dr. Caleb King
Research Statistician Tester
JMP Division, SAS Institute Inc.

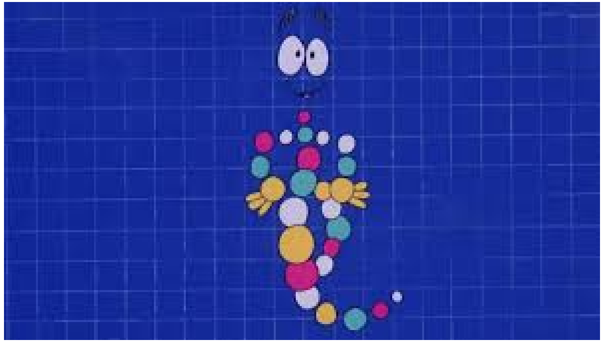
The "Typical" Screening Design

a.k.a. Factorial/Fractional Factorial Designs

- Each factor typically has two levels (usually expressed as 1/-1 or +/ -)
  - *Definitive screening designs* are an exceptional exception.
- The design is "orthogonal"
  - i.e. the main effect terms are uncorrelated with one another
  - However, they may be correlated or even aliased with interactions
- The design is "optimal" in some sense
  - D-optimal, A-optimal, I-optimal, etc.
- The number of runs (N) is always some multiple of 4
  - An orthogonal design only exists if this is the case
- What if I told you that there existed D-optimal designs where the number of runs is **NOT** a multiple of 4?

Welcome… to the Isle of Misfit Designs

- What makes these designs "misfits"?
  - They have run sizes that **<u>are not</u>** multiples of 4.
  - However, they also **<u>are not completely orthogonal designs</u>**.
    - There will be some correlation among the main effects
- Let's take a tour shall we?
  - But first…



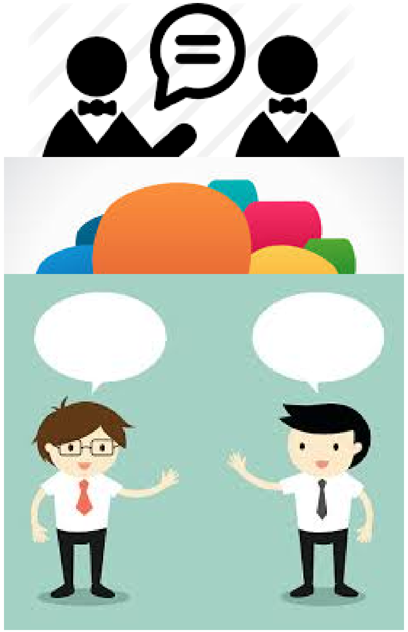
The Information Matrix

- A typical design matrix looks like $X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & -1 \end{bmatrix}$
- The covariance matrix is $(X^T X)^{-1}$ or a similar variation
- We'll actually look at the *information/precision matrix*, which is just $X^T X$
  - Most of the information about these "misfit" designs is more easily expressed through their information matrices.
- In the information matrix, we are primarily interested in the off-diagonal terms
  - Since we're focusing on two-level factors only, the diagonal terms will always be the same (N).

How Does it Work?

The Noisy Room Analogy

- Let’s consider a simple arbitrary information matrix
  - $$\begin{bmatrix} \ddots & \vdots & \vdots & \vdots & ⋰ \\ \cdots & N & a & a & \cdots \\ \cdots & a & N & a & \cdots \\ \cdots & a & a & N & \cdots \\ ⋰ & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$
- Think of that value “a” as a volume control
  - $a = 0$ (Ideal case!)
  - $a = \pm N$
  - $0 \leq |a| \leq N$
- The closer the volume is to 0, the more information we can gather about each factor with less “noise pollution”



A Quick Note on Optimality

- Factorial/fractional factorial screening designs are typically optimal for several different criteria
- Once you move away from those designs, your choice of optimality criterion may have a bigger impact
  - We'll focus on the D-criterion as it's the most commonly used
  - Remember, "optimal" doesn't necessarily mean "best overall"

Let the Tour Begin!

- Some final tour info…
  - There are four "exhibits" distinguished by how far the run size (N) is from a multiple of 4.
  - We'll primarily be looking at the optimal information matrices ($\boldsymbol{M}^*$) of each case as well as some other properties.

    - Variance Inflation Factors (VIF) – How much is the actual variance inflated beyond the ideal case (true variance/(1/N))
    - Variance Reduction Factors (VRF) – How much is the variance reduced from the previous design(s) (1-variance of current design/variance of previous design)
  - There won't be much emphasis on construction methods beyond very basic procedures.

All Aboard!!

Exhibit 0

$$X_{4,3} = \begin{bmatrix} 1 & + & + & + \\ 1 & + & - & - \\ 1 & - & + & - \\ 1 & - & - & + \end{bmatrix} \text{->} \quad \boldsymbol{M}^* = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

$$X_{8,7} = \begin{bmatrix} 1 & + & + & + & + & + & + & + \\ 1 & + & + & - & + & - & - & - \\ 1 & + & - & + & - & + & - & - \\ 1 & + & - & - & - & - & + & + \\ 1 & - & + & + & - & - & + & - \\ 1 & - & + & - & - & + & - & + \\ 1 & - & - & + & + & - & - & + \\ 1 & - & - & - & + & + & + & - \end{bmatrix} \text{->} \quad \boldsymbol{M}^* = \begin{bmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix}$$

Exhibit 1: N=4t+1

- $$\boldsymbol{M}^* = \begin{pmatrix} \boldsymbol{N} & \mathbf{-1} & \mathbf{-1} & \cdots & \mathbf{-1} \\ \mathbf{-1} & \boldsymbol{N} & \mathbf{1} & \cdots & \mathbf{1} \\ \mathbf{-1} & \mathbf{1} & \boldsymbol{N} & \cdots & \mathbf{1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{-1} & \mathbf{1} & \mathbf{1} & \cdots & \boldsymbol{N} \end{pmatrix}$$
- If the design is not saturated (# factors < N-1), then you can easily construct by simply adding a run to a factorial/fractional factorial design with N-1 runs.
- For saturated designs, achieving $\boldsymbol{M}^*$ is only possible if $\sqrt{2N-1}$ is an integer.
  - Requires more specialized construction techniques.
- The designs that achieve $\boldsymbol{M}^*$ are actually both D-optimal and A-optimal!
  - For cases where you can't achieve $\boldsymbol{M}^*$, D-optimal designs do exist, but they are more specialized.

Exhibit 1 Examples

$$X_{5,3} = \begin{bmatrix} 1 & + & + & + \\ 1 & + & - & - \\ 1 & - & + & - \\ 1 & - & - & + \\ 1 & - & - & - \end{bmatrix} \text{->}\ \boldsymbol{M}^* = \begin{bmatrix} 5 & -1 & -1 & -1 \\ -1 & 5 & 1 & 1 \\ -1 & 1 & 5 & 1 \\ -1 & 1 & 1 & 5 \end{bmatrix}$$

VIF=1.094
VRF=0.125

$$X_{5,4} = \begin{bmatrix} 1 & + & + & + & + \\ 1 & + & - & - & - \\ 1 & - & + & - & - \\ 1 & - & - & + & - \\ 1 & - & - & - & + \end{bmatrix} \text{->}\ \boldsymbol{M}^* = \begin{bmatrix} 5 & -1 & -1 & -1 & -1 \\ -1 & 5 & 1 & 1 & 1 \\ -1 & 1 & 5 & 1 & 1 \\ -1 & 1 & 1 & 5 & 1 \\ -1 & 1 & 1 & 1 & 5 \end{bmatrix}$$

VIF=1.1111

Exhibit 1 Examples

$$X_{9,8} = \begin{bmatrix} 1 & + & + & + & + & + & + & + & + \\ 1 & + & - & - & - & - & - & - & - \\ 1 & - & + & - & - & + & + & - & - \\ 1 & - & - & + & - & + & - & + & - \\ 1 & - & - & - & + & + & - & - & + \\ 1 & - & + & + & + & - & - & - & - \\ 1 & - & + & - & - & - & - & + & + \\ 1 & - & - & + & - & - & + & - & + \\ 1 & - & - & - & + & - & + & + & - \end{bmatrix} \text{->} \boldsymbol{M}^* = \begin{bmatrix} 9 & -5 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -5 & 9 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 9 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 9 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 9 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 9 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 & 9 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 & 1 & 9 & 1 \\ -1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 9 \end{bmatrix}$$

Intercept and Factor 1 -> VIF=1.469

Remaining Factors -> VIF=1.056

Exhibit 1 Detour

$$X_{5,3} = \begin{bmatrix} 1 & + & + & + \\ 1 & + & - & - \\ 1 & - & + & - \\ 1 & - & - & + \\ 1 & - & - & - \end{bmatrix} \text{->} \ \boldsymbol{M}^* = \begin{bmatrix} 5 & -1 & -1 & -1 \\ -1 & 5 & 1 & 1 \\ -1 & 1 & 5 & 1 \\ -1 & 1 & 1 & 5 \end{bmatrix}$$

VIF=1.09375
VRF=0.125

$$X'_{5,3} = \begin{bmatrix} 1 & + & + & + \\ 1 & + & - & - \\ 1 & - & + & - \\ 1 & - & - & + \\ 1 & 0 & 0 & 0 \end{bmatrix} \text{->} \ \boldsymbol{M}^* = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

Intercept VIF=1
Intercept VRF=0.2

Factor VIF=1.25
Factor VRF=0.0

Exhibit 2: N=4t+2

- $\boldsymbol{M}^* = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A} \end{pmatrix}$, where $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{N} & \mp\boldsymbol{2} & \mp\boldsymbol{2} & \mp\boldsymbol{2} \\ \mp\boldsymbol{2} & \boldsymbol{N} & \cdots & \boldsymbol{2} \\ \mp\boldsymbol{2} & \vdots & \ddots & \boldsymbol{2} \\ \mp\boldsymbol{2} & \boldsymbol{2} & \boldsymbol{2} & \boldsymbol{N} \end{pmatrix}$ [-2 for intercept]
- If the design is not saturated nor near-saturated (# factors < N-2), then you can easily construct by adding two orthogonal (or as close to orthogonal as you can get) rows to a standard design with N-2 runs.
- For saturated designs, achieving $\boldsymbol{M}^*$ is only possible if $N - 1$ is the sum of two squared integers.
  - For near-saturated designs, just delete a row from the saturated design.
  - For the saturated case, more specialized construction methods are needed.
- Designs that achieve $\boldsymbol{M}^*$ are also both D-optimal and A-optimal!

Exhibit 2 Examples

$$X_{6,3} = \begin{bmatrix} 1 & + & + & + \\ 1 & + & - & - \\ 1 & - & + & - \\ 1 & - & - & + \\ 1 & - & - & - \\ 1 & - & + & + \end{bmatrix} \text{->} \quad \boldsymbol{M}^* = \begin{bmatrix} 6 & -2 & 0 & 0 \\ -2 & 6 & 0 & 0 \\ 0 & 0 & 6 & 2 \\ 0 & 0 & 2 & 6 \end{bmatrix}$$

VIF=1.125
VRF=0.143 (5 runs)
VRF=0.25 (4 runs)

$$X_{6,5} = \begin{bmatrix} 1 & + & + & + & + & + \\ 1 & + & - & - & - & - \\ 1 & - & + & - & - & - \\ 1 & - & - & - & + & + \\ 1 & - & - & + & - & + \\ 1 & - & - & + & + & - \end{bmatrix} \text{->} \quad \boldsymbol{M}^* = \begin{bmatrix} 6 & -2 & -2 & 0 & 0 & 0 \\ -2 & 6 & 2 & 0 & 0 & 0 \\ -2 & 2 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & 2 & 2 \\ 0 & 0 & 0 & 2 & 6 & 2 \\ 0 & 0 & 0 & 2 & 2 & 6 \end{bmatrix}$$

VIF=1. 2

Exhibit 3: N=4t+3

- $\boldsymbol{M^*} = \begin{pmatrix} \boldsymbol{A_u} & \boldsymbol{\pm 1} \\ \boldsymbol{\pm 1} & \boldsymbol{A_v} \end{pmatrix}$ [+1 for intercept]
  - $\boldsymbol{A_j} = \begin{pmatrix} \boldsymbol{A_{11}} & \boldsymbol{\pm 1} & \boldsymbol{\pm 1} & \boldsymbol{\pm 1} \\ \boldsymbol{\pm 1} & \boldsymbol{A_{12}} & \cdots & \boldsymbol{-1} \\ \boldsymbol{\pm 1} & \vdots & \ddots & \boldsymbol{-1} \\ \boldsymbol{\pm 1} & \boldsymbol{-1} & \boldsymbol{-1} & \boldsymbol{A_{1j}} \end{pmatrix}, \boldsymbol{j} = \boldsymbol{u} \text{ or } \boldsymbol{v}$

    [+1 for intercept]

  - $\boldsymbol{A_{ij}} = \begin{pmatrix} \boldsymbol{N} & \boldsymbol{\mp 3} & \boldsymbol{\mp 3} & \boldsymbol{\mp 3} \\ \boldsymbol{\mp 3} & \boldsymbol{N} & \cdots & \boldsymbol{3} \\ \boldsymbol{\mp 3} & \vdots & \ddots & \boldsymbol{3} \\ \boldsymbol{\mp 3} & \boldsymbol{3} & \boldsymbol{3} & \boldsymbol{N} \end{pmatrix}$
  - $\boldsymbol{A_{iu}}$ is of size r, and $\boldsymbol{A_{iv}}$ is of size $r + 1$
- If # factors $\leq \frac{N+1}{2} + 2$, then you can easily construct by deleting a row from a standard design with N+1 runs.
- This case requires the most specialized constructions as it is the most difficult one to deal with.

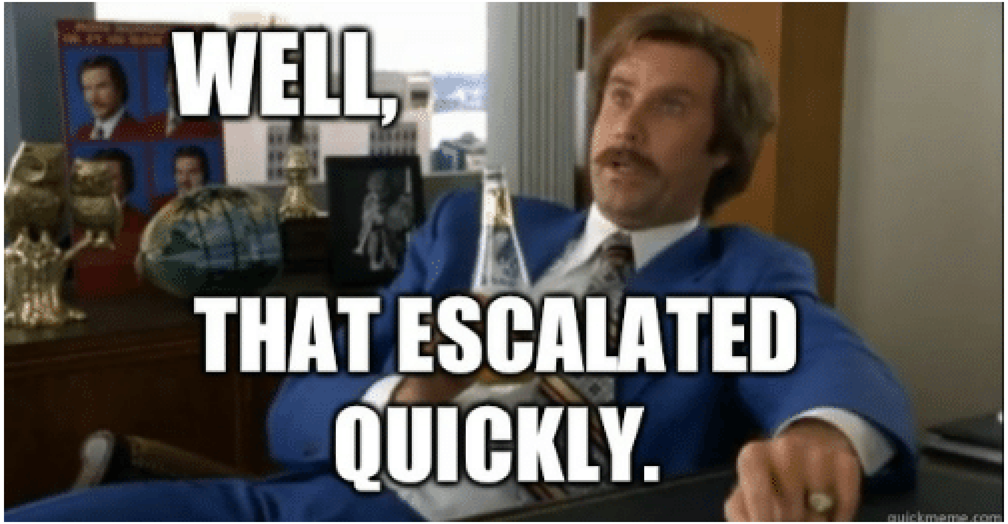


Exhibit 3 Examples

$$X_{7,3} = \begin{bmatrix} 1 & + & + & + \\ 1 & + & + & - \\ 1 & + & - & + \\ 1 & + & - & - \\ 1 & - & + & + \\ 1 & - & + & - \\ 1 & - & - & + \end{bmatrix} \text{->} \; \boldsymbol{M}^* = \begin{bmatrix} 7 & 1 & 1 & 1 \\ 1 & 7 & -1 & -1 \\ 1 & -1 & 7 & -1 \\ 1 & -1 & -1 & 7 \end{bmatrix}$$

VIF=1.094
VRF=0.167 (6 runs)
VRF=0.286 (5 runs)
VRF=0.375 (4 runs)

$$X_{7,6} = \begin{bmatrix} 1 & + & + & + & + & + & + \\ 1 & - & + & - & - & + & + \\ 1 & + & - & - & - & + & + \\ 1 & + & + & - & + & - & - \\ 1 & + & + & + & - & - & - \\ 1 & - & - & + & + & - & + \\ 1 & - & - & + & + & + & - \end{bmatrix} \text{->} \; \boldsymbol{M}^* = \begin{bmatrix} 7 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 7 & 3 & -1 & -1 & -1 & -1 \\ 1 & 3 & 7 & -1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 7 & 3 & -1 & -1 \\ 1 & -1 & -1 & 3 & 7 & -1 & -1 \\ 1 & -1 & -1 & -1 & -1 & 7 & 3 \\ 1 & -1 & -1 & -1 & -1 & 3 & 7 \end{bmatrix}$$

Intercept VIF=1. 167
Factor VIF=1.296

Exhibit 3 Examples

$$X_{7,5}^{D} = \begin{bmatrix} 1 & + & + & + & - & - \\ 1 & + & + & - & - & + \\ 1 & + & - & + & + & - \\ 1 & + & - & - & + & + \\ 1 & - & + & + & + & + \\ 1 & - & + & - & + & - \\ 1 & - & - & + & - & + \end{bmatrix} \text{->} \boldsymbol{M}^{*} = \begin{bmatrix} 7 & 1 & 1 & 1 & 1 & 1 \\ 1 & 7 & -1 & -1 & -1 & -1 \\ 1 & -1 & 7 & -1 & -1 & -1 \\ 1 & -1 & -1 & 7 & -1 & -1 \\ 1 & -1 & -1 & -1 & 7 & -1 \\ 1 & -1 & -1 & -1 & -1 & 7 \end{bmatrix}$$

$$X_{7,5}^{A} = \begin{bmatrix} 1 & + & + & + & + & + \\ 1 & + & - & - & + & + \\ 1 & - & - & - & + & + \\ 1 & + & - & + & - & - \\ 1 & + & + & - & - & - \\ 1 & - & + & + & - & + \\ 1 & - & + & + & + & - \end{bmatrix} \text{->} \boldsymbol{M}^{*} = \begin{bmatrix} 7 & 1 & 1 & 1 & 1 & 1 \\ 1 & 7 & -1 & -1 & -1 & -1 \\ 1 & -1 & 7 & 3 & -1 & -1 \\ 1 & -1 & 3 & 7 & -1 & -1 \\ 1 & -1 & -1 & -1 & 7 & 3 \\ 1 & -1 & -1 & -1 & 3 & 7 \end{bmatrix}$$

So What Does it All Mean?

- While the typical orthogonal designs have many great properties, they’re not the only designs out there.
  - There’s a whole world of optimal designs out there that can fit your resource constraints.
  - As we’ve seen, more runs always reduces variation.
- In limited cases, augmentation of a current design or loss of a run will still yield optimal properties.
  - When they correspond to basic construction methods.
- <u>Don’t tailor your budget to the design; instead, tailor the design to your budget!</u>

Where Can I Start?

- Immediate resources
  - Lots of articles in the literature (search "optimal weighing designs")
  - A table of saturated D-optimal designs: http://www.indiana.edu/~maxdet/fullPage.shtml#tableTop
- Or, if you're willing to be a bit patient…
  - A summary article has been submitted for publication 🤞
  - These techniques will be part of JMP 15 (coming this Fall!)

Thank you!!