Generating optimal space-filling designs with particle swarm optimization.

Rebekah L. Scott
Department of Mathematics and Statistics
Utah State University

Dr. Stephen J. Walsh
Department of Mathematics and Statistics
Utah State University

Introduction

In 1935, Ronald Fisher published *The Design of Experiments,* establishing classical designs for various types of experiments. With the rise of **computing power** came optimal design, where statisticians are able better **customize designs** according to the needs of the researchers running the experiment. Now, designs can be generated based on various criterions and metrics instead of simply selecting a design from a catalog.

Methods

Partial Swarm Optimization (PSO)—a clustering algorithm created by biologists in 1997—was used to generate optimal **space-filling designs**. Although new to the experimental design world, this algorithm is much better at finding optimal designs than the popular Coordinate Exchange algorithm. Space filling designs are able to fit high-ordered models because they seek to best "fill" the design space with experimental runs. In this research, the **MaxMin** objective was used, which seeks to maximize the minim distance between experimental points in the design (Wu). Designs were generated in Julia using PSO for the hypercube and simplex geometries for two and three factors (k = 2, 3) and various numbers of design points (n = 10, 20, 30). The distance metrics used include the Manhattan, Euclidean, Chebyshev, and Aitchison geometries.

Space-Filling Designs: Definitions

Maximin SF-design

$$\arg\max_{\mathbf{X}\in D} f(\mathbf{X}) \quad \text{where} \quad f(\mathbf{X}) = \min_{x_1, \boldsymbol{x_2} \in D} \rho(x_1, \boldsymbol{x_2})$$

Manhattan Distance, L^1

$$\rho\,(x_1, x_2) = \|\, x_1 - x_2 \,\|_{\mathrm{T}} = \sum_{i=1}^{k} |x_{1_i} - x_{2_i}|$$

Euclidean Distance, L^2

$$\rho\,(x_1, x_2) = \left(\sum_{i=1}^{k} |x_{1_i} - x_{2_i}|^2\right)^{1/2}$$

Chebyshev Distance, L^∞

$$\rho\,(x_1, x_2) = max\left(|x_{1_1} - x_{2_1}|, \ldots, |x_{1_k} - x_{2_k}|\right)$$

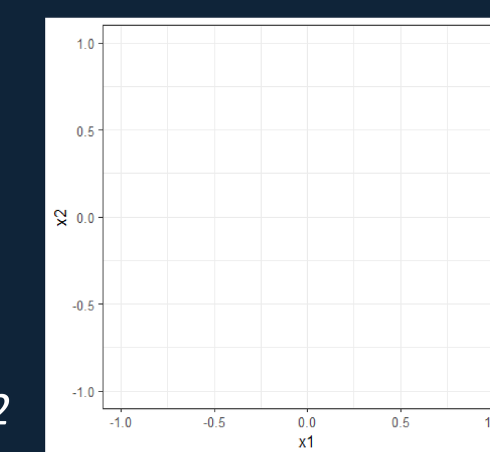
Aitchison Distance

$$\rho\,(x_1, x_2) = \sqrt{\frac{1}{2k}\sum_{i=1}^{k} \sum_{j=1}^{k} \left(\ln\frac{x_{1_i}}{x_{1_j}} - \ln\frac{x_{2_i}}{x_{2_j}}\right)^2}$$

Design Spaces

Process Variables: Hypercube

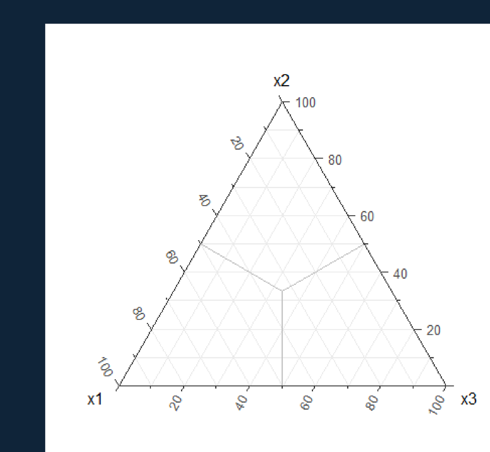
$$\chi = [-1, 1]^K$$



Empty hypercube, k = 2

Mixture Experiments: Simplex

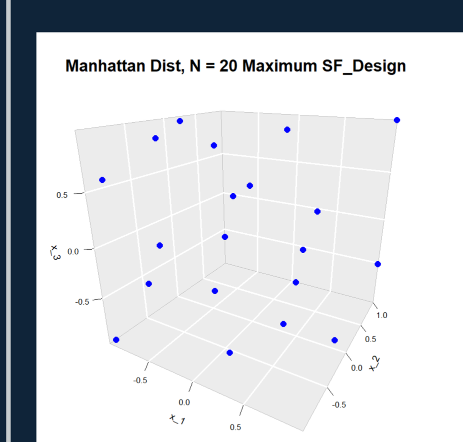
$$\Delta^{K-1} = \{\mathbf{x}' \in \mathbb{R}^K \mid \mathbf{x}'\mathbf{1}_K = 1, 0 \le \mathbf{x}'\mathbf{e}_i \le 1 \; for\; i = 1, \ldots, K\}$$



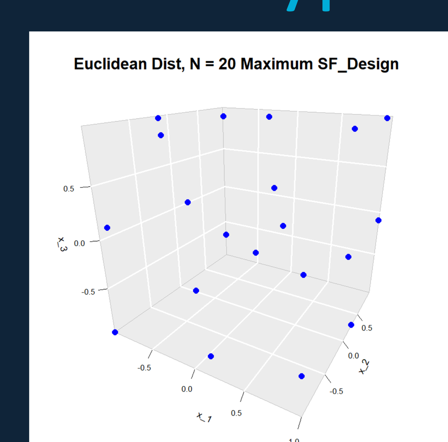
Empty simplex, k = 3

Results

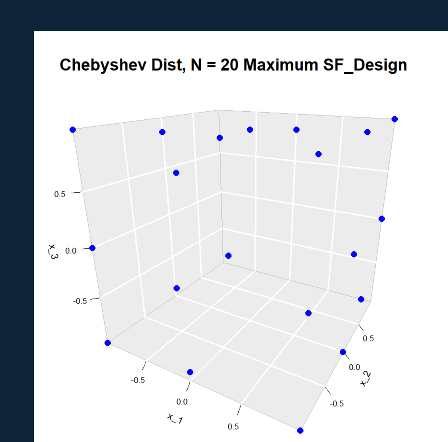
Process Variables: Hypercube



Maximum SF-design with Manhattan Distance on Hypercube

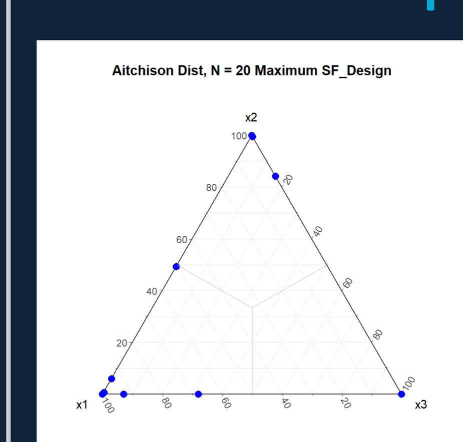


Maximum SF-design with Euclidean Distance on Hypercube

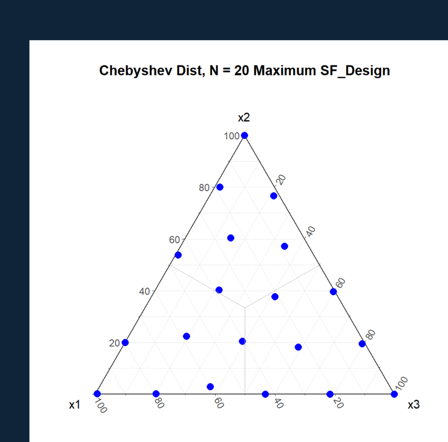


Maximum SF-design with Chebyshev Distance on Hypercube

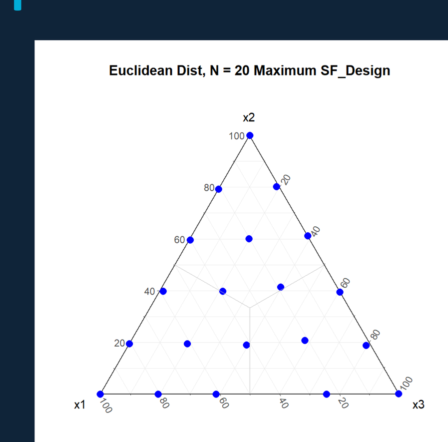
Mixture Experiments: Simplex



Maximum SF-design with Atchison Distance on Simplex



Maximum SF-design with Chebyshev Distance on Simplex

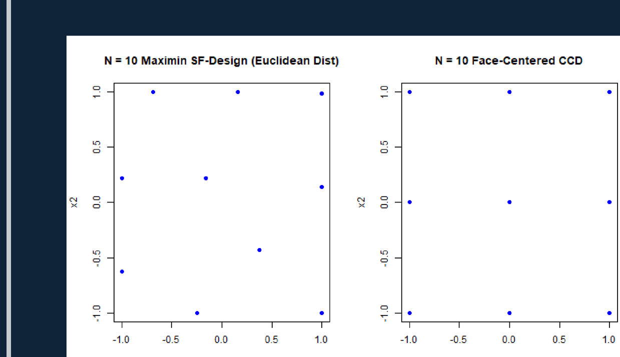


Maximum SF-design with Euclidean Distance on Simplex

Comparison to Classical Design:

Choosing the Experiment to Implement in Practice

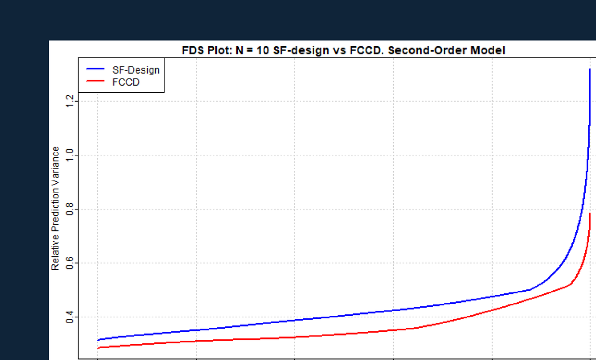
N = 10 SF-design vs. FCCD



Model Support: Data Analysis Options for Competing Designs

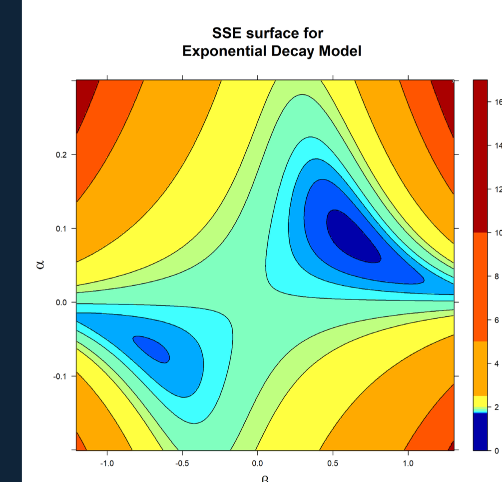
N = 10	Model Support?	
Model	**SF-Design**	**Face Centered CCD**
main-effects	yes	yes
main-effects + interactions	yes	yes
second-order	yes	yes
third-order	yes	no
fourth-order	yes	no

Fraction of Design Space plot comparing the SF-design (*N* = 10) prediction variance profile over study space indicates marginal tradeoff relative to classical design (face-centered CCD) for supporting Second order model.



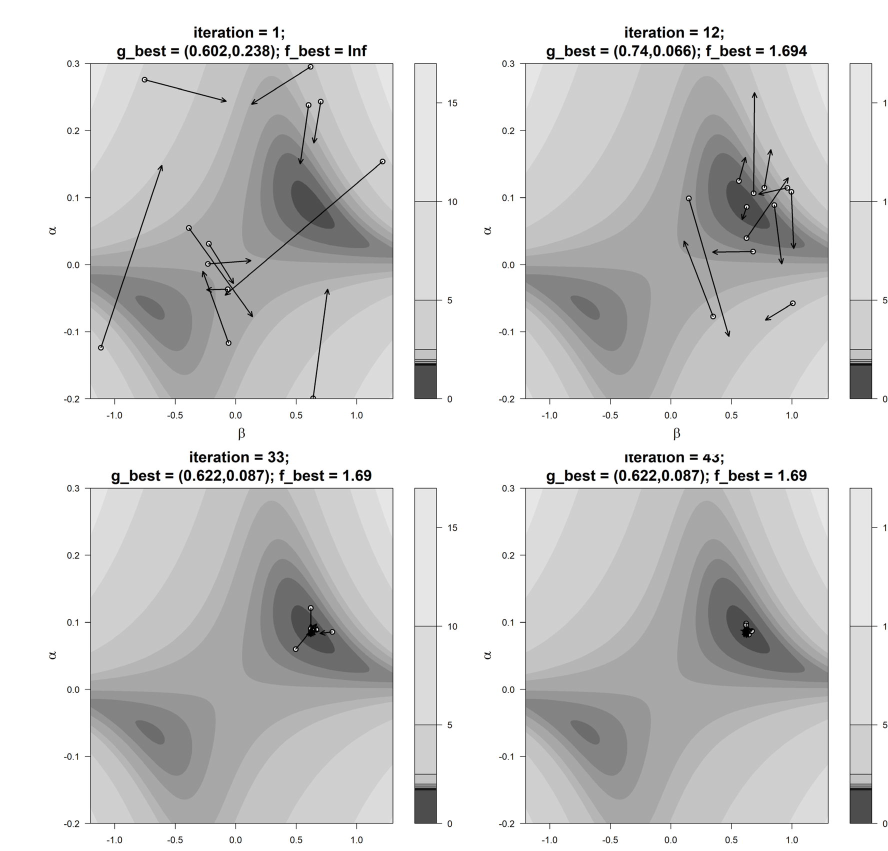
Particle Search Optm.

This optimization algorithm is extremely popular and used in applications such as neural networks, and signal processing. This meta-heuristic algorithm makes no assumptions about the nature of the objective function and does not use gradient techniques. It excels at finding global optimal on non-convex functions, making it an excellent choice for finding optimal designs.



PSO running on a simple 2-D objective function (below).

PSO converges quickly to the global optimal solution (left)



Conclusions

Interestingly, changing the distance metric in the objective function had a very small effect on the design, except for when using Aitchison geometry on the simplex, where points tended to cluster to the vertices. Investigating this strange behavior is an area of further research. There is only a small sacrifice in prediction variance with a space-filling design (compared with CCD with N = 10). Space-filling designs are the way to go for researchers wanting to support a higher-order model!

UtahStateUniversity