Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

A RECOMMENDATION SYSTEM FOR FIRST-ORDER NEARLY ORTHOGONAL-AND-BALANCED (NOAB) DESIGNS

Zachary C. Little

Department of Operational Sciences
Air Force Institute of Technology
2950 Hobson Way
Wright-Patterson AFB, OH 45433, USA

ABSTRACT

The construction of nearly orthogonal-and-balanced (NOAB) designs is examined for full first-order models in the framework of an algorithm selection problem, allowing for the examination of experimental design performance measures for various design sizes and maximum allowed imbalance settings. Based on a randomly-generated set of large design spaces, performances measures of D-criterion for good parameter estimation as well as estimated maximum unscaled prediction variance (UPV) are largely driven by choice of design size, with specific design space features found to impact the measures. In this multi-objective setting, prediction of design performance within the framework consistently results in designs that perform well over an entire sampled weight space for the multiple performance measures as well as for specific weights.

1 INTRODUCTION

Large design spaces of interest for complex, black-box systems often cannot be exhaustively explored, requiring space-filling experimental designs with possibly mixed factors (i.e., quantitative and qualitative with different numbers of levels) that allow for the construction of meta-models to represent system responses. A construction method is developed for first-order NOAB mixed-factor designs in (Vieira Jr. et al., 2013), though beyond a suggested range for design sizes there exists a need for greater knowledge of design performance for different design sizes and other construction parameter settings. The framework of an algorithm selection problem can aid in such understanding, even if a single selection rule is found to consistently result in top performing first-order NOAB designs for various problems.

2 ALGORITHM SELECTION PROBLEM FRAMEWORK

NOAB designs are constructed sequentially by factor using mixed-integer linear programming (MILP) formulations. Near orthogonality means that the maximum absolute pairwise correlation, ρ_{map}, for an encoded design matrix is less than or equal to 0.05, allowing for separation of individual factors. Near balance means that the maximum imbalance for all factor columns is close to zero, i.e. all levels of interest are represented nearly equally.

The algorithm selection problem requires four components/spaces. The *problem space* consists of 30 randomly-generated design spaces with between 8 and 20 factors overall, with categorical factors between 3 and 7 levels and discrete factors between 2 and 12 levels. The *feature space* includes the number of factors for each factor type (discrete and categorical) as well as statistical measures of the number of levels for each factor type, to include minimum, mean, maximum, Q1, median, Q3, sum, standard deviation, skewness, and kurtosis. The product of all numbers of levels (full factorial design size) and least common multiple of all numbers of levels are also included as meta-features. The *algorithm space* is comprised of combinations of $m = 2, 3, \dots, 10$ and maximum allowed imbalance $\delta^* = 0.05, 0.10, 0.15, 0.20, 0.25$. The smallest balance-feasible design size $n \geq m \cdot s$ is attempted for the number of design matrix columns s, with each MILP allowed up to two attempts of 30 seconds each to satisfy $\rho_{map} \leq 0.05$. The *performance space* is multi-objective where the aim is to minimize design size n and estimated maximum unscaled prediction variance (UPV), while maximizing the parameter estimation measure D-criterion. Predictions of the individual performance measures are used to develop an accurate recommendation system. The three measures are given the same scale by using linear desirability functions. An overall multiplicative desirability is used to ensure that no single measure performs too poorly, with weights given to each individual desirability. Synthesized efficiency (SEff) for a design of interest measures the overall desirability of that design relative to that of the top performing design for a specific weighting combination (sampled using a 5,000-point space-filling mixture design). The actual overall desirability for up to the top-5 predicted designs is compared with the desirability of the top performing design at each sampled weighting combination as well as for average SEff and minimum SEff. Spearman's rank correlation coefficient is also used to compare the actual and predicted rankings of overall desirability, average SEff, and minimum SEff.

A model-based meta-learner approach examines a set of possible meta-models for each performance measure of interest, where the model providing the smallest root mean square error (RMSE) is selected using 10-fold cross-validation. The meta-models considered include artificial neural networks (ANN), classification and regression trees (CART), multivariate adaptive regression splines (MARS), Gaussian processes with linear, polynomial, and radial kernels, random forests (RF), and support vector machines (SVM) with linear, polynomial and radial kernels. All observations for the problem to be predicted are held out from the training data set.

3 PREDICTION OF DESIGN PERFORMANCE MEASURES

Design size n is predetermined by the choice of m and δ^* using the balance-feasibility test from (Vieira Jr. et al., 2013). Using SVM with polynomial kernel and RF consistently results in the smallest RMSEs when predicting D-criterion and maximum UPV, respectively. The accurate prediction of individual design performance measures results in accurate recommendation and ranking of first-order NOAB designs for average and minimum SEff as well as multiplicative desirability over the weight space. Common selections for m across all problems are 6 and 7 for high average SEff (often near 0.89) and 6 for high minimum SEff (often near 0.5). Increasing δ^* generally relaxes balance constraints to achieve smaller n. The time required to construct designs for all (m, δ^*) combinations for a single problem is approximately between 2 and 14 hours. For the recommendation system, building two meta-models requires roughly 30 seconds and constructing a single design needs only between 3 and 19 minutes, so the developed framework and resulting recommendation system allows for efficient selection and construction of first-order NOAB designs.

4 FUTURE RESEARCH AND ADDITIONAL NOTES

Little et al. have derived extensions to the original construction method to allow for the creation of second-order NOAB designs (i.e., near orthogonality includes two-way interactions and quadratics), which may be examined in a similar framework. An analysis of computational requirements is warranted for various design spaces, whether by changing the allowed run time or implementing other stopping criteria. This research is in support of the Simulation and Analysis Facility (SIMAF), Air Force Life Cycle Management Center, Simulation and Analysis Division (AFLCMC/XZS). The views expressed in this abstract are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

REFERENCES

Vieira Jr., H., Sanchez, S.M., Kienitz, K.H. & Belderrain, M.C.N. (2013) Efficient, nearly orthogonal-and-balanced, mixed designs: an effective way to conduct trade-off analyses via simulation. *Journal of Simulation*, 7 (4): 264–275.