Symmetric Order of Addition Experiments

Nicholas Rios, George Mason University

Overview

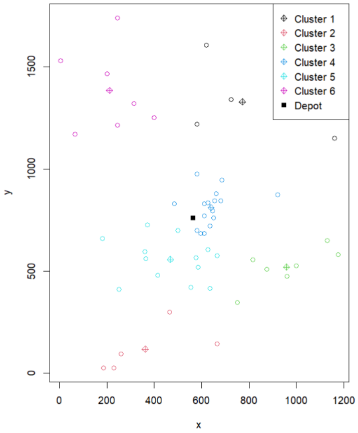
Motivating Example

- "Typical" delivery problem
- One truck makes many stops
- What is the shortest route?
  (Travelling Salesman Problem)

Motivating Example

New technology - delivery drones.

Motivating Example



- New approach: separate delivery points into m clusters
- Have trucks stop at the center of each cluster
- Send drones to all delivery points in the cluster, wait for them to return.

Motivating Example

- **Truck and Drone Delivery** problem
- One truck makes m stops (one per cluster)
- At each stop, they deploy drones to each delivery point in the cluster.
- What is the optimal route for the truck?

The Plan (High Level)

1. Identify a design, i.e., a subset of all possible routes, and retrieve the cost for each route in the design.

2. Fit a model to the data.

3. Use the model's parameter estimates to (greedily) search for the route with the lowest cost.

Challenges

- If m is large, then $m!$ is too large to examine all possible orders.
- e.g. if $m = 10$ then $10! = 3628800$. We can use symmetry to cut this in half, but it is still too large to examine.
- Existing Order-of-Addition models (and designs) inherently assume asymmetry, but this problem has underlying symmetric cost.

Overview

Order of Addition (OofA) Experiments

- We have m components and a response y
- Question: if we change the order of the m components, will this affect the response?
- A treatment is a permutation of (1,2, . . . , m).
- There are $m!$ treatments.
- We want to find the order of the m components that gives the optimal (maximum or minimum) response y^*.

Pairwise Ordering (PWO) Model

Table: OofA Design for $m = 3$, all possible orders

Order	z_{12}	z_{13}	z_{23}
(1,2,3)	1	1	1
(1,3,2)	1	1	-1
(3,1,2)	1	-1	-1
(3,2,1)	-1	-1	-1
(2,3,1)	-1	-1	1
(2,1,3)	-1	1	1

- Van Nostrand (1995) introduced a model based on pairwise order (PWO).
- PWO Model: $y = \beta_0 + \sum_{jk} z_{jk}\beta_{jk} + \epsilon$
- Some combinations of PWO factors are impossible (e.g. $z_{12} = 1, z_{13} = -1, z_{23} = 1$).

Component Poisition (CP) Model

- An alternative to the PWO coding scheme is the Component-Position (CP) model (Yang et al., 2021).

$$y(\mathbf{a}) = \mu_0 + \sum_{c=1}^{m} \sum_{j=1}^{m} \delta_c^{(j)} x_c^{(j)}(\mathbf{a}) + \epsilon$$

  for any order **a** of components.

- $x_c^{(j)}(\mathbf{a}) = 1$ if component c is in position j, and 0 otherwise.

- $\delta_c^{(j)}$ is the effect of placing component c in position j on the expected response

Component Position (CP) Model

Table: CP Design for $m = 3$, all possible orders

Order	$z_1^{(1)}$	$z_1^{(2)}$	$z_1^{(3)}$	$z_2^{(1)}$	$z_2^{(2)}$	$z_2^{(1)}$	$z_3^{(1)}$	$z_3^{(2)}$	$z_3^{(3)}$
(1,2,3)	1	0	0	0	1	0	0	0	1
(1,3,2)	1	0	0	0	0	1	0	1	0
(3,1,2)	0	1	0	0	0	1	1	0	0
(3,2,1)	0	0	1	0	1	0	1	0	0
(2,3,1)	0	0	1	1	0	0	0	1	0
(2,1,3)	0	1	0	1	0	0	0	1	0

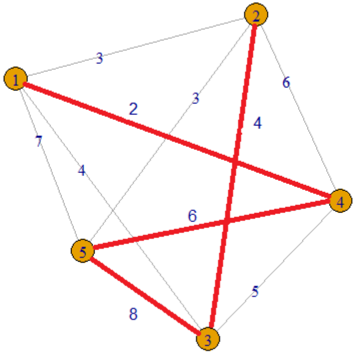
Overview

A Definition of Symmetry

- Let $\mathbf{a} = (a_1, \ldots, a_m)$ be a permutation of $(1, 2, \ldots, m)$.
- Let $\text{rev}(\mathbf{a}) = (a_m, \ldots, a_1)$ be the reversal of $\mathbf{a}$.
- Let $\tau(\mathbf{a})$ be the expected response given $\mathbf{a}$.
- We say that the order of addition problem is **symmetric** if $\tau(\mathbf{a}) = \tau(\text{rev}(\mathbf{a}))$ for all possible permutations $\mathbf{a}$
- Symmetric means reversing the order shouldn't change the expected outcome.

Graphical Representation - Symmetric Case

- Symmetry means we can use an undirected graph to represent the problem.
- Any permutation is a Hamiltonian path.



Problem Statement

- Let $\mathcal{A}^*$ be the set of all of $\frac{m!}{2}$ permutations that are distinct under reversals.

- Let $\mathcal{D}$ be the set of all possible subsets of $\mathcal{A}^*$. We want an optimal design D^*, i.e.

$$D^* = \arg\max_{D \in \mathcal{D}} \phi(D)$$

  for some (model-based) optimality criterion ϕ. (e.g. $D-$optimal)

- We wish to identify an optimal Hamiltonian path $\mathbf{a} \in \mathcal{A}^*$, i.e.

$$\mathbf{a}^* = \arg\min_{\mathbf{a} \in \mathcal{A}^*} \tau(\mathbf{a})$$

  where $\tau(\mathbf{a})$ is the expected response for a permutation $\mathbf{a}$.

Overview

Symmetric OofA Model

Table: Symmetric OofA Coding for $m = 3$

Order	x_{12}	x_{13}	x_{23}
(1,2,3)	1	0	1
(1,3,2)	0	1	1
(3,1,2)	1	1	0

- Let

$$x_{jk}(\mathbf{a}) = \begin{cases} 1 & \text{if } (j,k) \in \mathbf{a} \\ 0 & \text{otherwise} \end{cases}$$

- Model:
  $y = \beta_0 + \sum_{jk} x_{jk}(\mathbf{a})\beta_{jk} + \epsilon$
- For identifiability purposes, we constrain $\beta_{12} = \beta_{23} = \cdots = \beta_{(m-1)m} = 0$.
- Moment Matrix $M = \frac{1}{n}X^T X$, where X is the model matrix expansion of the design

Full Design is Optimal

Corollary

Under the proposed model, the full design (with all $m!/2$ runs included exactly once) is $\phi-$optimal for any ϕ that is concave and permutation invariant.

- Follows from Lin and Peng (2019), which proves a similar result for the asymmetric case.
- Requires permutation invariance (instead of signed permutation invariance).
- This allows us to compare all designs to the full design.

Moment Matrix of Full Design

Lemma 1: Moment Matrix

Under the proposed model, the full design has moment matrix

$$M_f = \frac{2}{m!} \begin{bmatrix} m!/2 & (m-1)!\mathbf{1}^T \\ (m-1)!\mathbf{1} & (m-1)!I_{\binom{m}{2}} + (m-2)!Q \end{bmatrix}$$

where $I_{\binom{m}{2}}$ is an identity matrix of dimension $\binom{m}{2}$, $\mathbf{1}$ is a $\binom{m}{2} \times 1$ column of ones, and Q is a matrix with columns and rows indexed by the pairs $(12, 13, 14, \ldots, (m-1)m)$ in lexicographically increasing order and elements

$$Q(ij, k\ell) = \begin{cases} 0 & \text{if } (i,j) = (k,\ell) \\ 2 & \text{if } i \neq k,\ell \text{ and } j \neq k,\ell \\ 1 & \text{otherwise.} \end{cases}$$

D-efficiency of Full Design

- From Lemma 1, we can show that the D-efficiency of the full design is

$$\left[\left(1+\frac{m-2}{m}\right)\left(\frac{2}{m}\right)^{\frac{(m-1)(m-2)}{2}-1}\left(\frac{2}{m(m-1)}\right)^{m-1}\right]^{\frac{1}{p}}.$$

- This allows us to quickly find the D-efficiency of the full design for any m, which is useful when relative efficiencies need to be calculated.

- Can also derive formula for $A-$efficiency.

Recursive Algorithm to Generate $\phi-$Optimal Designs

Algorithm 1: Recursively Generate Optimal Fractional Design

Inputs: Optimal fractional design for $m-1$ components D_{m-1}.
for $b = 1, 2, \ldots, m$ **do**
  Let B_b be the matrix that results from inserting a column of m's after the $(m-b)^{th}$ column of D_{m-1}.
end
$D = [B_1^T, B_2^T, \ldots, B_m^T]^T$
return D

Fractional Designs are $\phi-$Optimal

Theorem 1

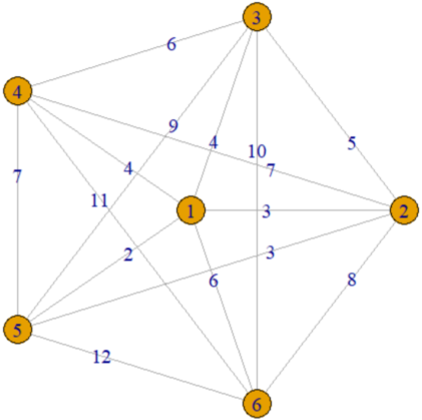
Suppose D is a design matrix generated by Algorithm 1. Then, D is $\phi-$optimal for any ϕ that is concave and permutation invariant.

- Applies to $D-$, $A-$ and many other popular optimality criteria.
- It is easy to use computer search to find optimal designs for small m (e.g. $m = 4, 5$ and then use recursion to find optimal designs for larger m.

Algorithm - Find Optimal Paths

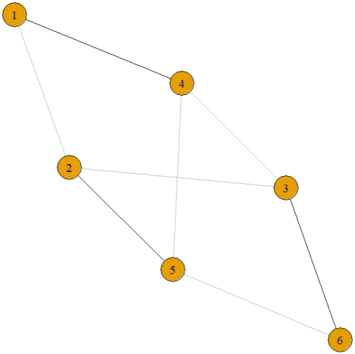
1. Let G^* be an empty graph on m vertices.
2. For all significant $\hat{\beta}_{jk} < 0$, add (j, k) to G^*.
3. If G^* contains no Hamiltonian paths, add baseline edges $(1, 2), (2, 3), \ldots, (m-1, m)$ to G^*.
4. Return the set of all Hamiltonian paths in G^*.

Toy Example - Finding Optimal Paths



- Example with $m = 6$ cities
- Minimal cost order (including return cost) is $\mathbf{a} = (5, 2, 1, 4, 3, 6)$.
- Significant negative edges are $(2, 5), (1, 4), (3, 6)$

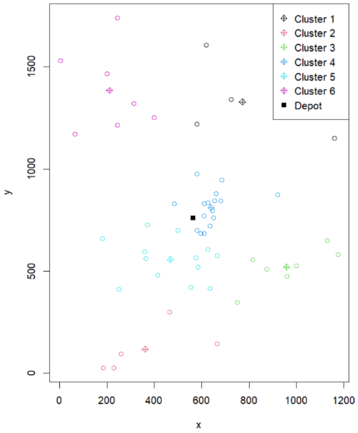
Toy Example - Finding Optimal Paths



- Added baseline edges $(1,2),(2,3),(3,4),(4,5),(5,6)$ to G^*
- Six possible Hamiltonian paths on G^*
- The optimal path is captured.

Overview

Motivating Example



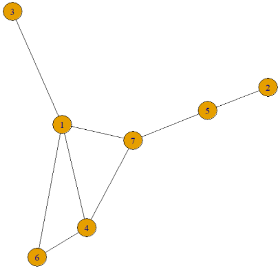
- 52 delivery locations in Berlin in 6 clusters
- $m = 7$ components: Depot (origin point) + 6 stops
- Used recursive algorithm to create optimal design based on a design for $m = 4$ components
- The optimal solution is $(1, 3, 2, 5, 7, 4, 6)$, i.e. start at depot (7), go to 4, 6, 1, 3, 2, 5, return.

Motivating Example

Parameter	Estimate	Std. Error	T	P-value
β_0	8959.74	24.48	365.97	0.00
β_{13}	-37.77	16.01	-2.36	0.02
β_{14}	-168.79	15.06	-11.21	0.00
β_{15}	51.64	15.47	3.34	0.00
β_{16}	-659.77	15.30	-43.12	0.00
β_{17}	-81.21	16.19	-5.02	0.00
β_{24}	46.25	15.07	3.07	0.00
β_{25}	-325.23	14.40	-22.59	0.00
β_{26}	77.39	14.70	5.26	0.00
β_{27}	0.14	15.30	0.01	0.99
β_{35}	3.61	15.09	0.24	0.81
β_{36}	217.72	14.40	15.12	0.00
β_{37}	78.53	15.47	5.08	0.00
β_{46}	-37.06	15.07	-2.46	0.01
β_{47}	-133.01	15.06	-8.83	0.00
β_{57}	-77.54	16.01	-4.84	0.00

Table: Parameter Estimates

Motivating Example



- Two possible optimal solutions: (2, 5, 7, 4, 6, 1, 3) and (3, 1, 6, 4, 7, 5, 2)
- Both are rotations of the optimal path

Overview

Conclusion

- First attempt to create designs and models for the Order-of-Addition problem where the costs are symmetric
- New model for representing the effect of order via edges in the adjacency matrix of an undirected graph
- Systematic recursive method for finding $\phi-$optimal fractions of the full design

Future Work

- Interactions between edges in a graph
- Cheaper designs that are highly efficient, but not exactly optimal might be found using stochastic search algorithms, e.g. Threshold Accepting (Winker et al., 2020)
- What if G is not a complete graph?

References I

Agatz, N., P. Bouman, and M. Schmidt (2018). Optimization approaches for the traveling salesman problem with drone. *Transportation Science 52*(4), 965–981.

Aidoo, R. P., E. O. Afoakwa, and K. Dewettinck (2014). Optimization of inulin and polydextrose mixtures as sucrose replacers during sugar-free chocolate manufacture–rheological, microstructure and physical quality characteristics. *Journal of Food Engineering 126*, 35–42.

Chandrasekaran, S. M., S. Bhartiya, and P. P. Wangikar (2006). Substrate specificity of lipases in alkoxycarbonylation reaction: Qsar model development and experimental validation. *Biotechnology and Bioengineering 94*(3), 554–564.

Chang, Y. S. and H. J. Lee (2018). Optimal delivery routing with wider drone-delivery areas along a shorter truck-route. *Expert Systems with Applications 104*, 307–317.

Chen, J., R. Mukerjee, and D. K. J. Lin (2020a). Construction of optimal fractional order-of-addition designs via block designs. *Statistics & Probability Letters 161*, 108728.

Chen, J. B., R. Mukerjee, and D. K. J. Lin (2020b). Construction of optimal fractional order-of-addition designs via block designs. *Statistics & Probability Letters*.

Ding, X., K. Matsuo, L. Xu, J. Yang, and L. Zheng (2015). Optimized combinations of bortezomib, camptothecin, and doxorubicin show increased efficacy and reduced toxicity in treating oral cancer. *Anti-Cancer Drugs 26*(5), 547–554.

Freitas, J. C., P. H. V. Penna, and T. A. Toffolo (2022). Exact and heuristic approaches to truck-drone delivery problems. *EURO Journal on Transportation and Logistics*, 100094.

García-Ródenas, R., J. C. García-García, J. López-Fidalgo, J. Á. Martín-Baos, and W. K. Wong (2020). A comparison of general-purpose optimization algorithms for finding optimal approximate experimental designs. *Computational Statistics & Data Analysis 144*, 106844.

Lin, D. K. J. and J. Peng (2019). Order-of-addition experiments: A review and some new thoughts (with discussion). *Quality Engineering 31*(1), 49–59.

Mee, R. W. (2020). Order-of-addition modeling. *Statistica Sinica 30*(3), 1543–1559.

Peng, J., R. Mukerjee, and D. K. J. Lin (2019). Design of order-of-addition experiments. *Biometrika 106*(3), 683–694.

Rajaonarivony, M., C. Vauthier, G. Couarraze, F. Puisieux, and P. Couvreur (1993). Development of a new drug carrier made from alginate. *Journal of pharmaceutical sciences 82*(9), 912–917.

References II

Rios, N. and D. K. J. Lin (2021). Order-of-addition mixture experiments. *Journal of Quality Technology*. Submitted for Publication.

Rios, N., P. Winker, and D. K. Lin (2022). Ta algorithms for d-optimal oofa mixture designs. *Computational Statistics & Data Analysis 168*, 107411.

Sljivic-Ivanovic, M. Z., I. D. Smiciklas, S. D. Dimovic, M. D. Jovic, and B. P. Dojcinovic (2015). Study of simultaneous radionuclide sorption by mixture design methodology. *Industrial & Engineering Chemistry Research 54*(44), 11212–11221.

Van Nostrand, R. (1995). Design of experiments where the order of addition is important. In *ASA proceedings of the Section on Physical and Engineering Sciences*, pp. 155–160. American Statistical Association Alexandria, VA.

Voelkel, J. G. (2019). The design of order-of-addition experiments. *Journal of Quality Technology 51*(3), 230–241.

Voelkel, J. G. and K. P. Gallagher (2019). The design and analysis of order-of-addition experiments: An introduction and case study. *Quality Engineering 31*(4), 627–638.

Wang, A., H. Xu, and X. Ding (2020). Simultaneous optimization of drug combination dose-ratio sequence with innovative design and active learning. *Advanced Therapeutics 3*(4), 1900135.

Winker, P., J. Chen, and D. K. J. Lin (2020). The construction of optimal design for order-of-addition experiment via threshold accepting. In *Contemporary Experimental Design, Multivariate Analysis and Data Mining*, Chapter 6, pp. 93–109. Springer.

Yang, J.-F., F. Sun, and H. Xu (2021). A component-position model, analysis and design for order-of-addition experiments. *Technometrics 63*(2), 212–224.

Zhao, Y., D. K. Lin, and M.-Q. Liu (2022). Optimal designs for order-of-addition experiments. *Computational Statistics & Data Analysis 165*, 107320.

Zhao, Y. L., D. K. J. Lin, and M. Liu (2020). Design for order of addition experiment. *Journal of Applied Statistics forthcoming.*