A Mixed Integer Optimization Approach for Model Selection in Screening Experiments

Alan Vazquez, Eric Schoen, Peter Goos
alan.vazquezalcocer@kuleuven.be

ENBIS Meeting 2019

Budapest, Hungary

KU LEUVEN

Outline

1. Motivating Example

2. Mixed Integer Optimization (MIO) Approach
    1. Basic idea
    2. Constraints on the model search
    3. Create a list of best models

3. Results and Discussion

Vazquez, Schoen and Goos (2018). A mixed integer optimization approach for model selection in screening experiments. Technical report 2018007. University of Antwerp. Available upon request.

Decolorization experiment

- Investigate the electrochemical decolorization of the RV5 compound on a textile surface (Fidaleo et al., 2016).
- 9 continuous factors at three levels.
- A 21-run definitive screening design was used.
- Response is the % of color removal.

Goal: Detect the influential linear effects, two-factor interactions and quadratic effects.

Data analysis via model selection

Run	A	B	C	D	E	F	G	H	I	Y
1	0	1	1	1	1	1	1	1	1	3.26
2	0	-1	-1	-1	-1	-1	-1	-1	-1	-3.29
3	1	0	-1	-1	-1	-1	1	1	1	4.70
4	-1	0	1	1	1	1	-1	-1	-1	-2.53
5	1	-1	0	-1	1	1	-1	-1	1	-2.68
6	-1	1	0	1	-1	-1	1	1	-1	3.44
7	1	-1	-1	0	1	1	1	1	-1	4.25
8	-1	1	1	0	-1	-1	-1	-1	1	-1.80
9	1	-1	1	1	0	-1	-1	1	-1	-0.37
10	-1	1	-1	-1	0	1	1	-1	1	1.11
11	1	-1	1	1	-1	0	1	-1	1	0.42
12	-1	1	-1	-1	1	0	-1	1	-1	0.40
13	1	1	-1	1	-1	1	0	-1	-1	2.72
14	-1	-1	1	-1	1	-1	0	1	1	1.67
15	1	1	-1	1	1	-1	-1	0	1	1.14
16	-1	-1	1	-1	-1	1	1	0	-1	-0.14
17	1	1	1	-1	-1	1	-1	1	0	0.00
18	-1	-1	-1	1	1	-1	1	-1	0	0.49
19	1	1	1	-1	1	-1	1	-1	-1	-0.69
20	-1	-1	-1	1	-1	1	-1	1	1	-0.70
21	0	0	0	0	0	0	0	0	0	-0.38

Y: *logit* of the percentage of color removal.

Information:

- 54 potential effects, 21 observations.
- Effect heredity.
- Screening design.

Desirable properties of a model selection method for screening

Property 1: Creates a list of models that are compatible with the data.

Property 2: Reveals the aliasing present in the screening design.

Property 3: Provides a framework to specify restrictions on the model search.

Available Model Selection Methods

Method	P1: List	P2: Aliasing	P3: Restrictions
LASSO and Extensions (Yuan et al., 2006; Choi et al., 2010; Bien et al., 2016)	✓	✗	✗
Dantzig Selector (Phoa et al., 2009)	✓	✗	✗
LARS and Extensions (Yuan et al., 2007, 2009)	✓	✗	✓
Forward Selection (Westfall et al., 1998)	✓	✗	✗
Simulated Annealing Model Search (Wolters et al., 2011)	✓	✓	✗

Outline

1. Motivating Example

2. Mixed Integer Optimization (MIO) Approach
    1. Basic idea
    2. Constraints on the model search
    3. Create a list of best models

3. Results and Discussion

MIO approach

<u>Basic Idea:</u> Best-subset selection

$$\min_{\widehat{\boldsymbol{\beta}}} \sum_{i=1}^{n} \left(y_i - \sum_{u=1}^{p} x_{iu} \hat{\beta}_u \right)^2$$

Subject to:

$$\sum_{u=1}^{p} I\left(\hat{\beta}_u \neq 0\right) = k$$

Indicator function

$\widehat{\boldsymbol{\beta}} = \left(\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p\right)^T$: Parameter estimates.

k: Subset size.

MIO problem

Bertsimas et al. (2016) formulated the best-subset selection problem as a mixed integer optimization problem.

Parameter estimates

$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_p \end{pmatrix}$$

Binary decision variables $z_u \in \{0,1\}$

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{pmatrix} = \mathbf{z}$$

Link between $\hat{\beta}_u$ and z_u:

$$\hat{\beta}_u(1 - z_u) = 0.$$

If $z_u = 0$, then $\hat{\beta}_u = 0$.

If $z_u = 1$, $\hat{\beta}_u$ can be different from zero.

Standard MIO problem

$$\min_{\hat{\boldsymbol{\beta}},\mathbf{z}} \sum_{i=1}^{n} \left(y_i - \sum_{u=1}^{p} x_{iu} \hat{\beta}_u \right)^2$$

Subject to:

$$\hat{\beta}_u (1 - z_u) = 0$$

$$\sum_{u=1}^{p} z_u = k$$

$$z_u \in \{0,1\}$$

$$\hat{\beta}_u \in \mathbb{R}$$

Framework to specify constraints on the model search.

Solvers: GUROBI, CPLEX, BARON, SCIP.

Outline

1. Motivating Example

2. Mixed Integer Optimization (MIO) Approach
    1. Basic idea
    2. Constraints on the model search
    3. Create a list of best models

3. Results and Discussion

Constraints on the model search

Effect heredity

z_u: linear effect of factor u.

z_{uv}: two-factor interaction between u and v.

- Strong effect heredity:

$$z_{uv} \leq z_u \text{ and } z_{uv} \leq z_v$$

- Weak effect heredity:

$$z_{uv} \leq z_u + z_v$$

- Effect heredity for quadratic effects:

$$z_{uu} \leq z_u$$

Final MIO problem

$$\min_{\hat{\boldsymbol{\beta}},\mathbf{z}} \sum_{i=1}^{n} \left(y_i - \sum_{u=1}^{p} x_{iu} \hat{\beta}_u \right)^2$$

Subject to:

$$\hat{\beta}_u (1 - z_u) = 0$$

$$\sum_{u=1}^{p} z_u = k$$

$$z_u \in \{0,1\}, \hat{\beta} \in \mathbb{R}$$

$$z_{uu} \leq z_u$$

$$z_{uv} \leq z_u + z_v$$

Categorical factors and heredity,

factor sparsity,

$\vdots$

Standard MIO problem.

- **Property 3: Search restrictions.**
- Output: The optimal feasible model of size k.

User-specified constraints.

Outline

1. Motivating Example

2. Mixed Integer Optimization (MIO) Approach
   1. Basic idea
   2. Constraints on the model search
   3. Create a list of best models

3. Results and Discussion

A sequential algorithm to list the best models

For a given model size k:

1. Solve MIO problem → Find optimal model S_1.
2. Solve MIO problem + extra constraint to exclude S_1 → Find second-best model S_2.
3. Solve … nts to excluc … model S_3.

Extra constraint:

$$\sum_{u \in S_1} z_u \leq k - 1$$

Ext

$\sum_{u}$ … $k - 1$

4. So … ex … S_M.

Output: List of the M best models of size k.

List of models

Repeat the sequential algorithm for several model sizes, $k = 1, 2, \ldots, k_0$.

Construct a list of the best models of different sizes. **Property 1: Compatible models with data.**

Study the list using raster plots. **Property 2: Visualize aliasing.**
Wolters and Bingham (2011)

The effects that appear consistently in the best feasible models are declared active.

Outline

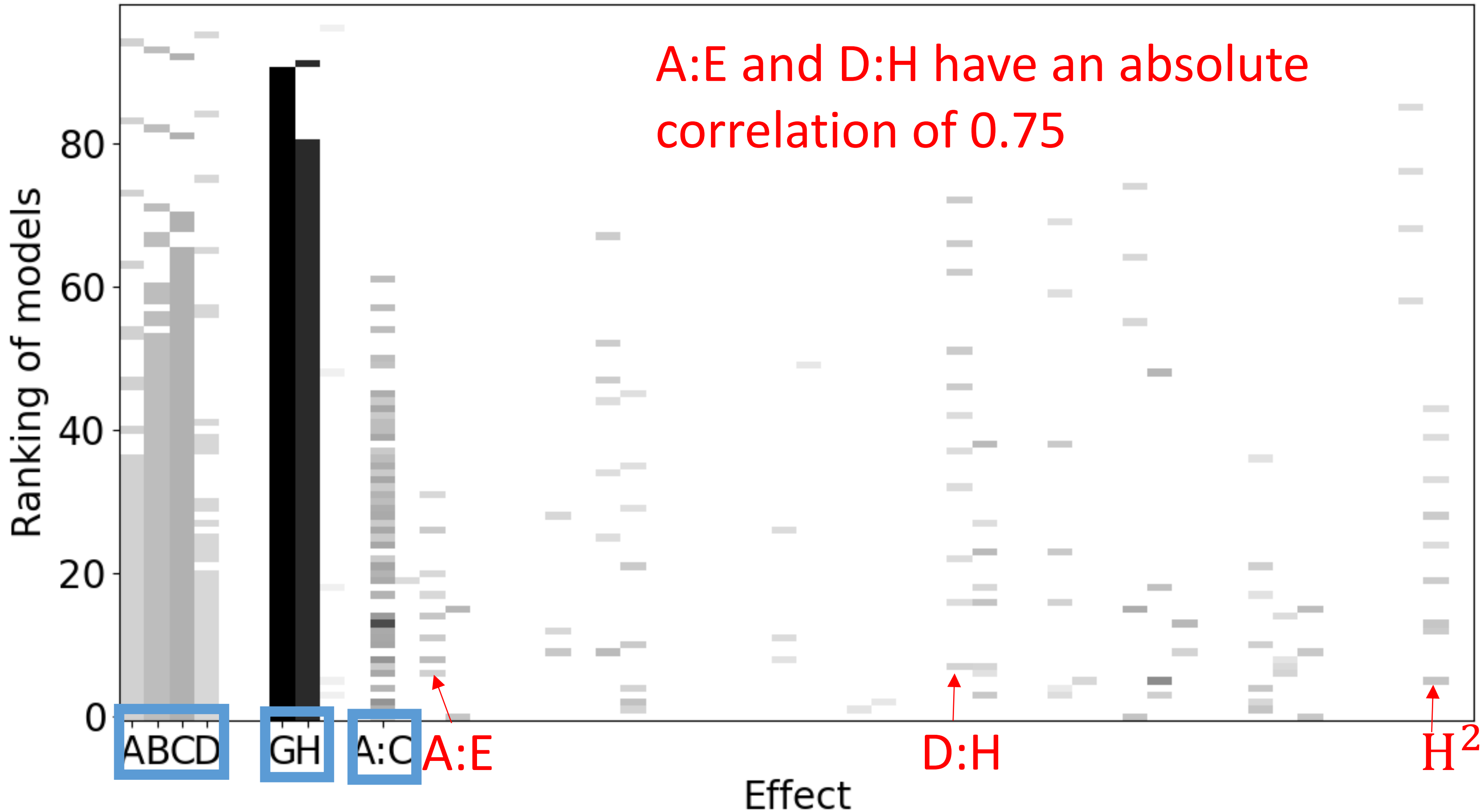
1. Motivating Example

2. Mixed Integer Optimization (MIO) Approach
   1. Basic idea
   2. Constraints on the model search
   3. Create a list of best models

3. Results and Discussion

Decolorization experiment

The 10 best models of sizes 1, 2, ..., 10.



The models follow weak effect heredity.

Discussion

- We developed a MIO approach to analyze data from screening designs.

- Unlike the benchmark methods, MIO has all the desirable properties of a good model selection method for screening.

- The MIO approach is deterministic and guarantees that the best-fitting models will be found.

- A Python/Gurobi implementation is available.

Special thanks to the ENBIS Knowledge Fund and the FWO, Belgium.

Appendix A: Computing times MIO approach

Specifications:

- Standard CPU: Intel(R) Core(TM) i7-4770 CPU @ 3.4 GHz, 16 GB.

- Python and Gurobi v7.5

Instance	(Runs, Effects)	Max. k	Additional Constraints	Time
Decolorization Experiment	(21, 54)	10	Heredity	3 m
TNO	(48, 91)	11	Heredity	30 m
DSD	(21, 65)	11	Heredity	10 m
Supersaturated	(24, 138)	6	None	> 1h 10m