SUPPLEMENT

Optimal Experimental Design in the Presence of Nested Factors

Peter Goos

Faculty of Bioscience Engineering and Leuven Statistics Research Centre

KU Leuven

Faculty of Applied Economics and StatUa Center for Statistics

Universiteit Antwerpen

Bradley Jones

SAS Institute

December 12, 2018

1 Magnetic bead experiment

Decrop et al. (2016, 2017) describe a response surface experiment concerning the optical manipulation of magnetic beads. The experiment involved a two-level categorical factor (bead type), a seven-level categorical factor (surfactant) and three quantitative factors (ionic buffer strength, buffer pH and surfactant concentration). The seven levels of the surfactant factor were 'None', 'Brij35', 'Pluronic-F68', 'Tween20', 'Tween40', 'Tween60' and 'Tween80'. Obviously, the factor surfactant concentration is only relevant in the event the level of the surfactant factor differs from 'None'. The investigators expected the effects of the surfactant concentration to depend on the type of surfactant. As they also expected surfactant concentration to be involved in two-way interactions with the factors bead type, ionic buffer strength and buffer pH, this experiment includes both conditional main effects and conditional quadratic effects of the nested factor, surfactant concentration, as well as conditional two-factor interaction effects involving it.

In this experiment, the surfactant factor is a seven-level branching factor, defining six nesting relationships and six nested factors. For all experimental tests using a surfactant, the level of the nested factor, surfactant concentration, was in the range from 0.05% to 1%. The factors bead type, ionic buffer strength and buffer pH are shared factors, when using the terminology of Hung et al. (2009).

Decrop et al. (2016, 2017) studied three levels for the shared quantitative factors ionic buffer strength and buffer pH and four levels for the nested factor surfactant concentration. They eventually used the following *a priori* model to create a D-optimal experimental design involving 80 runs:

$$\begin{aligned} Y = \beta_0 + \beta_{\text{bead}} x_{\text{bead}} + \sum_{j=1}^{6} \beta_{\text{surf}\,j} z_{\text{surf}\,j} + \sum_{i=1}^{2} (\beta_i x_i + \beta_{ii} x_i^2) + \beta_{12} x_1 x_2 + \sum_{i=1}^{2} \beta_i^{\text{bead}} x_i x_{\text{bead}} \\ + \sum_{i=1}^{2} \sum_{j=1}^{6} (\beta_i^{\text{surf}\,j} x_i + \beta_{ii}^{\text{surf}\,j} x_i^2) z_{\text{surf}\,j} + \sum_{j=1}^{6} \beta_{\text{surf}\,j}^{\text{bead}} z_{\text{surf}\,j} x_{\text{bead}} + \sum_{j=1}^{6} (\beta_3^{\text{surf}\,j} x_3 + \beta_{33}^{\text{surf}\,j} x_3^2) z_{\text{surf}\,j} \\ + \left(\sum_{j=1}^{6} z_{\text{surf}\,j} \right) \sum_{i=1}^{2} \beta_{i3} x_i x_3 + \left(\sum_{j=1}^{6} z_{\text{surf}\,j} \right) \beta_3^{\text{bead}} x_3 x_{\text{bead}} + \varepsilon, \end{aligned}$$

where

- Y denotes the response,

- x_{bead} is an effects-type coded two-level categorical variable which takes the value $+1$ for one bead type and the value -1 for the other,
- $z_{\text{surf } j}$ is an indicator variable which takes the value 1 if the jth surfactant is used and the value 0 otherwise,
- x_1, x_2 and x_3 represent the levels of the quantitative factors ionic buffer strength, buffer pH and surfactant concentration,
- ε is the error term, and
- $\beta_0, \beta_{\text{bead}}, \beta_{\text{surf 1}}, \ldots, \beta_{\text{surf 6}}, \beta_1, \beta_2, \beta_{11}, \beta_{22}, \beta_{12}, \beta_1^{\text{bead}}, \beta_2^{\text{bead}}, \beta_1^{\text{surf 1}}, \ldots, \beta_2^{\text{surf 6}}, \beta_{11}^{\text{surf 1}}, \ldots, \beta_{22}^{\text{surf 6}}, \beta_{\text{surf 1}}^{\text{bead}}, \ldots, \beta_{\text{surf 6}}^{\text{bead}}, \beta_3^{\text{surf 1}}, \ldots, \beta_3^{\text{surf 6}}, \beta_{33}^{\text{surf 1}}, \ldots, \beta_{33}^{\text{surf } j}, \beta_{13}, \beta_{23}$, and β_3^{bead} are the 60 regression coefficients.

This model reduces to a second-order response surface model in the factors ionic buffer strength and buffer pH for any bead type in the event no surfactant is used and all indicator variables are zero. However, whenever a surfactant is used, the model reduces to a second-order response surface model in the factors ionic buffer strength, buffer pH and surfactant concentration for any bead type.

One feature of the model is that the nested factor, surfactant concentration (x_3), only enters the model through cross-products terms involving the indicator variables $z_{\text{surf } j}$. The terms involving the cross-products $x_3 z_{\text{surf } j}$ and $x_3^2 z_{\text{surf } j}$ define the conditional main effects and the conditional quadratic effects, respectively. The cross-products $\left(\sum_{j=1}^{6} z_{\text{surf } j}\right) x_i x_3$ and $\left(\sum_{j=1}^{6} z_{\text{surf } j}\right) x_3 x_{\text{bead}}$ define conditional interaction effects of the surfactant concentration (x_3). In using these cross-products, rather than six cross-products of the form $z_{\text{surf } j} x_i x_3$ and six of the form $z_{\text{surf } j} x_3 x_{\text{bead}}$, Decrop et al. (2016, 2017) assumed that the conditional interaction effects do not differ across the six surfactants. Finally, because the number of nesting relationships (six) is smaller than the number of levels of the branching factor (seven), there is no need to drop one of the six indicator variable terms for the model to be identified. For this reason, all six terms of the form $\beta_{\text{surf } j} z_{\text{surf } j}$ are present in the model.

Decrop et al. (2016, 2017) used the SAS procedure OPTEX to create their D-optimal design. They allowed each quantitative shared factor to take three levels and the nested

factor to take four levels, and used a candidate set involving 450 potential test combinations ($2 \times 3^2 = 18$ test combinations without surfactant and $2 \times 3^2 \times 4 = 72$ test combinations for each of the six surfactants).

2 Illustration of the modified coordinate-exchange algorithm

This example illustrates the following operations unique to the modified coordinate-exchange algorithm:

1. Changing a branching factor's level from 0 to 1.
2. Changing a branching factor's level from 1 to 0.
3. Optimizing a nested factor's setting when the associated branching factor's level is 1.

First, the algorithm generates a starting design. Suppose that the following table contains a 4-run starting design for the proof-of-concept example in Section 2 of the main paper:

Run	z	x
1	0	−0.8
2	0	0.5
3	1	−0.5
4	1	0.7

The columns labeled z and x correspond to the branching factor (machine) and the dial factor, respectively, in the example. A zero value for z means that the old machine (without the dial) is used, while a one means that the new machine (with dial) is used. Note that the column for the dial factor x takes random values on the interval $[-1, +1]$ in the starting design.

The model matrix $\mathbf{X}$ corresponding to the starting design involves columns for the intercept, the term involving z and the term involving zx:

Int.	z	zx
1	0	0.0
1	0	0.0
1	1	−0.5
1	1	0.7

The D-optimality criterion value corresponding to that model matrix, $|\mathbf{X}'\mathbf{X}|$, is 2.88.

The first change to the starting design the modified coordinate-exchange algorithm will consider is a swap of the level of the branching factor, z, from 0 to 1, in the first run. Since such a swap means that the new machine (with dial) will be used, it requires the setting of the dial factor, x, to be optimized as well. Suppose that the algorithm first considers changing x to -1. The design matrix corresponding to these changes in z and x is

Run	z	x
1	1	−1.0
2	0	0.5
3	1	−0.5
4	1	0.7

The corresponding model matrix $\mathbf{X}$ is

Int.	z	zx
1	1	−1.0
1	0	0.0
1	1	−0.5
1	1	0.7

The D-optimality criterion corresponding to the modified design is 4.58. In a next step, the algorithm will consider changing x to 1 instead of -1 in the first run. The design matrix resulting from that change is

Run	z	x
1	1	1.0
2	0	0.5
3	1	−0.5
4	1	0.7

The corresponding model matrix is

Int.	z	zx
1	1	1.0
1	0	0.0
1	1	−0.5
1	1	0.7

The D-optimality criterion corresponding to this model matrix is 3.78, which is smaller than the value of 4.58 obtained for $x = -1$. Therefore, it is best to switch z to 1 and x to -1.

Now, consider row 3, to illustrate the effect of changing the level of the branching factor, z, from 1 to 0. This means that, for the third run, we would use the old rather than the new machine. The resulting design matrix is

Run	z	x
1	1	1.0
2	0	0.5
3	0	−0.5
4	1	0.7

The corresponding model matrix is

Int.	z	zx
1	1	−1.0
1	0	0.0
1	0	0.0
1	1	0.7

The D-optimality criterion value for that model matrix is 5.78, which is better than the previous best value of 4.58. Therefore, it is better to switch the value of z from 1 to 0 in row 3. Note that the value of x for this row is now irrelevant because zx is zero, irrespective of the value of x.

Now, consider row 4. The effect of changing z from 1 to 0 in that row on the design matrix is as follows:

Run	z	x
1	1	1.0
2	0	0.5
3	0	−0.5
4	0	0.7

The corresponding model matrix is

Int.	z	zx
1	1	−1.0
1	0	0.0
1	0	0.0
1	0	0.0

The D-criterion value corresponding to that model matrix is 0, which is worse than the current best determinant of 5.78. Therefore, it is better to switch z back to 1.

Now, since z is a branching factor with level 1, we need to optimize the setting of the associated nesting factor x. If x is set to −1, the D-criterion value becomes 0 again. So, consider the effect of setting x to 1:

Run	z	x
1	1	−1.0
2	0	0.5
3	0	−0.5
4	1	1.0

The corresponding model matrix is

Int.	z	zx
1	1	−1.0
1	0	0.0
1	0	0.0
1	1	1.0

The D-criterion value corresponding to that model matrix is 8, which is better than the previous best determinant of 5.78. Therefore, it is better to switch x to 1 in the fourth row.

References

Decrop, D., Brans, T., Gijsenbergh, P., Lu, J., Spasic, D., Kokalj, T., Beunis, F., Goos, P., Puers, R. and Lammertyn, J. (2016). Optical manipulation of single magnetic beads in a microwell array on a digital microfluidic chip, *Analytical Chemistry* **88**(17): 8596–8603.

Decrop, D., Lammertyn, J. and Goos, P. (2017). Optimal experimental design for efficient optical manipulation of magnetic beads seeded in a microwell array, *Journal of Quality Technology* **49**(4): 402–417.

Hung, Y., Joseph, V. R. and Melkote, S. N. (2009). Design and analysis of computer experiments with branching and nested factors, *Technometrics* **51**: 354–365.