Using Particle Swarm Optimization to identify D-optimal designs for binary response generalized linear models with mixed factors

Joshua Lukemire
Emory University

Design Problem

Our goal is to design experiments to obtain parameter estimates for the \( \beta \) factor model taking binary response

\[ \logit(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \]

D-optimal designs aim to minimize the area of the confidence ellipsoid for the parameter estimates, which corresponds to maximizing the determinant of the Fisher information matrix. The Fisher information matrix for an approximate design \( t \) is given by

\[ I_t = \sum_{i=1}^{n} (\nabla^2 \logit(p_i)) (X_i^T X_i) \]

Relative Efficiency When \( c-\text{Log}-\text{Log} \) is True Link

\[ -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 \]

\[ -3 -2 -1 0 1 2 3 \]

\[ 0.0 0.2 0.4 0.6 0.8 1.0 \]

Results

Structure of \( Z^2 \) Designs

We construct D-optimal designs for the model

\[ \logit(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \]

where \( x_1 \in [-1, 1] \) and \( x_2 \in [-1, 1] \). Our goal is to identify regions where minimally supported designs can be constructed as well as to identify regions in which designs can be constructed on the factor boundaries.

Simulation 1 Setup

- PSO settings: 25 particles, 100 maximum iterations, 500 maximum function evaluations
- For each combination of parameter values use PSO to identify D-optimal design using logit, probit, log-log, and complementary log-log links.
- Compare the relative efficiency of the logit-based design to the design identified using the correct link function.

Simulation 2 Results

- True Link
  - Quantile: 0.0100 1.0000 0.0000
  - Log-log: 0.0000 0.9900 0.0000
- Table 2: D-optimal designs obtained with (left) and without (right) conforming to the constraints.

Conclusions

- Particle swarm optimization is a powerful tool for identifying D-optimal designs for mixed factor experiments taking a binary response.
- For 2 factor experiments, minimally supported designs are often available. For those designs which are not minimally supported, it is often not optimal to place all experimental units at the factor boundaries.
- The designs obtained by PSO are generally highly robust against misspecification of the link function.

References


Irregular Design Regions

- Hypothetical plastic molding experiment similar to the one in (Anderson and Whitcomb, 2004).
- 2 continuous factors, temperature and pressure constrained to \( 560 \leq 10 \times \text{Temperature} + \text{Pressure} \leq 5800 \)

Figure 3: Plots of identified designs over the design space. The red diamonds correspond to the design that ignores the constraint, and the black dots correspond to the design that conforms to the constraint.

Table 2: D-optimal designs obtained with (left) and without (right) conforming to the constraints.