Exact D-optimal Design for Multi-factor Nonlinear Models
Yuanzhi Huang 1,2 Steven G. Gilmour 1 Kalliopi Mylona 1 Peter Goos 1,2
1 Southamton Statistical Sciences Research Institute, United Kingdom 2 Universiteit Antwerpen, Belgium

Multi-factor Nonlinear Models

In chemical and biological studies of mechanisms, statistical modeling can help scientists to find accurate information and conclusions.

Empirical models are common in industrialized experiments, which are lack of support from relevant scientific theories. As we assume the suitable form of the model, multiple controlled factors can be implicated.

Mechanistic models are of values in approximation and interpretation, which can be extrapolated from theoretical assumptions. Model function is usually nonlinear and more fragil in the use of parameters.

Mechanistic characteristics can feature even in empirical models (see 2nd ex.).

Optimal Design of Experiments

Design X consists of n runs. For i = 1, 2, · · · , n, the response is

\[ Y_i = f(X_i, \theta) + \varepsilon_i \]

where \( f(X_i, \theta) \) denotes the model function, \( \theta \) is the vector of parameters. As \( F = \partial f(X)/\partial \theta \), the local D-criterion aims to maximize

\[ \phi_D = \log |F^tF| \]

Essential Input of the Exchange Algorithm

1st Example: A Chemical Engineering Application

Yield \( \eta \) is response of mechanistic model (1) in Box and Draper (1987, chap.12), which involves flow rate \( R \), catalyst concentration \( C \) and temperature \( T \):

\[ \eta = \frac{(R_0 + C_0^2 \theta_0 \exp(\theta_0))(R_0 + C_0^2 \theta_0 \exp(\theta_0))}{x_0} + \varepsilon, \]

where \( x = 0.0028344 - 1/(T + 273) \) and \( \theta_0 \) equals the old estimate \( \theta_0 = 5.9, \theta_0 = 1.15, \theta_0 = 0.53, \theta_0 = -0.01, \theta_0 = 15475, \theta_0 = 7489 \).

For the X of 24 runs, the variable space \( \mathcal{X} = [1.5, 6] \times [1, 4] \times [70, 90] \).

Our reference \( X_{ref} \) is face-centered, of which \( \phi_D = -52.7712 \).

Implementation to the 1st Example

Exchanging rule for continuous optimization. stepwise improvement > 1.00001.

| Table 1: X_{ref} after Continuous Optimization of the Hybrid Method |
|---|---|---|---|---|---|---|
| (a) Point Exchange | (b) Coordinate Exchange |
| \( R \) | \( C \) | \( T \) | \( R \) | \( C \) | \( T \) | \( R \) | \( C \) | \( T \) |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |
| 1.5 | 1 | 70 | 1.5 | 1 | 70 | 1.5 | 1 | 70 |

Bayesian Optimal Design

Assume the Normal Prior for \( a_0 - a_4 \) and Log Normal Prior for \( (a_5 + 2.5) \). Prior standard deviations are set to \{0.1936, 0.3717, 0.0537, 0.492, 0.2611, 0.0822\}. \( \sqrt{5} \) times the fitted standard errors.

According to Chaloner and Verdinelli (1995), the expected criterion maximizes the multiple definite integral

\[ \varphi = \int \phi(X|\theta) \pi(\theta) d\theta. \]

If we choose Gauss-Hermite quadrature rule for the hybrid method, the sample size of \( \theta^* \) is 72 for reliable approximation.

References:


2nd Example: Depolymerization (Mountzouris et al., 1999)

The empirical hybrid model (2) explains conversion rate \( \xi \) in terms of substrate concentration \( S \), enzyme concentration \( E \) and transmembrane pressure \( P \):

\[ \xi = \exp(a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5) \frac{S}{\alpha_1 + S}, \]

where \( x_2 = \log_6(E/6.25) \) and \( x_2 = (P - 300)/100 \).

After the substitution \( a_0 = 0.4340, a_1 = 1.3140, a_2 = -1.0595, a_3 = -0.8224, a_4 = 0.4105, a_5 = -2.0633 \), \( \phi_D = 31.7538 \).

Exchanging rule for continuous optimization: improvement > 1.0000001.