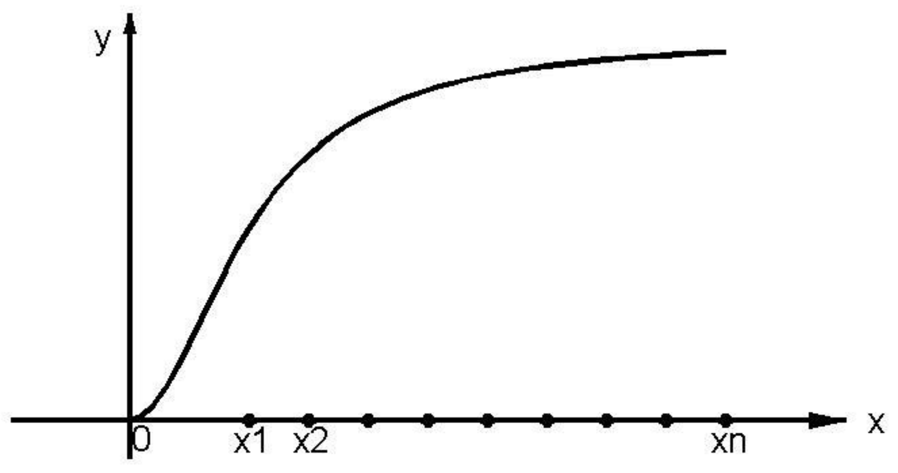
Optimal designs for regression models using the second-order least squares estimator

Yue Yin and Julie Zhou

Department of Mathematics and Statistics, University of Victoria

Background

Consider a regression model $y_i = g(\mathbf{x}_i;\boldsymbol{\theta}) + \varepsilon_i,\ i = 1,\cdots,n$, where y_i is the ith observed response at $\mathbf{x}_i$, $\boldsymbol{\theta} \in R^q$ is the unknown parameter vector, $g(\mathbf{x}_i;\boldsymbol{\theta})$ can be a linear or nonlinear function of $\boldsymbol{\theta}$, and ε_i's are i.i.d having mean 0 and variance σ^2.



Design problem

Find optimal distribution $\xi(\mathbf{x})$ of $\mathbf{x}$ to minimize the variance of $\hat{\boldsymbol{\theta}}$, the estimator of $\boldsymbol{\theta}$.

$$\xi(\mathbf{x}) = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \\ w_1 & w_2 & \ldots & w_n \end{pmatrix}$$

Research Question

The ordinary least squares estimator (OLSE) performs well under the assumption that the error distribution is normal or symmetric. However, it is not efficient if the error distribution is not symmetric.

We will explore the optimal designs when the error distribution is asymmetric.

Methods

From Wang and Leblanc (2008), the second-order least squares estimator (SLSE) $\hat{\boldsymbol{\gamma}}_{SLS}$ of $\boldsymbol{\gamma} = (\boldsymbol{\theta}^{\top}, \sigma^2)^{\top}$ minimizes

$$Q_n(\gamma) = \sum_{i=1}^{n} \begin{pmatrix} y_i - g(\mathbf{x}_i;\boldsymbol{\theta}) \\ y_i^2 - g^2(\mathbf{x}_i;\boldsymbol{\theta}) - \sigma^2 \end{pmatrix}^{\mathrm{T}} \mathbf{W}_i \begin{pmatrix} y_i - g(\mathbf{x}_i;\boldsymbol{\theta}) \\ y_i^2 - g^2(\mathbf{x}_i;\boldsymbol{\theta}) - \sigma^2 \end{pmatrix},$$

where $\mathbf{W}_i$'s are 2×2 positive semidefinite matrices.

The optimal SLSE has

$$Cov(\hat{\boldsymbol{\theta}}_{SLSE}) = (1-t)\sigma_0^2(\mathbf{G}_2 - t\mathbf{g}_1\mathbf{g}_1^{\mathrm{T}})^{-1},$$

where

$$\mathbf{g}_1(\mathbf{w}) = \mathrm{E}_{\xi}\left[\frac{\partial g(\mathbf{x};\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}\right], \mathbf{G}_2(\mathbf{w}) = \mathrm{E}_{\xi}\left[\frac{\partial g(\mathbf{x};\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \frac{\partial g(\mathbf{x};\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}^{\mathrm{T}}}\right],$$

$\boldsymbol{\theta}_0$ and σ_0 are true values for $\boldsymbol{\theta}$ and σ, respectively, and we have $0 \le t < 1$ for any error distribution, which is an indicator to reflect the skewness of error distribution.

Loss function

Let $\mathbf{A}(\mathbf{w}) = \mathbf{G}_2(\mathbf{w}) - t\mathbf{g}_1(\mathbf{w})\mathbf{g}_1^{\mathrm{T}}(\mathbf{w})$,

- D-optimality: $\phi_D(\mathbf{w}) = \det(\mathbf{A}(\mathbf{w})^{-1})$
- A-optimality: $\phi_A(\mathbf{w}) = \mathrm{trace}(\mathbf{A}(\mathbf{w})^{-1})$

Design criteria

D-optimality criterion

↓

Minimize $\det(Cov(\hat{\boldsymbol{\theta}}_{SLS}))$

↓

$\min_{\mathbf{w}}$ $\phi_D(\mathbf{w})$
subject to: $\Sigma w_i = 1$, $w_i \ge 0$

A-optimality criterion

↓

Minimize $\mathrm{trace}(Cov(\hat{\boldsymbol{\theta}}_{SLS}))$

↓

$\min_{\mathbf{w}}$ $\phi_A(\mathbf{w})$
subject to: $\Sigma w_i = 1$, $w_i \ge 0$

Results and discussion

CVX and SeDuMi programs in Matlab

Define

$$\mathbf{B}(\mathbf{w}) = \begin{pmatrix} 1 & \sqrt{t}\mathbf{g}_1^{\mathrm{T}}(\mathbf{w}) \\ \sqrt{t}\mathbf{g}_1(\mathbf{w}) & \mathbf{G}_2(\mathbf{w}) \end{pmatrix}.$$

$\mathbf{B}(\mathbf{w})$ is a linear function of $\mathbf{w}$, and we have $\det(\mathbf{A}(\mathbf{w})) = \det(\mathbf{B}(\mathbf{w}))$.

CVX program for D-optimal design

Minimize $\phi_D(\mathbf{w})$ ⇔ Minimize $-\log(\det(\mathbf{B}(\mathbf{w})))$ or $-(\det(\mathbf{B}(\mathbf{w})))^{1/(q+1)}$

Also, $\phi_A(\mathbf{w}) = \mathrm{trace}(\mathbf{A}(\mathbf{w})^{-1}) = \mathrm{trace}(\mathbf{C}(\mathbf{B}(\mathbf{w}))^{-1})$, where $\mathbf{C} = 0 \oplus \mathbf{I}_q$.

Define $\mathbf{B}_i = \begin{pmatrix} \mathbf{B}(\tilde{\mathbf{w}}) & e_i \\ e_i^{\mathrm{T}} & v_i \end{pmatrix}$ for $i = 2,\cdots,q+1$,

$$\mathbf{H}(\tilde{\mathbf{w}},\mathbf{v}) = \mathbf{B}_2 \oplus \cdots \oplus \mathbf{B}_{q+1} \oplus \mathbf{D}(\tilde{\mathbf{w}}),$$

where $\tilde{\mathbf{w}} = (w_1, w_2, \cdots, w_{n-1}, 1 - \sum_{i=1}^{n-1} w_i)^{\top}$, e_i is the ith unit vector in R^{q+1}, $\mathbf{v} = (v_2, \cdots, v_{q+1})^{\top}$ and $\mathbf{D}(\tilde{\mathbf{w}}) = diag(w_1, \cdots, w_{n-1}, 1 - \sum_{i=1}^{n-1} w_i)$.

SeDuMi program for A-optimal design

$\min_{\mathbf{w}}$ $\phi_A(\mathbf{w})$ subject to: $\Sigma w_i = 1$, $w_i \ge 0$ ⇔ $\min_{\tilde{\mathbf{w}},\mathbf{v}}$ $\Sigma_{i=2}^{q+1} v_i$ subject to: $\mathbf{H}(\tilde{\mathbf{w}},\mathbf{v}) \succcurlyeq 0$

Invariance properties of A-optimal design

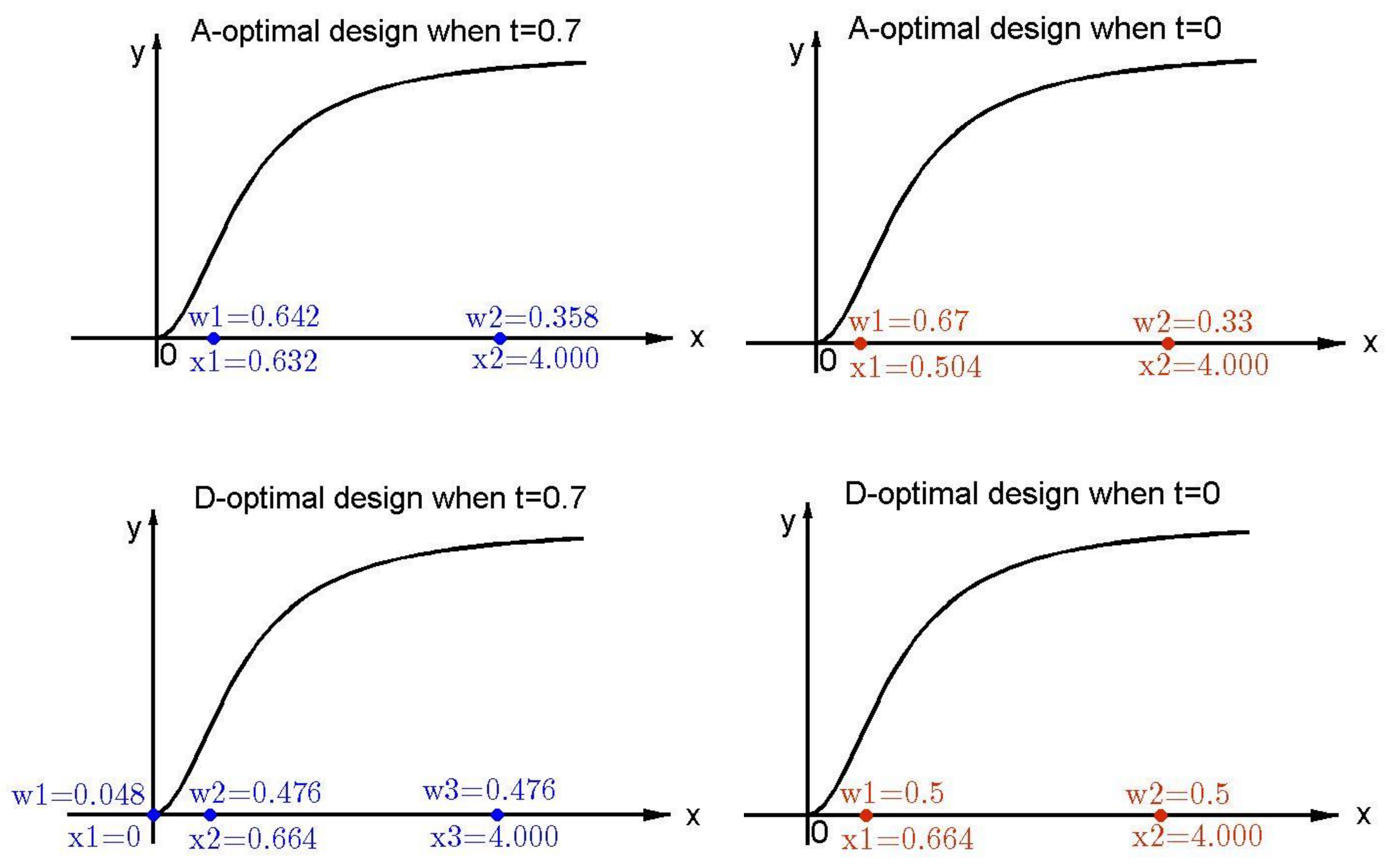
For symmetric design space, if the model follow certain properties, the A-optimal design is also symmetric.

Examples

Emax model:

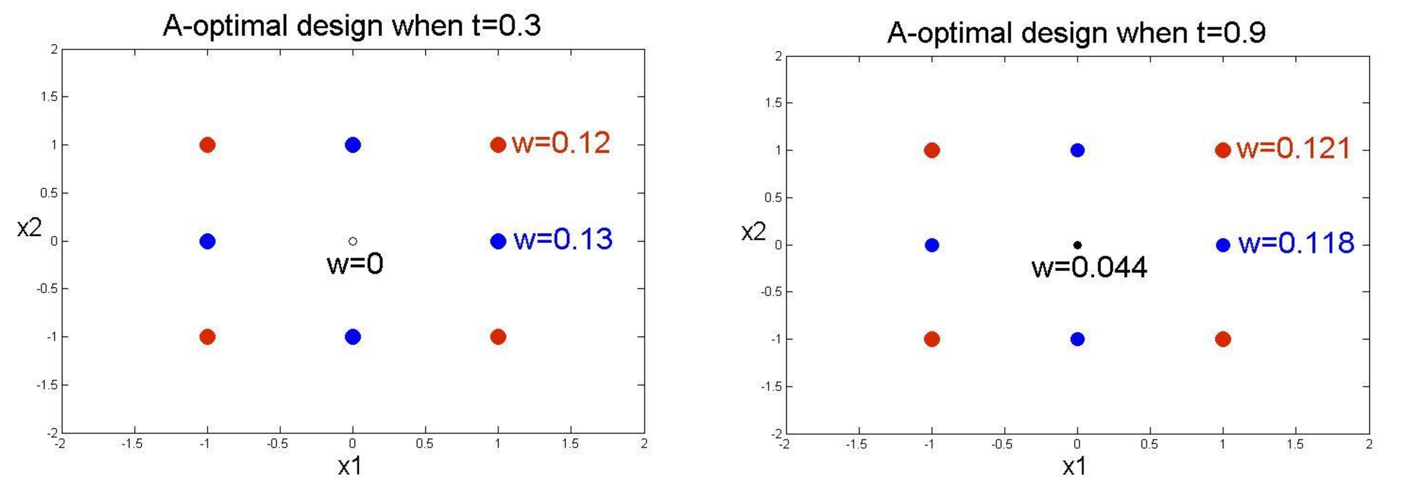
$$y_i = \frac{\alpha x_i^{\beta_2}}{\beta_1 + x_i^{\beta_2}} + \varepsilon_i,$$

design space S=[0,5], $\alpha = 1, \beta_1 = 1, \beta_2 = 2$, locally optimal designs:



Quadratic model with two variables:

$$y_i = \theta_1 x_{i1} + \theta_2 x_{i2} + \theta_3 (x_{i1})^2 + \theta_4 (x_{i2})^2 + \theta_5 x_{i1} x_{i2} + \varepsilon_i$$



References

[1].Wang, L. and Leblanc, A. (2008). Second-order nonlinear least squares estimation. *Annals of the Institute of Statistical Mathematics*, 60, 883-900.

[2].Yin,Y. and Zhou, J. (2016). Optimal designs for regression models using the second-order least squares estimator, submitted.