Inverse problems for Bayesian errors-in-variables regression models

Adam Gothorp[1], Eleanor Stillman[1], Paul Blackwell[1], Candice Majewski[2]

[1]School of Mathematics and Statistics, University of Sheffield

[2]Department of Mechanical Engineering, University of Sheffield

Introduction

The concept of inverse problems in linear regression can be split into three key ideas:

1. Discover relationship between response variable(s) and explanatory variable(s)
2. Take some 'desired' value(s) for the response(s)
3. Find values for the explanatory variable(s) that would produce desired response(s), given relationship found in forward model

Area of research. The aim of this PhD is to improve the efficiency and effectiveness of 3D printing; here, this is by modelling flow characteristics of the powder used in terms of other powder properties, and then choosing those properties to optimise flow characteristics.

Why Bayesian?

1. Data involves replicate measurements on multiple groups of powder, which implies measurement error, which implies errors-in-variables models; the error is then naturally accounted for in a Bayesian setting
2. Inverting the relationship between the response variable(s) and the explanatory variable(s), particularly in a multivariate case, is relatively straightforward

Forward model

The forward model describes finding the relationship between the explanatory variable(s) and the response variable(s).

The example given here will consider a bivariate response and two explanatory variables:

$$\begin{pmatrix}\widetilde{Y}_{1i}\\ \widetilde{Y}_{2i}\end{pmatrix} = \begin{pmatrix}\beta_{01} & \beta_{11} & \beta_{21}\\ \beta_{02} & \beta_{12} & \beta_{22}\end{pmatrix}\begin{pmatrix}1\\ \widetilde{X}_{1i}\\ \widetilde{X}_{2i}\end{pmatrix} + \begin{pmatrix}\varepsilon_{1i}\\ \varepsilon_{2i}\end{pmatrix},$$

$$\begin{pmatrix}\varepsilon_{1i}\\ \varepsilon_{2i}\end{pmatrix} \sim \mathrm{N}\left(0, T_{\widetilde{Y}}\right),$$

$$T_{\widetilde{Y}} \sim \mathrm{Wishart}_2(0.001\mathrm{I}_2, 2)$$

$$Y_{1ij_1} = \widetilde{Y}_{1i} + \eta_{1ij_1}, \eta_{1ij_1} \sim \mathrm{N}(0, \tau_{Y_1}),$$

$$Y_{2ij_2} = \widetilde{Y}_{2i} + \eta_{2ij_2}, \eta_{2ij_2} \sim \mathrm{N}(0, \tau_{Y_2}),$$

$$\tau_{Y_1} \sim \mathrm{Gamma}(0.001, 0.001),$$

$$\tau_{Y_2} \sim \mathrm{Gamma}(0.001, 0.001),$$

$$\begin{pmatrix}X_{1ik}\\ X_{2ik}\end{pmatrix} = \begin{pmatrix}\widetilde{X}_{1i}\\ \widetilde{X}_{2i}\end{pmatrix} + \begin{pmatrix}\delta_{1ik}\\ \delta_{2ik}\end{pmatrix}, \begin{pmatrix}\delta_{1ik}\\ \delta_{2ik}\end{pmatrix} \sim \mathrm{N}\left(0, T_X\right)$$

$$T_X \sim \mathrm{Wishart}_2(0.001\mathrm{I}_2, 2)$$

$$\begin{pmatrix}\widetilde{X}_{1i}\\ \widetilde{X}_{2i}\end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix}\mu_{X_1}\\ \mu_{X_2}\end{pmatrix}, T_{\widetilde{X}}\right)$$

$$T_{\widetilde{X}} \sim \mathrm{Wishart}_2(\mathrm{I}_2, 2)$$

The model is written in JAGS (Just Another Gibbs Sampler) and run in R using the package *rjags*. Uninformative priors are placed on the β matrix, the τ precisions and the T precision matrices.

The plots in Figures 1 and 2 show the fitted values $\widehat{Y}_i$ plotted against the posterior means of $\widetilde{Y}_i$, where $\widehat{Y}_i$ is taken to be the product of the posterior mean β matrix and the posterior mean of $\widetilde{X}_i$.

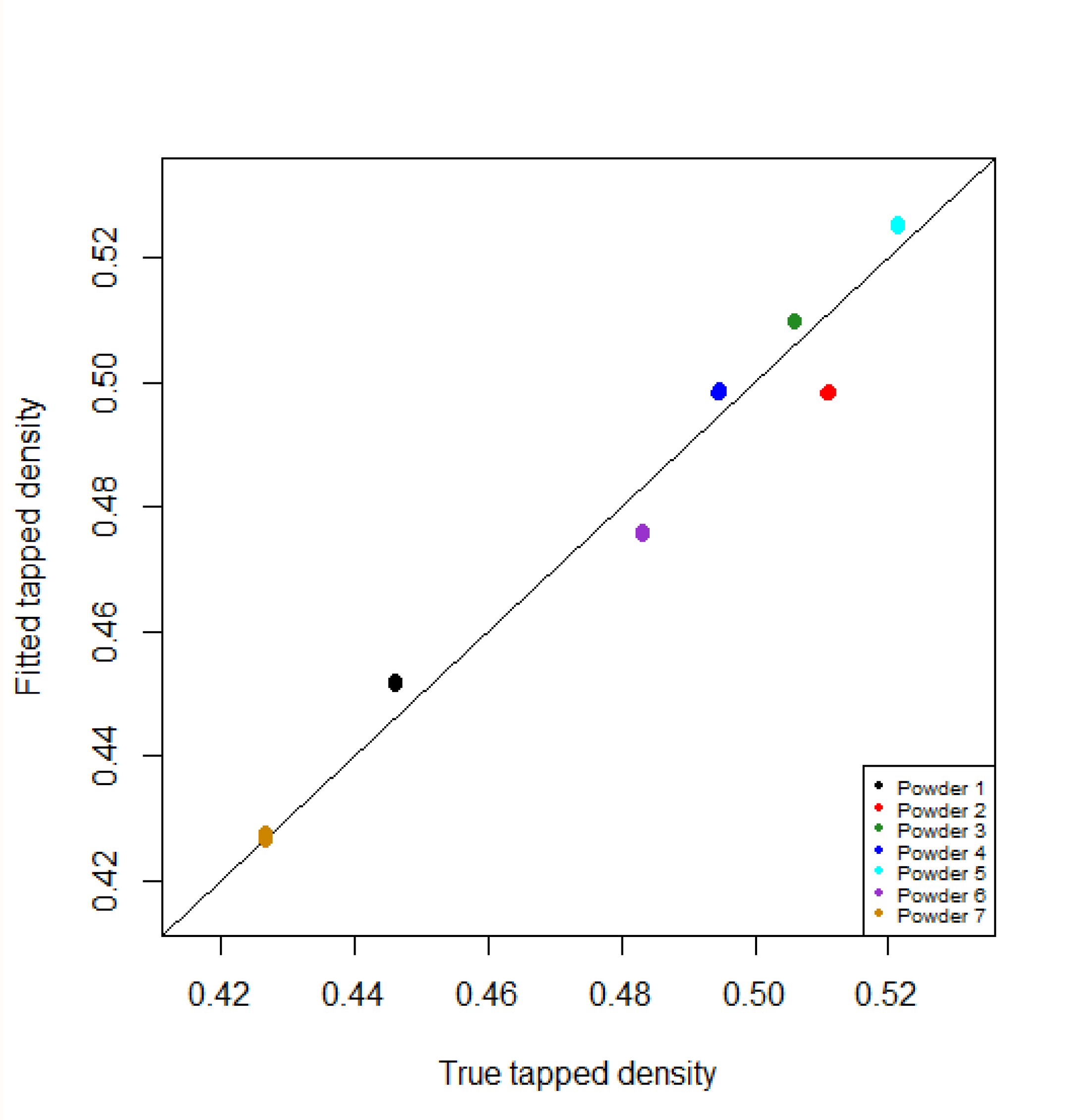

Figure 1: $\widehat{Y}_i$ vs. $\widetilde{Y}_{i,post}$ for tapped density

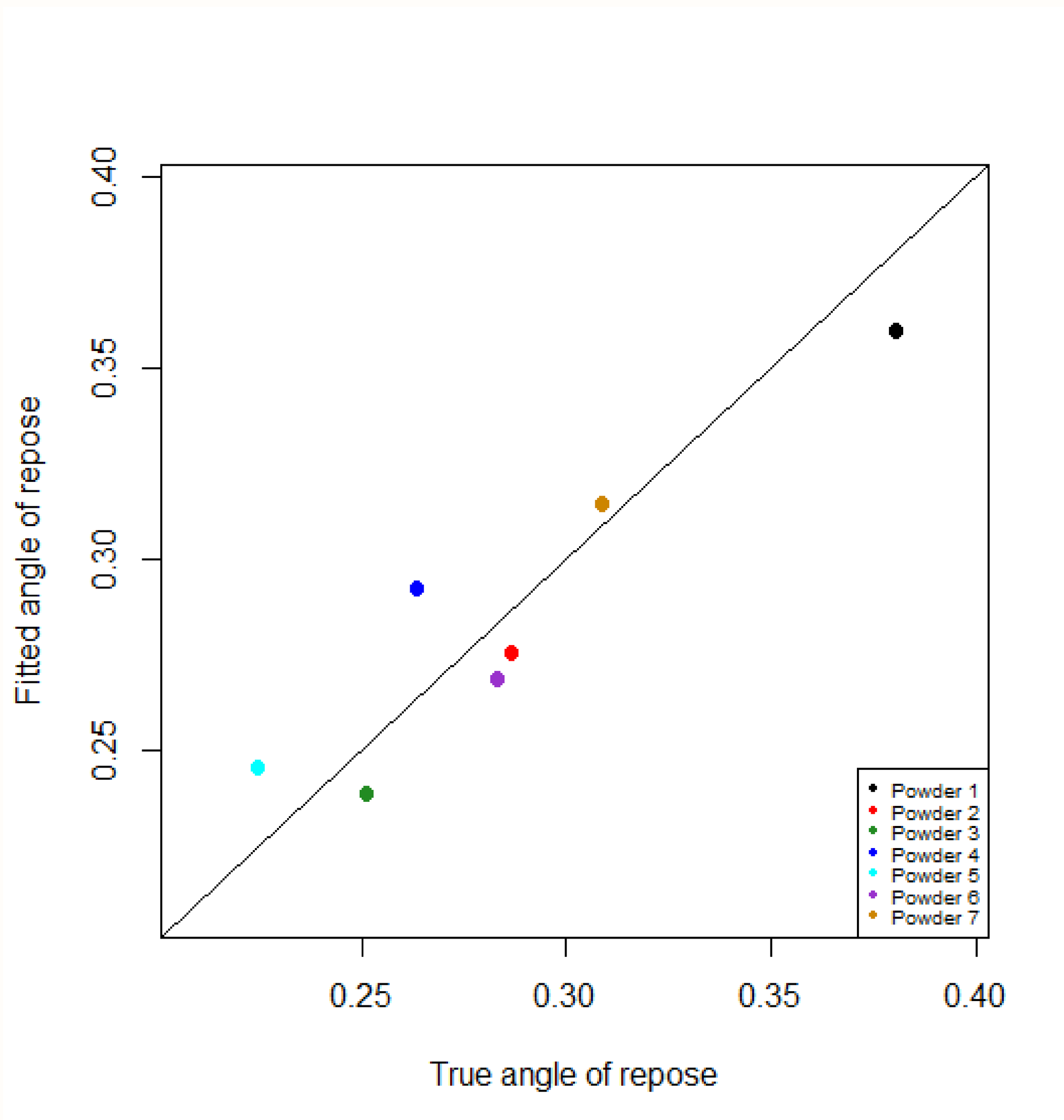

Figure 2: $\widehat{Y}_i$ vs. $\widetilde{Y}_{i,post}$ for angle of repose

Backward model

The idea here:

1. Testing the backward model — produce 'desired' response Y^*, with specific values of X^*
2. Take the product of the posterior mean of β with the vector X^*, then assume X^* to be unknown
3. Take $r = 1, \ldots, 2000$ random samples from joint posteriors of β (matrix of regression coefficients), $T_{\widetilde{Y}}$ and $T_{\widetilde{X}}$
4. For each r, fit the model below and take a random sample from the posterior of X^*
5. Combine together to make posterior distribution of X^*

$$\begin{pmatrix}Y_1^*\\ Y_2^*\end{pmatrix} = \begin{pmatrix}\beta_{01}^r & \beta_{11}^r & \beta_{21}^r\\ \beta_{02}^r & \beta_{12}^r & \beta_{22}^r\end{pmatrix}\begin{pmatrix}1\\ X_1^*\\ X_2^*\end{pmatrix} + \begin{pmatrix}\varepsilon_1^*\\ \varepsilon_2^*\end{pmatrix}$$

with the prior distributions

$$\begin{pmatrix}X_1^*\\ X_2^*\end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix}\mu_{X_1}\\ \mu_{X_2}\end{pmatrix}, T_{\widetilde{X}}^r\right),$$

$$\begin{pmatrix}\varepsilon_1^*\\ \varepsilon_2^*\end{pmatrix} \sim \mathrm{N}\left(0, T_{\widetilde{Y}}^r\right).$$

Note that the superscript used in the β matrix refers to a particular sample from the posterior distribution of β, not a power of β.

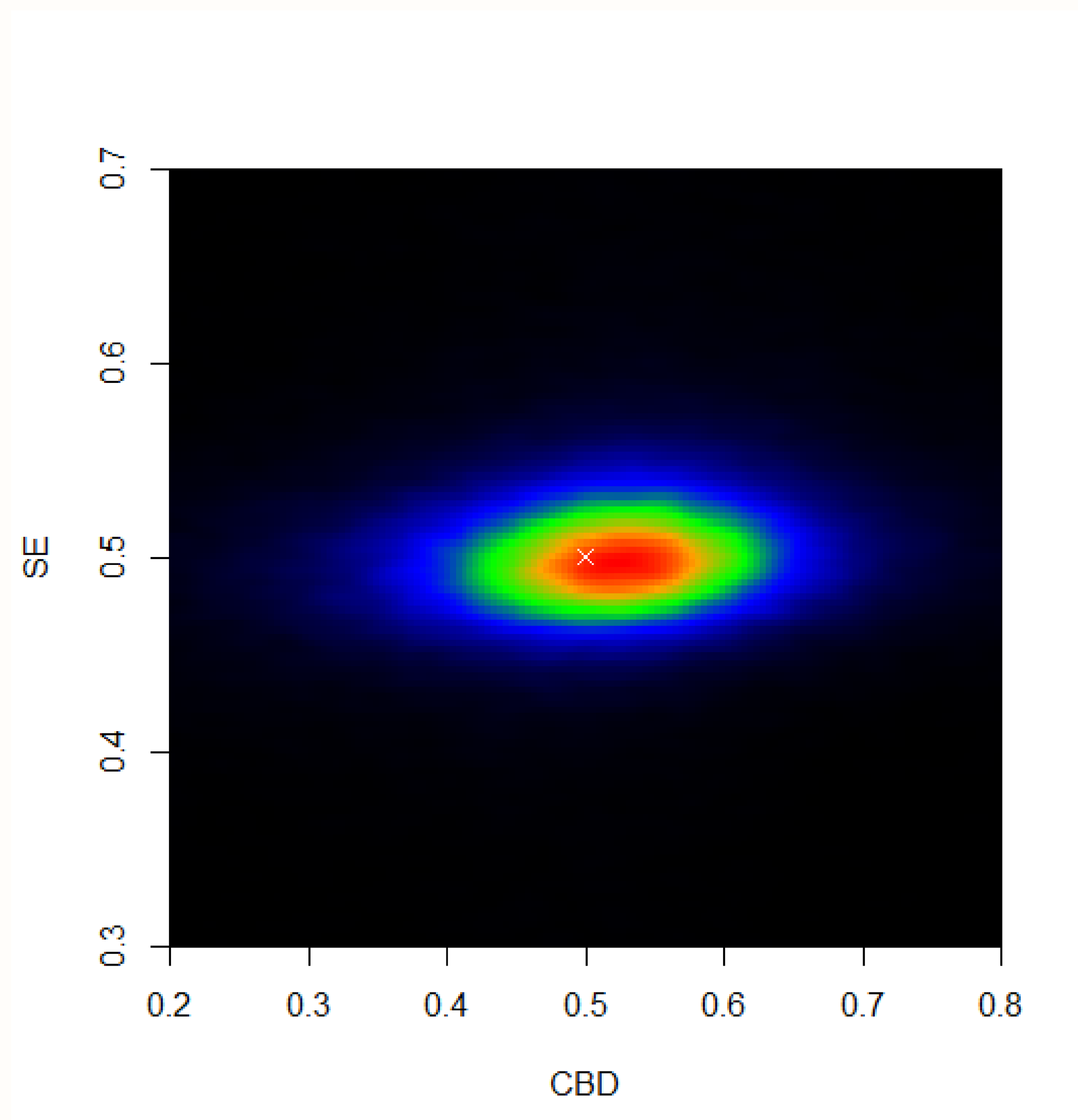

Figure 3: Joint posterior distribution of CBD and SE given 'desired' tapped density of 0.5327 and 'desired' angle of repose of 0.2248

Future work

1. Experiment with prior distribution on X^*
2. Produce more plots showing how probability of producing desired response varies as values of explanatory variables vary
3. Experiment with more complicated forward models and see how that affects backward model

References

[1] A. Gelman et al., *Bayesian Data Analysis, Third Edition.* CRC Press, Taylor Francis Group, Florida, 2013.

[2] M. Plummer and others., 2003, 'JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling', *Proceedings of the 3rd international workshop on distributed statistical computing*, Vienna, Austria.

[3] M. Plummer, A. Stukalov, M. Denwood. Packages 'rjags', *Vienna, Austria*, 2016.

MathODS Virtual Conference 2020