Analysis of measurement data subject to additive and multiplicative effects

Kavya Jagan
National Physical Laboratory, Teddington, Middlesex, UK
E-mail: kavya.jagan@npl.co.uk

Alistair Forbes
National Physical Laboratory, Teddington, Middlesex, UK
E-mail: alistair.forbes@npl.co.uk

Abstract. Many calibration problems in metrology involve fitting a model to measurement data representing the response of a system. The standard assumption about the measurement system is that the random effects associated with the measurements are drawn from the same distribution. In practice however, the uncertainty associated with the measurements may depend on the magnitude of the response being measured. For example, the measurement of displacement using laser interferometry follows such a model in which the uncertainty has a dependence on length due to the uncertainty associated with the refractive index of the air. For such systems, the model is inherently nonlinear for which standard approaches such linear least-squares estimation provide only approximate solutions.

This paper describes a Bayesian approach for analysing such data. The data is modelled as a linear or non-linear response subject to additive and multiplicative noise components. We assume that prior, possibly vague, information about the variances associated with additive and multiplicative noise components is given in terms of Gamma distributions. The Bayesian posterior distribution for such models cannot be expressed analytically in closed form and a Metropolis-Hastings Markov chain Monte Carlo algorithm is used to sample from the posterior distribution. This method is illustrated on data relating to the radioactive decay of Pb211.

Keywords: Additive and multiplicative random effects, Metropolis-Hastings

1. Introduction

Suppose that the model for an indication $y_i$ is

$$ y_i|\alpha \sim N(h_i(\alpha), \sigma^2(\alpha)),$$

$$ \sigma^2(\alpha) = \sigma^2_A + h_i(\alpha)^2 \sigma^2_M, \quad i = 1, \ldots, m, \quad (1) $$

where $h_i(\alpha)$, the modelled response, depends on parameters $\alpha = (\alpha_1, \ldots, \alpha_n)^T$, $n \leq m$. We assume that $\sigma_A$ and $\sigma_M$ are not known exactly but we may have some information about them from a previous calibration, for instance. This model arises if the measuring
Analysis of measurement data subject to additive and multiplicative effects

system is subject to additive (‘A’) and multiplicative (‘M’) effects. For example, the measurement of displacement using laser interferometry follows such a model in which the uncertainty has a dependence on length due to the uncertainty associated with random perturbations of the refractive index of the air. The primary aim is to make inferences about model parameters $\alpha$, but we may also wish to update our knowledge of $\sigma_A$ and $\sigma_M$.

The contribution of each effect can be made more explicit as in

$$y_i = (1 + \beta_i)h_i(\alpha) + \epsilon_i, \quad \beta_i \in N(0, \sigma_M^2), \quad \epsilon_i \in N(0, \sigma_A^2).$$

In a Bayesian framework, information about $\sigma_A$ and $\sigma_M$ can be incorporated into the analysis by assigning a prior distribution to the precision parameters $\phi_A = 1/\sigma_A^2$ and $\phi_M = 1/\sigma_M^2$. Suitable prior distributions for these precision parameters are Gamma distributions. The parameters of the Gamma priors are estimates of the standard deviations $\sigma_{0_A}$ and $\sigma_{0_M}$ and associated degrees of belief on these estimates, $m_{0_A}$ and $m_{0_M}$ respectively. Their density functions up to a normalising constant are given by, e.g.,

$$p(\phi_A) \propto \phi_{m_{0_A}/2}^{-1} \exp \left\{ -\frac{\phi_A}{2m_{0_A}} \sigma_{0_A}^2 \right\}.$$ 

The model in equation (1) defines the likelihood

$$p(y|\alpha, \phi_A, \phi_M) \propto \prod_{i=1}^m \frac{1}{\sigma_i(\alpha)} \exp \left\{ -\frac{1}{2\sigma_i^2(\alpha)} (y_i - h_i(\alpha))^2 \right\}.$$ 

Assigning a non-informative prior for $\alpha$ and given the observed response $y$, by Bayes theorem [3] the posterior distribution for $\alpha$, $\phi_A$ and $\phi_M$ is given by

$$p(\alpha, \phi_A, \phi_M|y) \propto p(y|\alpha, \phi_A, \phi_M)p(\phi_A)p(\phi_M),$$ 

where

$$p(\alpha, \phi_A, \phi_M|y) \propto \prod_{i=1}^m \frac{1}{\sigma_i(\alpha)} \exp \left\{ -\frac{1}{2\sigma_i^2(\alpha)} (y_i - h_i(\alpha))^2 \right\} \phi_{m_{0_A}/2}^{-1} \exp \left\{ -\frac{\phi_A}{2m_{0_A}} \sigma_{0_A}^2 \right\} \phi_{m_{0_M}/2}^{-1} \exp \left\{ -\frac{\phi_M}{2m_{0_M}} \sigma_{0_M}^2 \right\}.$$ 

Since the posterior distribution can be expressed analytically only up to a normalising constant, computational approaches have to be employed to determine summary statistics associated with the model parameters. One approach is to draw samples of the parameters using the Metropolis-Hastings MCMC algorithm.

1.1. Metropolis-Hastings algorithm

Markov chain Monte Carlo methods (MCMC) [1, 2] are sequential, iterative schemes designed to produce samples from a probability distribution $p(\alpha)$. They are constructed in such a way that sampled value at a particular iteration $\{a_q\}$ only depends on the value $a_{q-1}$ at the previous iteration. (They therefore satisfy the Markov property, hence their name.) The sampling scheme for generating the next element of the chain has to
Analysis of measurement data subject to additive and multiplicative effects

be designed in such a way that, eventually, the samples \( a_q \) represent samples from \( p(\alpha) \). One commonly used method to achieve this is the Metropolis-Hastings MCMC method.

Suppose we wish to sample \( \{ a_q \} \) from a target distribution \( p(\alpha) \). Given a draw \( a_{q-1} \), a proposed draw \( a^* \) for the next member of the sequence is generated at random from a jumping distribution \( p_0(a^*|a_{q-1}) \). Then \( a_q \) is set to \( a^* \) with acceptance probability

\[
P_q = \min\{1, r_q\}, \quad r_q = \frac{p(a^*)p_0(a_{q-1}|a^*)}{p(a_{q-1})p_0(a^*|a_{q-1})}.
\]  

(3)

Otherwise, \( a_q \) is set to \( a_{q-1} \). The simplest way to implement the acceptance step is to draw \( u_q \) from the uniform distribution \( U(0,1) \) and if \( u_q < r_q \), set \( a_q = a^* \), otherwise set \( a_q = a_{q-1} \).

2. Numerical example

In this section, we illustrate the performance of the algorithm on data relating to the decay of a radioactive isotope of lead, viz., Pb211. The model can be expressed as

\[
y_i \in N(h_i(\alpha), \phi^{-1}_A + h_i(\alpha)^2\phi^{-1}_M),
\]

where \( h_i(\alpha) = \alpha_1 \exp\{-\alpha_2 x_i\} + \alpha_3 \).

Residuals resulting from a non-linear least squares fit of the exponential decay model to the data is shown in figure 1a. The variation in the residuals decreases over time as the amount of radioactive material decreases. This suggests that the data is subject to additive and multiplicative random effects. Assuming vague Gamma priors for \( \phi_A \) and \( \phi_M \) with parameters \( m_{0,A} = 2, \sigma_{0,A} = 0.001, m_{0,M} = 2 \) and \( \sigma_{0,M} = 0.001 \) respectively, 100 chains of 10000 samples each were generated from the posterior distribution in equation (2). Figure 1b shows the data along with the exponential decay fit derived from the posterior parameter estimates given by the means of the samples. (The data has been scaled to range from 0 to 1 resulting in arbitrary units; rescaling back to standard units is straightforward.)

Figure 1c shows the histogram of samples from the posterior distribution of \( \phi_M \) and a superimposed plot of its vague prior distribution. The posterior distribution for \( \phi_M \) is represented by the histogram. The means of the posterior distributions for \( \phi_A \) and \( \phi_M \) are leads to posterior estimates of \( \hat{\sigma}_A \) and \( \hat{\sigma}_M \), respectively. The quantity \( (\hat{\sigma}_A^2 + \hat{\sigma}_M^2)^{1/2} \) ranges from 0.001 1 for the first data point down to 0.000 2 for the last data point. A unit weighting of the data significantly overweights the first data points and underweights the last data points. The MCMC approach determines an optimal weighting based on the actual data. Moreover, the posterior distribution for \( \alpha \) takes into account the fact that \( \phi_A \) and \( \phi_M \) are not known exactly.

3. Conclusion

This paper describes a Bayesian approach for analysing data that is subject to additive and multiplicative random effects, each parametrized by a variance parameter. The
Analysis of measurement data subject to additive and multiplicative effects

(a) Residuals resulting from a non-linear least squares fit.

(b) Data points and the exponential decay fit derived from the posterior parameter estimates.

(c) Histogram of samples from the posterior distribution of $\phi_M$ and superimposed plot of its prior distribution.

Figure 1: Exponential decay illustration.

The posterior distribution for the fitted parameters takes into account the fact that these variance parameters are also estimated from the data and have associated uncertainties. The Bayesian posterior distribution for such models cannot be expressed analytically in closed form and a Metropolis-Hastings algorithm has been described to sample from the posterior distribution. This method is illustrated on radioactive decay data.

Acknowledgement. This work has been supported by the UK’s National Measurement System Programme for Data Science.

4. References

[1] Siddhartha Chib and Edward Greenberg. Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327–335, 1995.

[2] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton, Fl., third edition, 2014.

[3] D. S. Sivia. *Data Analysis: a Bayesian Tutorial*. Clarendon Press, Oxford, 1996.