Bayesian inference examples

Engineering Risk Analysis Group, Technische Universität München.
Arcisstr. 21, 80333 Munich, Germany.

August 22, 2019

Introduction

Let $\boldsymbol{\theta}$ be a set of uncertain variables, and suppose we have a set of observed or measured data points $\mathcal{D}$. Using Bayes' Theorem, our belief about $\boldsymbol{\theta}$ can be updated by

$$f\left(\boldsymbol{\theta}|\mathcal{D}\right) = \frac{1}{Z}\, f\left(\boldsymbol{\theta}\right) f\left(\mathcal{D}|\boldsymbol{\theta}\right), \tag{1}$$

where the *prior distribution* $f\left(\boldsymbol{\theta}\right)$ represents the initial knowledge about the parameters; the *likelihood function* $f\left(\mathcal{D}|\boldsymbol{\theta}\right)$ stands for the probability of observing data $\mathcal{D}$ conditional on the parameter vector $\boldsymbol{\theta}$; the *model evidence* $Z = f\left(\mathcal{D}\right) = \int f\left(\mathcal{D}|\boldsymbol{\theta}\right) f\left(\boldsymbol{\theta}\right) \mathrm{d}\boldsymbol{\theta}$ acts as a normalizing constant; and the *posterior distribution* $f\left(\boldsymbol{\theta}|\mathcal{D}\right)$ represents the updated belief about $\boldsymbol{\theta}$ after observing the data $\mathcal{D}$.

In the following, we describe 3 examples that are commonly used as benchmarks in Bayesian inference with engineering application. These problems are solved in MATLAB® and Python using the following approaches:

- Bayesian inference with subset simulation (BUS-SuS) [3] :
  Function `BUS_SuS`.
- Adaptive Bayesian inference with subset simulation (aBUS-SuS) [3]:
  Function `aBUS_SuS`.
- Improved transitional Markov chain Monte Carlo (iTMCMC) [3]:
  Function `iTMCMC`.

For further questions or bugs in the codes please write to:
`felipe.uribe@tum.de`.

1 One-dimensional posterior

Consider a one-dimensional inference problem [3]

$$f\left(\theta \mid \mathcal{D}\right) \propto f\left(\theta\right) f\left(\mathcal{D} \mid \theta\right) = \mathcal{N}(\theta; 0, 1) \cdot \mathcal{N}(\theta; \mu, \sigma^2),$$

where the likelihood has parameter $\sigma^2 = 0.04$. The only 'data' point is the mean $\mathcal{D} = \mu = 5$. This problem has analytical solution; the posterior mean and standard deviation are 4.81 and 0.196, respectively. Moreover, the model evidence is $Z = 2.36 \times 10^{-6}$.

2 Multi-modal posterior

We consider the simulation of a multi-modal posterior given by the truncated mixture of bi-variate Gaussian densities [2]

$$f(\boldsymbol{\theta} \mid \mathcal{D}) \propto f(\boldsymbol{\theta})\, f(\mathcal{D} \mid \boldsymbol{\theta}) = \mathcal{U}(\boldsymbol{\theta}; [a_1, a_2], [b_1, b_2]) \cdot \sum_{i=1}^{n} w_i \mathcal{N}(\boldsymbol{\theta}; \boldsymbol{\mu}_i, \sigma^2 \mathbf{I}_2), \tag{2}$$

where the prior is the uniform distribution in the square domain $[a_1, b_1] \times [a_2, b_2]$.

The following values for the parameters are used: The number of densities in the mixture is $n = 10$, the bounds for the prior are $a_1 = a_2 = 0$, $b_1 = b_2 = 10$, the variance is $\sigma^2 = 0.01$, and the weights are given by $\{w_i\}|_i^n = 0.1$. The 'data' points are given by the means $\mathcal{D} = \{\boldsymbol{\mu}_i\}|_i^n$, which are generated by simulating a realization from the prior. Posterior samples of (2) are shown in Figure 1.

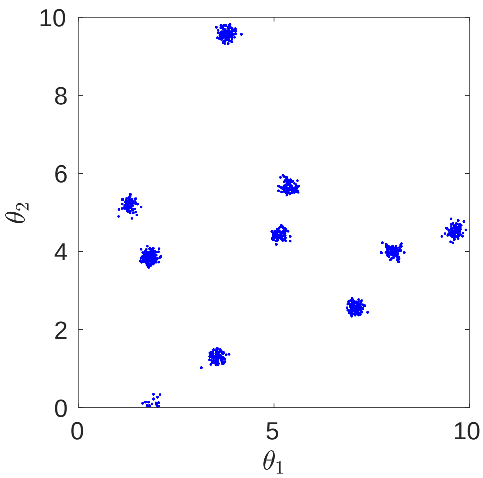

Figure 1: Samples from the multi-modal posterior distribution in Example 2.

3 Two-degree-of-freedom shear building

3.1 Model description

Consider a two-degree-of-freedom shear building model without damping, as shown in Figure 2. This type of representation is commonly used to model moment-resisting frame structures. The objective is to identify its uncertain interstory stiffness parameters based on measurements of the first two eigenfrequencies. According to [1], this problem is not globally identifiable, which means that there exist multiple combinations of values of the model parameters that can approximate the measured data reasonably well.

The story masses are taken as deterministic values, $m_1 = 16531$ and $m_2 = 16131$ kg. The interstory stiffness are parametrized as, $k_1 = \bar{k}_1\theta_1$ and $k_2 = \bar{k}_2\theta_2$, where θ_1 and θ_2 are the stiffness parameters to be identified. The nominal values for the interstory stiffnesses are chosen as, $\bar{k}_1 = \bar{k}_2 = 29.7 \times 10^6$ N/m. The vector of uncertain parameters is $\boldsymbol{\theta} = [\theta_1, \theta_2]$.

3.2 Definition of the prior

Motivated by the positive nature of the stiffness parameters the prior distribution of $\boldsymbol{\theta}$ is assumed to be given by the product of two log-normal PDFs with modes $\mathrm{Mo}_{\theta_1} = 1.3$ and $\mathrm{Mo}_{\theta_2} = 0.8$, and standard deviations $\bar{\sigma}_{\theta_1} = \bar{\sigma}_{\theta_2} = 1$ [1].

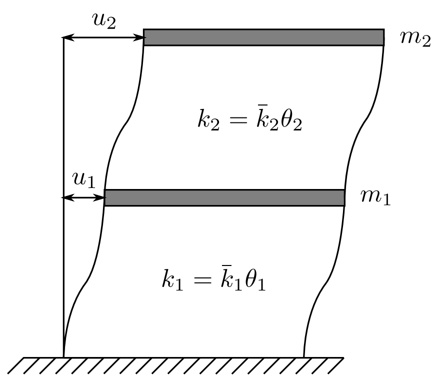

Figure 2: Shear building model (without damping).

The mode and variance of the log-normal PDF are equal to

$$\mathrm{Mo} = \exp(\mu - \sigma^2) \qquad \Longrightarrow \qquad \sigma^2 = \mu - \ln(\mathrm{Mo}) \tag{3}$$

$$\bar{\sigma}^2 = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1). \tag{4}$$

By substituting Eq. 3 into Eq. 4, the following equations can be solved to obtain the parameter μ for each distribution

$$\bar{\sigma}^2_{\theta_1} - \exp(2\mu_{\theta_1} + [\mu_{\theta_1} - \ln(\mathrm{Mo}_{\theta_1})])(\exp(\mu_{\theta_1} - \ln(\mathrm{Mo}_{\theta_1})) - 1) = 0$$

$$\bar{\sigma}^2_{\theta_2} - \exp(2\mu_{\theta_2} + [\mu_{\theta_2} - \ln(\mathrm{Mo}_{\theta_2})])(\exp(\mu_{\theta_2} - \ln(\mathrm{Mo}_{\theta_2})) - 1) = 0,$$

which result into $\mu_{\theta_1} = 0.51024$ and $\mu_{\theta_2} = 0.16958$. Now, using the mean and mode values in Eq. 3, we obtain the standard deviations of the underlying Gaussian $\sigma_{\theta_1} = 0.49787$ and $\sigma_{\theta_2} = 0.62667$.

These are the parameters of the log-normal distributions, and therefore, the prior PDF is defined as

$$f(\boldsymbol{\theta}) = \prod_{j=1}^{2} \frac{1}{\theta_j \sigma_{\theta_j} \sqrt{2\pi}} \exp\left(-\frac{(\ln(\theta_j) - \mu_{\theta_j})^2}{2\sigma^2_{\theta_j}} \right). \tag{5}$$

3.3 Definition of the likelihood

Using modal data $\mathcal{D}$ corresponding to measured eigenfrequencies, the likelihood function for the stiffness parameters $\boldsymbol{\theta}$ can be formulated as [1]

$$f(\mathcal{D}|\boldsymbol{\theta}) = \exp\left(\frac{-J(\boldsymbol{\theta})}{2\sigma^2_{\varepsilon}} \right), \tag{6}$$

where $J(\boldsymbol{\theta})$ is the modal measure-of-fit function given by [6]

$$J(\boldsymbol{\theta}) = \sum_{j=1}^{2} \mu^2_{\varepsilon_j} \left(\frac{f_j^2(\boldsymbol{\theta})}{\tilde{f}_j^2} - 1 \right)^2, \tag{7}$$

here $\mu_{\varepsilon_1} = \mu_{\varepsilon_2} = 1$ are the means and $\sigma^2_{\varepsilon} = 1/2^{i-1}$ the variance of the prediction error (for a given $i = 1, ..., 9$ simulation level to explore the effect of different values of σ^2_{ε} on the posterior). Furthermore, the $f_j(\boldsymbol{\theta})$ are the natural eigenfrequencies obtained with the model and the $\mathcal{D} = \tilde{f}_j$ are the natural eigenfrequencies used as the data in the model updating, these values are $\tilde{f}_1 = 3.13$ and $\tilde{f}_2 = 9.83$ Hz.

The equation of motion for un-damped free vibration encountered in structural engineering is expressed as

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{0},$$

where the mass matrix $\mathbf{M}$ and stiffness matrix $\mathbf{K}$ are given by:

$$\mathbf{M} = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \qquad \mathbf{K} = \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{bmatrix}.$$

This equation can be solved by assuming a harmonic solution of the form:

$$\mathbf{u} = \mathbf{\Phi}_j \sin(\omega_j t),$$

where ω_j are the circular natural eigenfrequencies and $\mathbf{\Phi}_j$ are the eigenvectors or modal shapes. Performing differentiation of the harmonic solution and substituting into the equation of motion yields

$$-\omega_j^2 \mathbf{M}\mathbf{\Phi}_j \sin(\omega_j t) + \mathbf{K}\mathbf{\Phi}_j \sin(\omega_j t) = \mathbf{0},$$

which is equivalent to

$$\left(\mathbf{K} - \omega_j^2 \mathbf{M}\right) \mathbf{\Phi}_j = \mathbf{0}, \qquad \text{for } j = 1, 2.$$

This equation is known as the eigenvalue problem of un-damped linear systems, which is a set of homogeneous algebraic equations. The basic form of this eigenvalue problem can be expressed as:

$$\mathbf{K}\mathbf{\Phi}_j = \omega_j^2 \mathbf{M}\mathbf{\Phi}_j,$$

which is solved for ω_j^2 and $\mathbf{\Phi}_j$. Finally, the corresponding natural eigenfrequencies in Hz are obtained as $f_j = \omega_j/(2\pi)$. Note that our problem only depends on the stiffness matrix since the mass matrix is assumed deterministic.

Combining Eqs. 6 and 7, the likelihood function is equal to

$$f(\mathcal{D}|\boldsymbol{\theta}) = \exp\left(-\sum_{j=1}^{2} \frac{\mu_{\varepsilon_j}^2}{2\sigma_\varepsilon^2}\left(\frac{f_j^2(\boldsymbol{\theta})}{\tilde{f}_j^2} - 1\right)^2\right). \tag{8}$$

3.4 Definition of the posterior

From Eq. 1, the posterior distribution can be expressed as

$$f(\boldsymbol{\theta}|\mathcal{D}) = Z_{\mathcal{D}}^{-1} f(\mathcal{D}|\boldsymbol{\theta})\, f(\boldsymbol{\theta})$$

where $Z_{\mathcal{D}} = f(\mathcal{D}) = \int f(\mathcal{D}|\boldsymbol{\theta})\, f(\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta}$ is a normalizing constant known as the model evidence or marginal likelihood. Since the aim is to draw samples from the posterior distribution, only its shape is needed [4, 7]. This particularity allows us to apply several Bayesian inference algorithms only using our knowledge about the prior and likelihood functions, i.e., $f(\boldsymbol{\theta}|\mathcal{D}) \propto f(\mathcal{D}|\boldsymbol{\theta})\, f(\boldsymbol{\theta})$.

Hence, using Eqs. 5 and 8, the target distribution to be used in the sampling process for the updating of the posterior distribution is given by

$$f(\boldsymbol{\theta}|\mathcal{D}) \propto \pi(\boldsymbol{\theta}) = \exp\left(-\sum_{j=1}^{2} \frac{\mu_{\varepsilon_j}^2}{2\sigma_\varepsilon^2}\left[\frac{f_j^2(\boldsymbol{\theta})}{\tilde{f}_j^2} - 1\right]^2\right) \prod_{j=1}^{2} \frac{1}{\theta_j \sigma_{\theta_j}\sqrt{2\pi}} \exp\left(-\frac{(\ln(\theta_j) - \mu_{\theta_j})^2}{2\sigma_{\theta_j}^2}\right). \tag{9}$$

Posterior samples of (9) are shown in Figure 3.

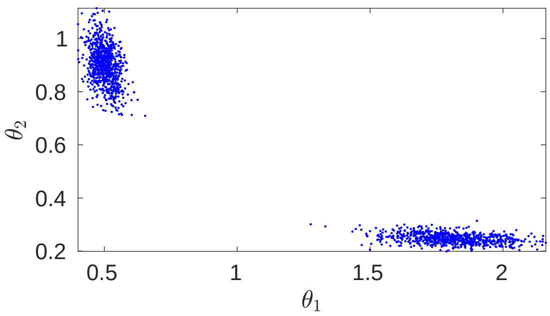

Figure 3: Samples from the posterior distribution in Example 3.

References

[1] Beck, J. L. and S.-K. Au (2002) - "Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation". *Journal of Engineering Mechanics*, 128(4), 380-391.

[2] Beck, J. L. and K. M. Zuev (2013) - "Asymptotically independent Markov sampling: A new Markov chain Monte Carlo scheme for Bayesian inference". *International Journal for Uncertainty Quantification*, 3(5), 445-474.

[3] Betz, W., I. Papaioannou, J. L. Beck and D. Straub (2018) - "Bayesian inference with subset simulation: Strategies and improvements". *Computer Methods in Applied Mechanics and Engineering*, 331, 72-93.

[4] Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin (2013) - *Bayesian Data Analysis*. Chapman and Hall/CRC, Third Edition.

[5] Straub, D. and I. Papaioannou (2015) - "Bayesian updating with structural reliability methods". *Journal of Engineering Mechanics*, 141(3), 04014134.

[6] Vanik, M. W., J. L. Beck and S. K. Au (2000) - "Bayesian probabilistic approach to structural health monitoring". *Journal of Engineering Mechanics*, 126(7), 738-745.

[7] Walsh, B. (2004) - *Markov chain Monte Carlo and Gibbs sampling*. Lecture notes for EEB 581, version 26.