Report on Bayesian ETAS

Suhas Shastry

June 9, 2018

Backgroud

Earthquakes: Aftershock and Mainshock

For a given earthquake catalog, earthquakes are categorized into aftershocks and mainshocks. Mainshock is the one with higher magnitude and is an early event. This mainshock causes unrest in the region and triggers several aftershocks of lesser magnitude. If an aftershock is of more magnitude than mainshock, then the aftershock is categorized as mainshock and the mainshock as foreshock. For simplicity reason, foreshock category is dropped.

Hawkes Process

Earliest models of earthquakes considered time-space Poisson process. As mainshocks increases the chance of aftershock, independence assumption of Poisson process is violated. Hence a modified version of Poisson process called Hawkes Process is more suitable for earthquakes modeling. Hawkes process a point process in which, background events are independent Poisson events. These background events trigger or excite few more events with diminishing rate. Hawkes process is depicted in Figure 1. This process can be written mathematically as

$$\lambda^*(t) = \lambda + \sum_{t_i<t} \mu(t - t_i)$$

Here $\lambda^*(t)$ is the overall rate of the event, λ is the background rate and $\mu(.)$ is excitation rate.

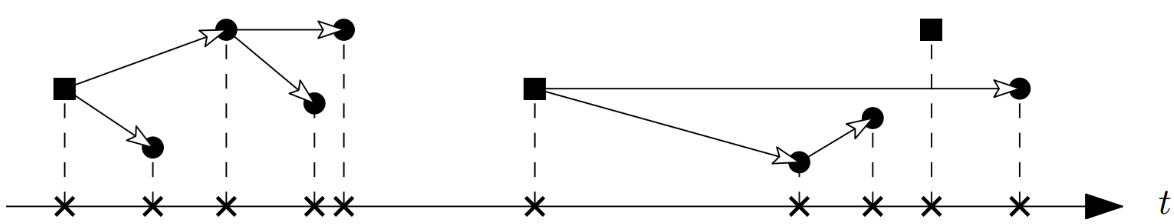

Figure 1: Hawkes process. Box represents background event and dot represents excited event.

ETAS Model

De-clustering is a process of identifying and separating independent mainshocks from dependent aftershocks. *Ogata* (1988) applied Hawkes process to de-cluster earthquake catalog by extending Omori law and Gutenberg-Richter law.

Omori Law

This is an empirical law proposed by Fusakichi Omori in 1894. The law states that frequency of an aftershock decreases roughly with the reciprocal of time after mainshock. Modified version of Omori law proposed by Utsu is most commonly used. This can be written mathematically as

$$n(t) = \frac{(p-1)c^{p-1}}{(t+c)^p}$$

Here p and c are appropriate constants.

Gutenberg-Richter law

This is also an empirical formula which refers to the relationship between the magnitude and total number of earthquakes in a region in a given period. Mathematically,

$$N = 10^{a-bM}$$

N is the number of events with magnitude greater than M and a and b are constants.

Ogata's ETAS is a time-space Hawkes process in which excitation process is described by Omori law and Gutenberg-Richter law. For ease of work in Bayesian context, time-space model is simplified to temporal ETAS model. Mathematically temporal ETAS is given by,

$$\lambda(t|H_t) = \mu + \sum_{t_i<t} \kappa(m_i|K,\alpha)h(t_i|c,p)$$

Where, $m_i > M_0$; $\kappa(m_i|K,\alpha) = Ke^{\alpha(m_i-M_0)}$ and

$$h(t_i|c,p) = \frac{(p-1)c^{p-1}}{(t-t_i+c)^p}$$

Parameters considered in the above model are $\theta = (\mu, K, \alpha, c, p)$. Data $Y = (Y_1, Y_2, \ldots, Y_n)$ where $Y_i = (t_i, m_i)$ is an earthquake event from the catalog. Estimates for these parameters are obtained using classical statistics. Estimator $\hat{\theta}$ is obtained using Maximum Likelihood Estimation using nonlinear optimization procedure such as the Davidon-Fletcher-Powell algorithm. Motivation for using Bayesian statistics are that,

a. Estimated $\hat{\theta}$ will not be equal to θ.

b. $SE(\hat{\theta})$ is very difficult to calculate.

Bayesian ETAS

Likelihood

Likelihood of the ETAS model can be re-written as,

$$p(Y|\theta) = \prod_{i=1}^{n} \lambda(t_i|H_t,\theta)e^{-\int_0^\infty \lambda(z|\theta,Hz)dz}$$

Where H_t represents all the events occurred before time t. Logarithmic version of likelihood is given by,

$$\log p(Y|\theta) = \sum_{i=1}^{n}\left(\mu + \sum_{j=1}^{i-1} \frac{Ke^{\alpha(m_i-M_0)(p-1)c^{p-1}}}{(t-t_i+c)^p}\right) - \mu T - \sum_{i=1}^{n} Ke^{\alpha(m_i-M_0)}\left(1 - \frac{c^{p-1}}{(T-t_i+c)^{p-1}}\right)$$

Posterior is very difficult to obtain from this likelihood equation because the likelihood function is multi-modal and the components of θ are highly correlated. Since MCMC methods can also suffer from serious convergence issues, latent variable for each earthquake event is introduced to make sure elements of θ are independent.

$$B_i = \begin{cases} 0 & If\ it\ is\ a\ mainshock \\ 1 & If\ it\ is\ an\ aftershock\ of\ Y_j \end{cases}$$

These latent variables are modified into sets so that they can be used in likelihood equation.

$$S_j = \{t_i; B_i = j\}, 0 \leq j\ < n$$

S_0 is a set of all mainshock events. S_1 is a set of aftershock events whose mainshock is Y_1. $|S_0|$ denotes the number of elements in set S_0. By introducing these latent variables, log-likelihood equation becomes

$$p(Y|\theta, B) = e^{-\mu T}\mu^{|S_0|}\prod_{i=1}^{n}\left(e^{-\kappa(m_j|K,\alpha)H(T-t_j|c,p)}\kappa(m_i|K,\alpha)^{|S_j|}\prod_{t_i\in S_j}h(t_i-t_j|c,p)\right)$$

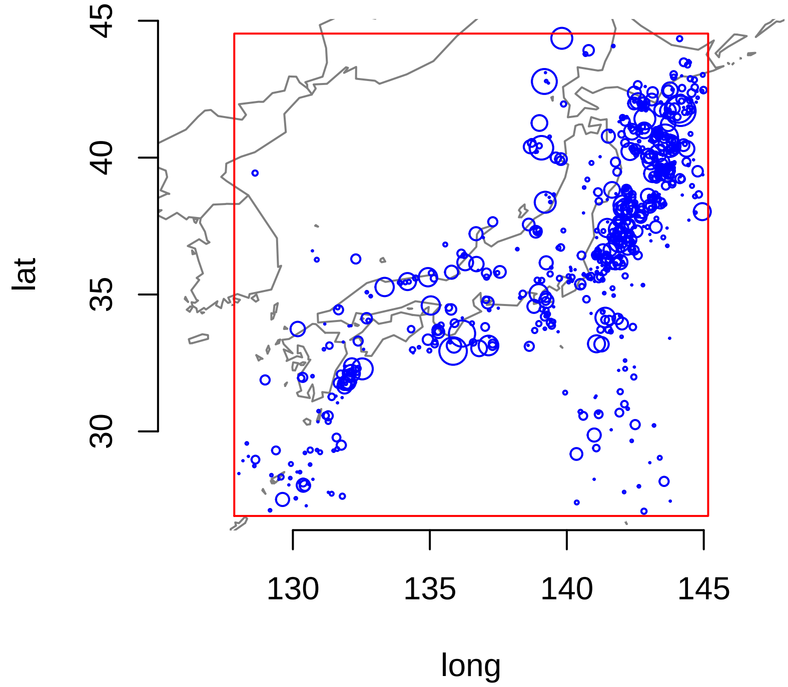

Figure 2: Earthquakes with M greater or equal to 6.0 in Japan region.

Prior

Non-informative priors are used for all the parameters. For B_is, discrete Uniform, for μ, conjugate Gamma prior and for K, α, c and p, Uniform priors are used. By using these priors, posterior distribution is obtained by Bayes rule. Analytical obtaining these distributions is very hard or even impossible, hence posterior distribution is obtained by simulation using random walk Monte Carlo Markov Chain which uses Metropolis-Hastings algorithm.

$$\pi(B_i) = \frac{1}{j}$$

$$\pi(\mu) = Gamma(0.1, 0.1)$$

$$\pi(\log K) \propto Uniform$$

$$\pi(\log \alpha) \propto Uniform$$

$$\pi(c) = U(0, 8)$$

$$\pi(p) = U(0, 8)$$

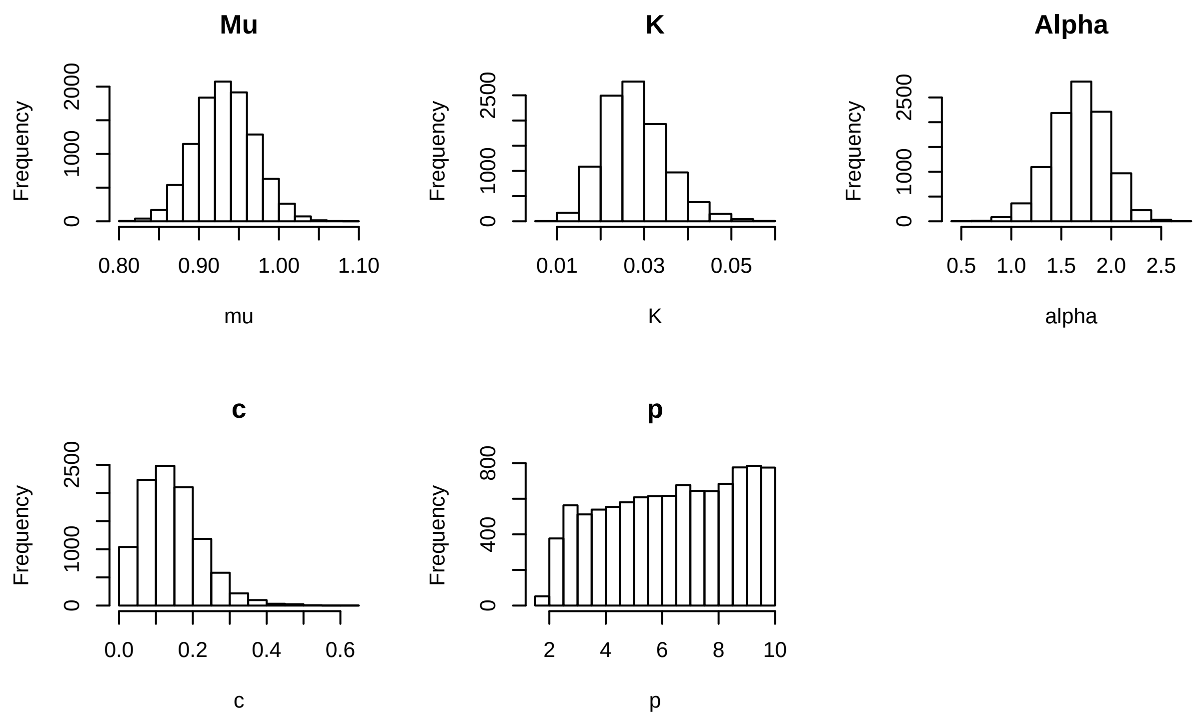

Figure 3: Posterior distribution of all the parameters

Simulation

Japan Catalog

Earthquakes happened in Japan from 1926 to 2007 was considered for simulation. Only those earthquakes with magnitude greater than or equal to 6.0 was considered. Figure 2 is the image of all those earthquakes. There were 701 such quakes. Simulation was done with the help of `bayesianETAS` R package. This internally uses MCMC for posterior distribution simulation. This simulation is very sensitive to initial values of θ. Hence θ_{MLE} obtained from classical ETAS was used as initial values.

$$\theta_{init} = (\mu = 0.012, K = 0.55, \alpha = 1.74, c = 0.137, p = 1.034)$$

Posterior

For 10000 simulations with 500 burn-in iterations, distribution obtained is plotted in Figure 3. Five parameters have stabilized after 500 burn-in iterations. This can be seen in Figure 4.

Limitations and Way forward

1. Calculation of log-likelihood has two nested summations. Hence simulation takes $\mathcal{O}(n^2)$ time for likelihood estimation. This is a serious limitation for catalog with more than 1000 earthquakes.

2. ETAS algorithm is applicable only for region which is of shape of a convex polygon.

3. Way ahead, this should be extended to catalog of whole sphere to get a holistic idea on earthquakes.

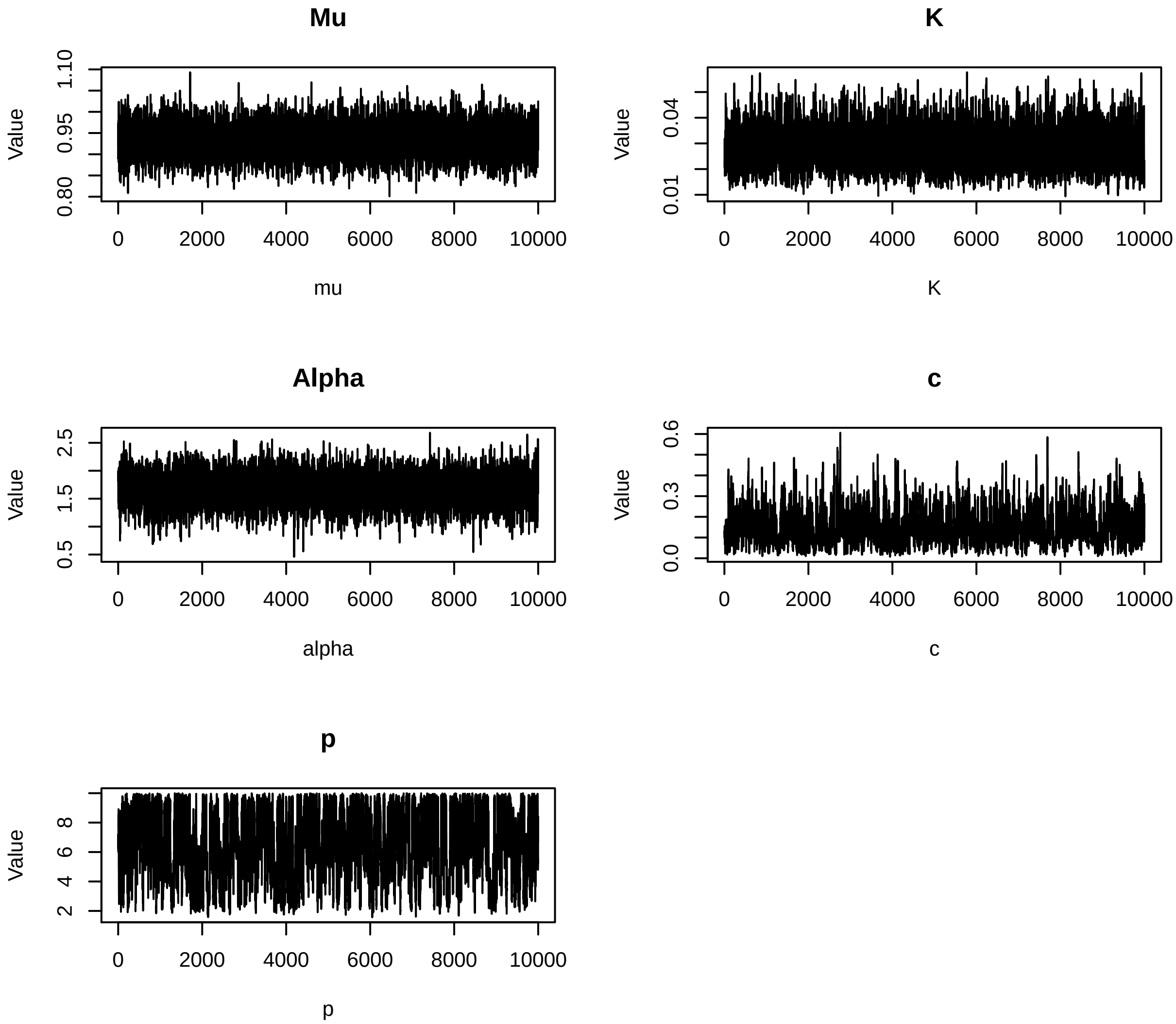

Figure 4: Trace of all parameters

4. Bayesian ETAS model considered was only a temporal model. Interpretation using temporal model is difficult and it cannot give whole picture of de-clustering. Hence time-space model should be considered in future.

References

1. Y. Ogata. Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association, 83(401):9-27, 1988

2. J. Zhuang, Y. Ogata, and D. Vere-Jones. Stochastic declustering of space-time earthquake occurrences. Journal of the American Statistical Association, 97(458):369-380, 2002

3. Gordon J Ross - Bayesian ETAS