Epidemic-Type Aftershock Sequence model approximation with Inlabru

Francesco Serafini, Mark Naylor, Finn Lindgren

10/01/2022

Contents

1 Definition of ETAS

We use the Inlabru R-package to approximate a specific form of *Epidemic-type Aftershock Sequence* (ETAS, Ogata 1988, 2011; Ogata and Zhuang 2006) model. The ETAS model is a self-exciting process (or Hawkes process, Hawkes 1971a, 1971b; Laub, Lee, and Taimre 2021) model, which is a point process model with conditional intensity, in its general form, given by:

$$\lambda(t, \mathbf{s}, m|\mathcal{H}_t) = \left(\mu + \sum_{h:t_h<t} K g_m(m_h) g_t(t-t_h) g_s(\mathbf{s}-\mathbf{s}_h)\right) \pi(m),$$

where $t \in (T_1, T_2)$ indicates the time, $\mathbf{s} \in W \subset \mathbb{R}^2$ is the 2D location, $m \in [M_0, M_c]$ is the magnitude, and $\mathcal{H}_t$ is the history of the process up to time t.

The functions $g_m(\cdot), g_t(\cdot), g_s(\cdot)$ are non-negative triggering functions representing the influence of past observations on the evaluation point. Specifically, $g_m(\cdot)$ accounts for the magnitude of past events and has to be an increasing function of m_h; $g_t(\cdot)$ accounts for the time passed between the past observation and the evaluation point and has to be a decreasing function of this quantity; $g_s(\cdot)$ accounts for the *distance* between the past observation's location and the evaluation point's location and has to be a decreasing function of this *distance*. The function $\pi(\cdot)$ is a density on the magnitude domain usually given by some form of Gutenberg-Richter law (GR-law Gutenberg and Richter 1956).

We choose specific forms of the triggering functions for reasons that will be explained in details in Section **??**. We use

$$g_m(m_h) = \exp\Big(\alpha(m_h - M_0)\Big),$$

with parameter $\alpha \geq 0$.

$$g_t(t-t_h) = \left(\frac{t-t_h}{c}+1\right)^{-p},$$

with parameters $c>0, p>1$. It is a slightly modified version of the unnormalized Omoris's law (Omori 1894; Utsu, Ogata, et al. 1995).

$$g_s(\mathbf{s}-\mathbf{s}_h) = \det(2\pi\Sigma)^{-1/2}\exp\Big(-\frac{1}{2}(\mathbf{s}-\mathbf{s}_h)^T\Sigma^{-1}(\mathbf{s}-\mathbf{s}_h)\Big),$$

which is a Multivariate Gaussian density evaluated at $\mathbf{s}$ with mean $\mathbf{s}_h$ and covariance matrix Σ. This means that *Sigma* has to be a symmetric positive semi-definite matrix. The *distance* between past observations and evaluation point is the Mahalanobis distance (Mahalanobis 1936) given by:

$$d(\mathbf{s},\mathbf{s}_h) = (\mathbf{s}-\mathbf{s}_h)^T\Sigma^{-1}(\mathbf{s}-\mathbf{s}_h).$$

The density chosen for the magnitude domain is a form of GR-law and its parameters are estimated independently from the others. For this reason, we are going to consider it as given in the following steps.

1.1 Alternative Parametrization and Notation

To ensure all the constraints on the parameters hold (i.e., $\mu, \alpha, K, c > 0$ and $p > 1$, Σ valid covariance matrix) we consider a different parametrization:

- $\mu = \exp(\theta_1)$
- $K = \exp(\theta_2)$
- $\alpha = \exp(\theta_3)$
- $c = \exp(\theta_4)$
- $p - 1 = \exp(\theta_5)$
- $\Sigma = \text{Bart}(\theta_6, \theta_7, \theta_8)$

With parameters $\theta_1, ..., \theta_8 \in (-\infty, \infty)$

In this way, the magnitude and time triggering functions are given by

$$g_m(m_h) = \exp\Big(\exp(\theta_3)(m_h - M_0)\Big),$$

and

$$g_t(t-t_h) = \left(\frac{t-t_h}{\exp(\theta_4)}+1\right)^{-1-\exp(\theta_5)}.$$

The only parameter of the space triggering function is the covariance matrix

$$\Sigma = \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$$

which is defined by three quantities: the two marginal standard deviations $\sigma_1, \sigma_2 > 0$ and the correlation coefficient $-1 \leq \rho \leq 1$. In order to have free-constraints parameters, we are going to use the Bartlett decomposition of the matrix Σ for which

$$\Sigma = LAA^TL^T,$$

where, L and A are 2×2 matrices with values in $\mathbb{R}$.

Specifically, L is the Cholesky factor (Benoit 1924) of Σ, such that $\Sigma = LL^T$ and

$$L = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}$$

The matrix A is composed by

$$A = \begin{pmatrix} c_1 & 0 \\ n_{21} & c_2 \end{pmatrix}$$

where $c_i^2 \sim \chi^2_{n-i+1}$ and $n_{21} \sim N(0,1)$ independently. In practice, we provide a Cholesky decomposition using prior estimates of σ_1, σ_2 and ρ and consider $\theta_6, \theta_7, \theta_8 \sim N(0,1)$ such that $h(\theta_6, 2) = c_1$, $h(\theta_7, 1) = c_2$ and $\theta_8 = n_{21}$. The function $h(x,k)$ takes x having a standard normal distribution and returns values having a Chi-squared distribution with k degrees of freedom. The obtained Σ matrix is given by

$$\Sigma = \begin{pmatrix} \sigma_1^2 c_1^2 & \sigma_1\sigma_2\left(\rho c_1^2 + c_1 n_{21}\sqrt{1-\rho^2}\right) \\ \sigma_1\sigma_2\left(\rho c_1^2 + c_1 n_{21}\sqrt{1-\rho^2}\right) & \sigma_2^2\left((1-\rho^2)c_2^2 + \left(n_{21}\sqrt{1-\rho^2} + \rho c_1\right)^2\right) \end{pmatrix}$$

2 Inlabru Approximation

The Inlabru R-package (Bachl et al. 2019) provides an automated way to perform Bayesian analyses. It relies on the INLA algorithm (Rue, Martino, and Chopin 2009; Rue et al. 2017) which provides an alternative to Monte Carlo Markov Chain (MCMC, Robert, Casella, and Casella 1999) methods for estimating the posterior distributions of the parameters. Contrarily to MCMC, INLA is deterministic (MCMC is based on simulations), and easier to reproduce. Also, INLA is generally faster than MCMC methods and allows to consider complex spatio-temporal structures, which would be difficult to implement using an MCMC method. Furthermore, the optimization routine used to obtain the posterior of the parameters is automated, and does not need to be explicitly programmed by the end-user, who has the only duty to set up the components of the model. This makes Inlabru easy to programme and gives access to complex models to a greater number of users, who can focus mainly on the hypotheses incorporated in the model, rather than the algorithm to obtain the results. Also, being the same algorithm provides a common-ground to compare the performance of models based on different hypotheses.

We use the INLA algorithm through the Inlabru R-package to approximate the ETAS model. The key of our approach is to approximate the different components of the log-likelihood using three Poisson Counts model with log-intensity linear in the parameters. The parametrization seen in Section 1.1 potentially allows to consider the parameters as Gaussian processes which can be used to incorporate different hypotheses in the models. Possible examples are: parameters varying in time, parameters varying in space, parameters as functions of covariates, mixtures of the previous ones.

3 Priors

For now are considering Gaussian priors for all the parameters. However, we can use transformations of the parameters distributed according to a target density $\pi(\cdot)$. The transformation is composed by two transformations. Assuming X $N(0,1)$ and $F(x) = \Pr(X \leq x)$. The first transformation is $Z = F(X)$, the variable Z has a uniform distribution in $(0,1)$. The second transformation is $X' = F_\pi^{-1}(Z)$, where $F_\pi^{-1}(\cdot)$ is the inverse function of $F_\pi(\cdot)$ which is the distribution function associated with the target density $\pi(\cdot)$. The resulting values of X' are distributed according to $\pi(\cdot)$.

References

Bachl, Fabian E., Finn Lindgren, David L. Borchers, and Janine B. Illian. 2019. "inlabru: An R Package for Bayesian Spatial Modelling from Ecological Survey Data." *Methods in Ecology and Evolution* 10: 760–66. https://doi.org/10.1111/2041-210X.13168.

Benoit, Commandant. 1924. "Note Sur Une méthode de résolution Des équations Normales Provenant de l'application de La méthode Des Moindres Carrés à Un Système d'équations Linéaires En Nombre Inférieur à Celui Des Inconnues (Procédé Du Commandant Cholesky)." *Bulletin géodésique* 2 (1): 67–77.

Gutenberg, Beno, and Carl F Richter. 1956. "Earthquake Magnitude, Intensity, Energy, and Acceleration: (Second Paper)." *Bulletin of the Seismological Society of America* 46 (2): 105–45.

Hawkes, Alan G. 1971a. "Point Spectra of Some Mutually Exciting Point Processes." *Journal of the Royal Statistical Society: Series B (Methodological)* 33 (3): 438–43.

———. 1971b. "Spectra of Some Self-Exciting and Mutually Exciting Point Processes." *Biometrika* 58 (1): 83–90.

Laub, Patrick J, Young Lee, and Thomas Taimre. 2021. *The Elements of Hawkes Processes*. Springer Nature.

Mahalanobis, Prasanta Chandra. 1936. "On the Generalized Distance in Statistics." In. National Institute of Science of India.

Ogata, Yosihiko. 1988. "Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes." *Journal of the American Statistical Association* 83 (401): 9–27.

———. 2011. "Significant Improvements of the Space-Time ETAS Model for Forecasting of Accurate Baseline Seismicity." *Earth, Planets and Space* 63 (3): 217–29.

Ogata, Yosihiko, and Jiancang Zhuang. 2006. "Space-Time ETAS Models and an Improved Extension." *Tectonophysics* 413 (1-2): 13–23.

Omori, Fusakichi. 1894. *On the After-Shocks of Earthquakes*. Vol. 7. The University.

Robert, Christian P, George Casella, and George Casella. 1999. *Monte Carlo Statistical Methods*. Vol. 2. Springer.

Rue, Håvard, Sara Martino, and Nicolas Chopin. 2009. "Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations." *Journal of the Royal Statistical Society: Series b (Statistical Methodology)* 71 (2): 319–92.

Rue, Håvard, Andrea Riebler, Sigrunn H Sørbye, Janine B Illian, Daniel P Simpson, and Finn K Lindgren. 2017. "Bayesian Computing with INLA: A Review." *Annual Review of Statistics and Its Application* 4: 395–421.

Utsu, Tokuji, Yosihiko Ogata, et al. 1995. "The Centenary of the Omori Formula for a Decay Law of Aftershock Activity." *Journal of Physics of the Earth* 43 (1): 1–33.