State-space modeling of conflict data

Andreas Lindholm

May 4, 2020

1 Data

- y_t monthly (month index = t) death count in state-based conflict events for a certain gridpoint
- u_t^i various "exogenous inputs", such as (possibly nonlinearly transformed) population numbers, economical indicators, earlier death counts in neighboring grid points etc.

2 Model

$$
\begin{aligned}
x_{t+1} &= a x_t + c^1 u_t^1 + c^2 u_t^2 + \ldots c^p u_t^p + w_t, && w_t \sim \mathcal{N}\left(0, (1-a^2)(1-\sigma_e^2)\right) && \text{(1a)}\\
s_{t+1} &= b s_t + d^1 u_t^1 + d^2 u_t^2 + \ldots d^p u_t^p + v_t, && v_t \sim \mathcal{N}\left(0, 1-b^2\right) && \text{(1b)}\\
z_t &= Be\left(\beta\left[\exp(\exp(s_t-1)-1)-1\right]_+\right) && && \text{(1c)}\\
y_t &= z_t \cdot \text{round}_{\text{nni}}\left[\exp\left(\alpha(x_t+e_t)\right)-1\right] && e_t \sim \mathcal{N}\left(0, \sigma_e^2\right) && \text{(1d)}
\end{aligned}
$$

with

$$
\begin{aligned}
x_0 &\sim \mathcal{N}\left(0, 1-\sigma_e^2\right), && \text{(1e)}\\
s_0 &\sim \mathcal{N}\left(0, 1\right). && \text{(1f)}
\end{aligned}
$$

where

$$
\begin{aligned}
\text{round}_{\text{nni}}[\cdot] &: \text{ Round to nearest } \textbf{n}\text{on-}\textbf{n}\text{egative } \textbf{i}\text{nteger}\\
Be(\cdot) &: \text{ Bernoulli distribution}
\end{aligned}
$$

Here x_t and s_t are hidden states, whereas z_t is part of the observation model.

Note that with $c^1 = \ldots c^p = 0$ and $d^1 = \ldots d^p = 0$, the states (plus noise) have the marginal distributions $x_t + e_t \sim \mathcal{N}(0,1)$ and $s_t \sim \mathcal{N}(0,1)$.

3 Particle filter

A particle filter can, in principle, be applied out of the box to model (1). However, some care can be taken to gain computational efficiency:

- By approximating (1d) as $\frac{\log(y_t+1)}{\alpha} = x_t + e_t$ if $z_t \neq 0$ (that is, ignoring $\text{round}_{\text{nni}}$ and consider y_t continuous instead of discrete), we have a conditionally linear-Gaussian model when $y_t > 0$, which allows for a Rao-Blackwellized particle filter with $p(x_t \mid \ldots)$ handled analytically.

- To run a R-B particle filter also when $y_t = 0$, z_t needs (?) to be sampled. If $z_t = 0$, no "measurement update" is made, and if $z_t = 1$, then the observation is $x_t + e_t \leq \frac{\ln 1.5}{\alpha}$ made, which implies that $p(x_t + e_t \mid \dots)$ becomes a truncated Gaussian. This can be handled using moment matching ($m_t^{(i)}$ and $P_t^{(i)}$ is the predictive mean and variance for x_t at time t, and the version with tilde is the mean and variance after the "measurement update" with $z_t = 1$),

$$q = \frac{\frac{\ln 1.5}{\alpha} - m_t^{(i)}}{\sqrt{P_t^{(i)} + s_e^2}}, \quad \phi \text{ normal pdf, } \Phi \text{ normal cdf,}$$

$$\tilde{m}_t^{(i)} = m_t^{(i)} - \sqrt{P_t^{(i)} + s_e^2}\,\frac{\phi(q)}{\Phi(q)},$$

$$\tilde{P}_t^{(i)} = \max(P_t^{(i)}\left(1 - q\frac{\phi(q)}{\Phi(q)} - \left(\frac{\phi(q)}{\Phi(q)}\right)^2\right) - s_e^2, 0).$$

- The case $y_t = 0$ is treated as a discrete observation (which it is), whereas $y_t > 0$ is treated as a continuous observation (which it, strictly speaking, isn't). In order to correctly "compare" two particles i, j with $z_t^{(i)} = 0$ and $z_t^{(j)} = 1$ if $y_t = 0$ in the particle filter weighting, we have to make sure we use probabilities, and not densities, which means

$$p(y_t = 0 \mid z_t^{(i)} = 0) = 1$$

$$p(y_t = 0 \mid z_t^{(j)} = 1, m_t^{(j)}, P_t^{(j)}) = \Phi\left(\frac{\frac{\ln 1.5}{\alpha} - m_t^{(j)}}{\sqrt{P_t^{(j)} + \sigma_e^2}}\right)$$

4 Parameter inference

Ideally, all parameters should be estimated jointly. That has not been found feasible so far. The somewhat ad-hoc attempts made so far on the parameter estimation are

Parameter	Value	Estimation procedure
α	2.3	To have simulated data matching the observed histogram of y_t (independent of a and b)
β	0.005	To have simulated data matching the observed histogram of y_t (independent of a and b)
a	0.999*	Maximum likelihood (evaluated using particle filter; Figure 1) for a subset of the data with $c^1 = \cdots = c^p = 0$ and $d^1 = \cdots = d^p = 0$
b	0.995*	Maximum likelihood (evaluated using particle filter; Figure 1) for a subset of the data with $c^1 = \cdots = c^p = 0$ and $d^1 = \cdots = d^p = 0$
σ_e^2	0.27	Not systematically explored yet
$c_1, c_2, \dots, c^p$	Only very initial results	Maximum likelihood (using eg. PSAEM), conditional on previous parameters
$d_1, c_2, \dots, d^p$	Only very initial results	Maximum likelihood (using eg. PSAEM), conditional on previous parameters

*The partcile filter used in this maximum likelihood evaluation might have contained some errors. Should be checked.

5 Evaluation

It is not clear how to evaluate a model in a sound way. Always predicting zero scores quite well. . . Some thoughts:

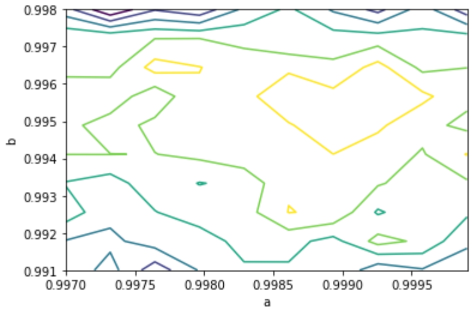

Figure 1: Maximum likelihood inference for a and b. Level curves for likelihood function, estimated using particle filter and data from a subset of 1000 grid points. Note: This figure might not have been made with a correct particle filter.

- A probabilistic precision/recall evaluation of the classes $y_t = 0$ (no lethal violence) and $y_t \neq 0$ (lethal violence)
- A further evaluation within $y_t \neq 0$ (lethal violence):

$$\frac{1}{T_{y_t \neq 0}} \sum_{y_t \neq 0} |\log(y_t + 1) - \log(\widehat{y}_t + 1)| \tag{2}$$

- The calibration is probably good (because of the design (1a)) but could be analyzed as well