A Gibbs Sampler to Find the Change Point in a Time Series of Count Data

Joshua P. French

January 3, 2014

In this example we will talk about a Bayesian approach for detecting the change point in a time series of counts.

Suppose that we have a times series of counts $y_1, y_2, \ldots, y_n$. Since the responses are counts, it is natural to model them using a Poisson distribution. We will suppose that there is a change point at some time during the series. Specifically, we'll assume that $y_1, y_2, \ldots, y_k \overset{i.i.d.}{\sim} \text{Poisson}(\lambda)$ and $y_{k+1}, y_{k+2}, \ldots, y_{112} \overset{i.i.d.}{\sim} \text{Poisson}(\phi)$. We denote the vector of responses by $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. For prior distributions, we assume $\lambda \sim \text{Gamma}(a, b)$, $\phi \sim \text{Gamma}(c, d)$, and that k has a discrete uniform distribution on the values $1, 2, \ldots, 112$.

We can approximate the posterior distributions for parameters λ, ϕ, and k using the Gibbs sampler. The full conditional distributions of the parameters are:

$$p(\lambda \mid \phi, k, \mathbf{y}) = \text{Gamma}\left(a + \sum_{i=1}^{k} y_i, k + b\right),$$

$$p(\phi \mid \lambda, k, \mathbf{y}) = \text{Gamma}\left(c + \sum_{i=k+1}^{n} y_i, n - k + d\right),$$

and

$$\begin{aligned} p(k \mid \lambda, \phi, \mathbf{y}) &= \frac{\exp\left\{k(\phi - \lambda) + \log(\lambda/\phi) \sum_{i=1}^{k} y_i\right\}}{\sum_{k=1}^{n} \exp\left\{k(\phi - \lambda) + \log(\lambda/\phi) \sum_{i=1}^{k} y_i\right\}} \\ &= \exp\left\{g(k, \lambda, \phi) - \ln\left(\sum_{k=1}^{n} g(k, \lambda, \phi)\right)\right\}, \end{aligned}$$

where $g(k, \lambda, \phi) = \exp\left\{k(\phi - \lambda) + (\log(\lambda/\phi) \sum_{i=1}^{k} y_i\right\}$.

After specifying some initial values for ϕ and k, as well as the parameters a, b, c, and d, we will first sample a λ value from its full conditional distribution. We will then use this value of λ and the current value of k to sample a new ϕ from its full conditional distribution. Lastly, we will use the value of λ and the new value of ϕ to sample a new value of k from its full conditional distribution. We then iterate through this procedure the desired number of times. The posterior distributions of the parameters are approximated by their respective samples from the Gibbs sampler algorithm.