A new spatial model for points above a threshold

April 22, 2014

1 Introduction

2 Statistical model

Let $Y_t(\mathbf{s}) \in \mathcal{R}$ be the observed value at location $\mathbf{s}$ on day t. To avoid bias in estimating tail parameters, we model the thresholded data

$$\tilde{Y}_t(\mathbf{s}) = \begin{cases} Y_t(\mathbf{s}) & Y_t(\mathbf{s}) > T \\ T & Y_t(\mathbf{s}) \leq T \end{cases} \tag{1}$$

where T is a pre-specified threshold.

We first specify a model for the complete data, $Y_t(\mathbf{s})$, and then study the induced model for thresholded data, $\tilde{Y}_t(\mathbf{s})$. The full data model is given in Section 2.2 assuming a skew normal distribution with a different variance each day. Computationally, the values below the threshold are updated using standard Bayesian missing data methods as described in Section 3. The skew normal representation is from (Minozzo and Ferracuti, 2012) and is the sum of a normal and half-normal random variable.

2.1 Half-normal distribution

Let $u = |z|$ where $Z \sim N(\mu, \sigma^2)$. Specifically, we consider the case where $\mu = 0$. Then U follows a half-normal distribution which we denote as $U \sim HN(0, 1)$, and the density is given by

$$f_U(u) = \frac{\sqrt{2}}{\sqrt{\pi\sigma^2}} \exp\left(-\frac{u^2}{2\sigma^2}\right) I(u > 0) \tag{2}$$

When $\mu = 0$, the half-normal distribution is also equivalent to a $N_{(0,\infty)}(0, \sigma^2)$ where $N_{(a,b)}(\mu, \sigma^2)$ represents a normal distribution with mean μ and standard deviation σ that has been truncated below at a and above at b.

2.2 Complete data

Consider the spatial process

$$Y_t(\mathbf{s}) = X_t(\mathbf{s})\beta + e_t(\mathbf{s}) \tag{3}$$
$$e_t(\mathbf{s}) = \delta z_t(\mathbf{s}) + v_t(\mathbf{s}) \tag{4}$$

where $z_t(\mathbf{s}) = z_{tl}$ if $s \in P_{tl}$ where $P_{t1}, \ldots, P_{tL}$ form a partition, and $z_{tl} \stackrel{iid}{\sim} N_{(0,\infty)}(0, \sigma_t^2)$, $\delta \in (-1, 1)$ controls skew, and $v_t(\mathbf{s})$ is a spatial Gaussian process with mean zero and variance $\sigma_t^2(1 - \delta^2)$.

Then $Y_t(\mathbf{s})$ is skew normal within each partition (Minozzo and Ferracuti, 2012). We model this with a Bayesian hierarchical model as follows. Let $w_{t1}, \ldots, w_{tL}$ be partition centers so that

$$P_{tl} = \{\mathbf{s}_t : l = \arg\min_k ||\mathbf{s}_t - w_{tk}||\}.$$

Then

$$Y_t(\mathbf{s}) \mid \Theta, z_{t1}, \dots, z_{tL} = X_t(\mathbf{s})\beta + \delta z_t(\mathbf{s}) + v_t(\mathbf{s}) \tag{5}$$
$$z_{tl}(\mathbf{s}) \mid \Theta \sim N_{(0,\infty)}(0, \sigma_t^2) \tag{6}$$
$$v_t(\mathbf{s}) \mid \Theta \sim \text{Matérn } (0, \Sigma) \tag{7}$$
$$\sigma \sim IG(\alpha, \beta) \tag{8}$$
$$\delta \sim Unif(-1, 1) \tag{9}$$
$$w_{tk} \sim Unif(\mathcal{D}) \tag{10}$$

where $\Theta = \{w_{t1}, \dots, w_{tL}, \beta, \sigma_t, \delta, \rho, \nu\}$; $l = \arg\min_k ||\mathbf{s} - w_k||$; Σ_t is a Matérn covariance matrix with variance $\sigma_t^2(1 - \delta^2)$, spatial range ρ and smoothness ν; and $\mathcal{D}$ is the spatial domain of interest.

3 Computation

The MCMC for this model is fairly straightforward. First, we impute values below the threshold. Then, we update Θ using random walk MH or Gibbs sampling when appropriate. Finally, we make spatial predictions. Each requires the joint distribution for the complete data given Θ. As defined in 5, the distribution of $Y_t(\mathbf{s}) \mid \Theta$ is the usual multivariate normal distribution with a Matérn spatial covariance structure.

3.1 Imputation

We can use Gibbs sampling to update $\tilde{Y}_t(\mathbf{s})$ for observations that are below T, the thresholded value. Given Θ, $Y_t(\mathbf{s})$ has truncated normal full conditional with these parameter values. So we sample $Y_t(\mathbf{s}) \sim N_{(-\infty,T)}$

3.2 Parameter updates

To update Θ given the current value of the complete data $\mathbf{Y}_1, \dots, \mathbf{Y}_T$, we use a standard Gibbs updates for all parameters except for the knot locations which are done using a Metropolis update. See Appendix A.1 for details regarding Gibbs sampling and $|u_t(\mathbf{s})|$.

3.3 Spatial prediction

Given $\mathbf{Y}_t$ the usual Kriging equations give the predictive distribution for $Y_t(\mathbf{s}^*)$ at prediction location $(\mathbf{s}^*)$

4 Data analysis

5 Conclusions

Acknowledgments

Appendix A.1: Posterior distributions

Conditional posterior of $z_{tl} \mid \ldots$

For simplicity, drop the subscript t and define

$$R(\mathbf{s}) = \begin{cases} Y(\mathbf{s}) - X(\mathbf{s})\beta & s \in P_l \\ Y(\mathbf{s}) - X(\mathbf{s})\beta - \delta z(\mathbf{s}) & s \notin P_l \end{cases}$$

Let

$$\begin{aligned} R_1 &= \text{the vector of } R(\mathbf{s}) \text{ for } s \in P_l \\ R_2 &= \text{the vector of } R(\mathbf{s}) \text{ for } s \notin P_l \\ \Omega &= \Sigma^{-1}. \end{aligned}$$

Then

$$\begin{aligned} \pi(z_l|\ldots) &\propto \exp\left\{-\frac{1}{2\sigma^2}\left[\frac{1}{(1+\delta^2)}\begin{pmatrix} R_1 - \delta z_l \mathbf{1} \\ R_2 \end{pmatrix}^T \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \begin{pmatrix} R_1 - \delta z_l \mathbf{1} \\ R_2 \end{pmatrix} + z_l^2\right]\right\} I(z_l > 0) \\ &\propto \exp\left\{-\frac{1}{2}\left[\Lambda_z z_l^2 - 2\mu_z z_l\right]\right\} I(z_l > 0) \end{aligned}$$

where

$$\begin{aligned} \mu_z &= \frac{\delta}{(1-\delta^2)}(R_1^T \Omega_{11} + R_2^T \Omega_{21})\mathbf{1} \\ \Lambda_z &= \frac{\delta^2 \mathbf{1}^T \Omega_{11} \mathbf{1}}{\sigma^2(1-\delta^2)} + \frac{1}{\sigma^2}. \end{aligned}$$

Then $Z_l|\ldots \sim N_{(0,\infty)}(\Lambda_z^{-1}\mu_z, \Lambda_z^{-1})$

Conditional posterior of $\beta \mid \ldots$

Let $\beta \sim \mathrm{N}_p(0, \Lambda_0)$ where Λ_0 is a precision matrix. Then

$$\begin{aligned} \pi(\beta \mid \ldots) &\propto \exp\left\{-\frac{1}{2}\beta^T \Lambda_0 \beta - \sum_{t=1}^{T} \frac{1}{2\sigma_t^2(1-\delta)^2}[\mathbf{Y}_t - X_t\beta - \delta z_t]^T \Sigma^{-1}[\mathbf{Y}_t - X_t\beta - \delta z_t]\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left[\beta^T \Lambda_\beta \beta - 2\sum_{t=1}^{T} \frac{1}{\sigma_t^2(1-\delta^2)}[\beta^T X_t \Sigma^{-1}(\mathbf{Y}_t + z_t)]\right]\right\} \\ &\propto \mathrm{N}(\Lambda_\beta^{-1}\mu_\beta, \Lambda_\beta^{-1}) \end{aligned}$$

where

$$\mu_\beta = \sum_{t=1}^{T} \frac{1}{(1-\delta)^2} \left[X_t^T \Sigma^{-1} (\mathbf{Y}_t + \delta z_t) \right]$$

$$\Lambda_\beta = \Lambda_0 + \sum_{t=1}^{T} \frac{1}{\sigma^2(1-\delta^2)} X_t^T \Sigma^{-1} X_t.$$

Conditional posterior of $\sigma^2 \mid \ldots$

Let $\sigma_t^2 \overset{iid}{\sim} \mathrm{IG}(\alpha_0, \beta_0)$. For simplicity, drop the subscript t. Then

$$\begin{aligned}
\pi(\sigma^2 \mid \ldots) &\propto (\sigma^2)^{-\alpha_0 - L/2 - n/2 - 1} \exp\left\{ -\frac{\beta_0}{\sigma^2} - \sum_{l=1}^{L} \frac{z_l^2}{2\sigma^2} - \frac{(Y(\mathbf{s}) - \boldsymbol{\mu})^T \Sigma^{-1} (Y(\mathbf{s}) - \boldsymbol{\mu})}{2\sigma^2(1-\delta^2)} \right\} \\
&\propto (\sigma^2)^{-\alpha_0 - L/2 - n/2 - 1} \exp\left\{ -\frac{1}{\sigma^2} \left[\beta_0 + \sum_{l=1}^{L} \frac{z_l^2}{2} + \frac{(Y(\mathbf{s}) - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{Y} - \mu)}{2(1-\delta^2)} \right] \right\} \\
&\propto \mathrm{IG}(\alpha^*, \beta^*)
\end{aligned}$$

where

$$\alpha^* = \alpha_0 + \frac{L}{2} + \frac{n}{2}$$

$$\beta^* = \beta_0 + \sum_{l=1}^{L} \frac{z_l^2}{2} + \frac{(Y(\mathbf{s}) - \boldsymbol{\mu})^T \Sigma^{-1} (Y(\mathbf{s}) - \boldsymbol{\mu})}{2(1-\delta^2)}$$

and L is the number of partitions.

Appendix A.2: MCMC Details

Priors

References

Minozzo, M. and Ferracuti, L. (2012) On the existence of some skew-normal stationary processes. *Chilean Journal of Statistics (ChJS)*, **3**, 157–170.