A new spatial model for points above a threshold

June 26, 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

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

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 a skew Gaussian spatial process

$$Y_t(\mathbf{s}) = X_t(\mathbf{s})\beta + \sigma_1 z_t(\mathbf{s}) + v_t(\mathbf{s}) \tag{3}$$

where $z_t(\mathbf{s}) = z_{tl}$ if $s \in P_{tl}$ where $P_{t1}, \dots, P_{tL}$ form a partition, and $z_{tl} \overset{iid}{\sim} N_{(0,\infty)}(0, 1)$, $\sigma_1 \in \mathcal{R}$, $\sigma_2, \sigma_0 \in \mathcal{R}^+$, and $v_t(\mathbf{s})$ is a spatial Gaussian process with mean zero and variance σ^2_{tl}. It can be shown (Zhang and El-Shaarawi, 2010) that $Y_t(\mathbf{s})$ follows a skew normal distribution with skewness parameter $\alpha = \frac{\sigma_1}{\sigma_{tl}}$. We can then reexpress the model in (3) as

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

where $z_t(\mathbf{s}) = z_{tl}$, $z_{tl} \overset{iid}{\sim} N_{(0,\infty)}(0, \sigma^2_{tl})$, and $v_t(\mathbf{s})$ is defined as before.

We model this with a Bayesian hierarchical model as follows. Let $w_{t1}, \dots, w_{tL}$ be partition centers so that

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

Then

$$
\begin{aligned}
Y_t(\mathbf{s}) \mid \Theta, z_{t1}, \ldots, z_{tL} &= X_t(\mathbf{s})\beta + \alpha z_t(\mathbf{s}) + v_t(\mathbf{s}) && (5) \\
z_{tl}(\mathbf{s}) \mid \Theta &\sim N_{(0,\infty)}(0, \sigma^2_{tl}) && (6) \\
v_t(\mathbf{s}) \mid \Theta &\sim \text{Matérn } (0, \Sigma) && (7) \\
\sigma^2_{tl} &\overset{iid}{\sim} IG(\alpha, \beta) && (8) \\
\alpha &\sim N(0, 10) && (9) \\
w_{tk} &\sim Unif(\mathcal{D}) && (10)
\end{aligned}
$$

where $\Theta = \{w_{t1}, \ldots, w_{tL}, \beta, \sigma_t, \delta, \rho, \nu\}$; $l = \arg\min_k ||\mathbf{s} - w_k||$; Σ_t is a Matérn covariance matrix with variance $\sigma^2_{tl}(1 - \delta^2)$, 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 $\mathbf{Y}_1, \ldots, \mathbf{Y}_T$, 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.

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

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

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

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

where

$$\begin{aligned} \mu_l &= (R_1^T \Omega_{11} + R_2^T \Omega_{21})\mathbf{1} \\ \Lambda_l &= \alpha^2 \mathbf{1}^T \Omega_{11} \mathbf{1} + \frac{1}{\sigma_l^2}. \end{aligned}$$

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

Conditional posterior of $\beta \mid \ldots$

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

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

where

$$\mu_\beta = \sum_{t=1}^{T} \left[X_t^T \Sigma^{-1} (\mathbf{Y}_t - \alpha z_t) \right]$$
$$\Lambda_\beta = \Lambda_0 + \sum_{t=1}^{T} X_t^T \Sigma^{-1} X_t.$$

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

In the case where $L = 1$, then σ^2 has a conjugate posterior distribution. Let $\sigma_t^2 \overset{iid}{\sim} \mathrm{IG}(\alpha_0, \beta_0)$. For simplicity, drop the subscript t. Then

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

where

$$\alpha^* = \alpha_0 + \frac{1}{2} + \frac{n}{2}$$
$$\beta^* = \beta_0 + \frac{z^2}{2} + \frac{1}{2} (\mathbf{Y} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{Y} - \boldsymbol{\mu}).$$

In the case that $L > 1$, a random walk Metropolis Hastings step will be used to update σ_{lt}^2.

Conditional posterior of $\alpha \mid \ldots$

Let $\alpha \sim N(0, \tau_\alpha)$ where τ_α is a precision term. Then

$$\begin{aligned}
\pi(\alpha \mid \ldots) &\propto \exp\left\{ -\frac{1}{2} \tau_\alpha \alpha^2 + \sum_{t=1}^{T} \frac{1}{2} [\mathbf{Y}_t - X_t \beta - \alpha z_t]^T \Omega [\mathbf{Y}_t - X_t \beta - \alpha z_t] \right\} \\
&\propto \exp\left\{ -\frac{1}{2} [\alpha^2 (\tau_\alpha + \sum_{t=1}^{T} z_t^T \Omega z_t^T) - 2\alpha \sum_{t=1}^{T} [z_t^T \Omega (\mathbf{Y}_t - X_t \beta) \right\} \\
&\propto \exp\left\{ -\frac{1}{2} [\tau_\alpha^* \alpha^2 - 2\mu_\alpha] \right\}
\end{aligned}$$

where

$$\mu_\alpha = \sum_{t=1}^{T} z_t^T \Omega (\mathbf{Y}_t - X_t \beta)$$
$$\tau_\alpha^* = t_\alpha + \sum_{t=1}^{T} z_t^T \Omega z_t.$$

Then $\alpha \mid \ldots \sim N(\tau_\alpha^{*-1} \mu_\alpha, \tau_\alpha^{*-1})$

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. URLhttp://chjs.deuv.cl/Vol3N2/ChJS-03-02-04.pdf.

Zhang, H. and El-Shaarawi, A. (2010) On spatial skewGaussian processes and applications. *Environmetrics*, 33–47. URLhttp://onlinelibrary.wiley.com/doi/10.1002/env.982/abstract.