SCR Likelihood with a spatial log Gaussian Cox process and GMRF

David Finn

September 27, 2016

1 Likelihood function

1.1 Detection probability

Let ω_{itk} be the "capture response" recorded on occasion $t \in \{1, \ldots, T\}$ by detector $k \in \{1, \ldots, K\}$ for individual i. This could be binary (1 or 0) if all that is recorded is whether or not the individual was detected at all on the occasion by the detector, or it could be a count if the detector is able to record the number of times that the individual was detected on the occasion. Let $p_{itk}(\boldsymbol{s})$ be the probability of observing ω_{itk} for an individual at $\boldsymbol{s}$. (Note: subscript i is there as a shorthand way of denoting that this function is evaluated for capture response ω_{itk}.)

Let $g_{tk}(\boldsymbol{s})$ be the probability that on occasion t detector k detects an individual with activity centre $\boldsymbol{s}$, i.e., $g_{tk}(\boldsymbol{s}) = \mathbb{P}(\omega_{itk} > 0|s)$, and $p_{\cdot\cdot}(\boldsymbol{s})$ be the probability that an individual with activity centre $\boldsymbol{s}$ was detected by at least one detector on at least one occasion. Then

$$p_{\cdot\cdot}(\boldsymbol{s}) \quad = \quad 1 - \prod_t \prod_k \{1 - g_{tk}(\boldsymbol{s})\} \tag{1}$$

1.2 Poisson point process

Assume that individuals occur in the plane according to a Poisson process with intensity $\lambda(\boldsymbol{s})$ at location $\boldsymbol{s}$. If an individual at $\boldsymbol{s}$ is detected with probability $p_{\cdot\cdot}(\boldsymbol{s})$ then detected individuals occur in the plane according to a Poisson process with intensity $\tilde{\lambda}(\boldsymbol{s}) = \lambda(\boldsymbol{s})p_{\cdot\cdot}(\boldsymbol{s})$. Define $\Lambda = \int_S \lambda(\boldsymbol{s})ds$ and $\tilde{\Lambda} = \int_S \tilde{\lambda}(\boldsymbol{s})ds$.

1.3 Likelihood

It follows from the above that the number of individuals detected (n) is a Poisson random variable with rate parameter $\tilde{\Lambda}$. Let $P_{it}(\boldsymbol{s}_i)$ be the probability of observing $\{\omega_{itk}\}_k = (\omega_{it1}, \ldots, \omega_{itK})$ on occasion t for an individual with activity centre at $\boldsymbol{s}$. (If individuals are detected independently between detectors then $P_{it}(\boldsymbol{s}_i) = \prod_k p_{itk}(\boldsymbol{s}_i)$.) If individuals are detected independently across occasions, then the probability of observing the capture response data $\Omega_i = \{\omega_{itk}\}_{tk}$ (outer subscript shows indices "spanned" by set) for an individual with activity centre at $\boldsymbol{s}_i$ is $\prod_t P_{it}(\boldsymbol{s}_i)$. The conditional probability, given detection by at least one detector on at least one occasion is $[\prod_t P_{it}(\boldsymbol{s}_i)]/p_{\cdot\cdot}(\boldsymbol{s}_i) = P_i(\boldsymbol{s}_i)/p_{\cdot\cdot}(\boldsymbol{s}_i)$, where $P_i(\boldsymbol{s}_i) = \prod_t P_{it}(\boldsymbol{s}_i)$. And the probability density function of detected points, evaluated at $\boldsymbol{s}_i$ is $\tilde{\lambda}(\boldsymbol{s}_i)/\tilde{\Lambda}$. So if we define $\boldsymbol{S} = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n)$, the probability of observing $\Omega = \{\omega_{itk}\}_{itk}$ can be written as

$$
\begin{aligned}
\mathbb{P}(\Omega|S) &= \frac{\tilde{\Lambda}^n e^{-\tilde{\Lambda}}}{n!} \prod_{i=1}^{n} \frac{P_i(\boldsymbol{s}_i)}{p_{\cdot\cdot}(s_i)} \frac{\tilde{\lambda}(s_i)}{\tilde{\Lambda}} \\
&\propto e^{-\tilde{\Lambda}} \prod_{i=1}^{n} P_i(\boldsymbol{s}_i)\lambda(s_i) \qquad (2)
\end{aligned}
$$

We can't use this because we don't observe $\boldsymbol{s}$, so we need to marginalise over $\boldsymbol{s}$, and since the $\boldsymbol{s}_i$s are iid, we can write this as:

$$
\mathbb{P}(\Omega) \propto e^{-\tilde{\Lambda}} \prod_{i=1}^{n} \int_s P_i(\boldsymbol{s})\lambda(\boldsymbol{s})ds \qquad (3)
$$

1.4 Discrete approximation to the likelihood

Write $P_i(\boldsymbol{s})\lambda(\boldsymbol{s})$ as $\exp\{\eta(s; \lambda, \boldsymbol{\theta}, \Omega_i)\}$, where $\boldsymbol{\theta}$ is the parameters of $p_{itk}(\boldsymbol{s})$. (The reason for writing it thus is that the likelihood then involves $\log[\lambda(\boldsymbol{s}_j)]$, which is equal some fixed effect plus a GMRF random variable. Laplace approximation involves the second derivatives with respect to the GMRF variables, which is equivalent to the second derivatives with respect to the $\log[\lambda(\boldsymbol{s}_j)]$s - see below.) We now approximate the integral by a weighted sum of $\exp\{\eta(s; \lambda, \boldsymbol{\theta}, \Omega_i)\}$ evaluated at a discrete set of M points $\boldsymbol{s} = (\boldsymbol{s}_1, \ldots, s_M)$ spanning the integration area and weight α_j at location $\boldsymbol{s}_j$, and this is the approximate likelihood for $\lambda, \boldsymbol{\theta}$:

$$
\begin{aligned}
\mathcal{L}(\lambda, \boldsymbol{\theta}) &\approx e^{-\tilde{\Lambda}} \prod_i \sum_j \alpha_j \exp\{\eta(\boldsymbol{s}_j; \lambda, \boldsymbol{\theta}, \Omega_i)\}) \qquad (4) \\
&= \exp\left(-\sum_{j=1}^{M} \alpha_j p_{\cdot\cdot}(\boldsymbol{s}_j) e^{\log[\lambda_j]} + \sum_{i=1}^{n} \log\left[\sum_{j=1}^{M} \alpha_j P_i(\boldsymbol{s}_j) e^{\log[\lambda(\boldsymbol{s}_j)]} \right] \right).
\end{aligned}
$$

The corresponding (approximate) log-likelihood is

$$
l(\lambda, \boldsymbol{\theta}) = -\sum_{j=1}^{M} \alpha_j p_{\cdot\cdot}(\boldsymbol{s}_j) e^{\log[\lambda_j]} + \sum_{i=1}^{n} \log\left[\sum_{j=1}^{M} \alpha_j P_i(\boldsymbol{s}_j) e^{\log[\lambda(\boldsymbol{s}_j)]} \right] \qquad (5)
$$

1.5 Laplace approximation

Now $\log[\lambda(\boldsymbol{s}_j)] = \boldsymbol{x}_j\boldsymbol{\beta} + \xi_j$ $(j = 1, \ldots, M)$, where $\boldsymbol{x}_j$ is a vector of explanatory variables at location j and $\boldsymbol{\beta}$ is a parameter vector. We assume that $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_M)$ is a GMRF with mean zero and variance $\boldsymbol{\Sigma}_\xi$:

$$
\boldsymbol{\xi} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_\xi). \qquad (6)
$$

Denote its pdf $f_\xi(\boldsymbol{\xi})$, and to reflect the fact that $\log[\lambda(\boldsymbol{s})]$, and hence $l(\lambda, \boldsymbol{\theta})$ depends on $\boldsymbol{\xi}$ and $\boldsymbol{\beta}$, we write the approximate log likelihood as $l(\boldsymbol{\xi}, \boldsymbol{\beta}, \boldsymbol{\theta})$.

The marginal likelihood (integrating out the GMRF) is then

$$\mathcal{L}(\boldsymbol{\beta},\boldsymbol{\theta}) = \int \cdots \int \mathcal{L}(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta}) f_\xi(\boldsymbol{\xi})\ \partial\xi_1 \cdots \partial\xi_M \tag{7}$$

where $\mathcal{L}(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta})$ is Equation (4) written as a function of $\boldsymbol{\xi}$ and $\boldsymbol{\beta}$, rather than λ We approximate this using Laplace approximation:

$$\tilde{\mathcal{L}}(\boldsymbol{\beta},\boldsymbol{\theta}) = \sup_{\boldsymbol{\xi}} \{\mathcal{L}(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta}) f_\xi(\boldsymbol{\xi})\} \frac{(2\pi)^{\frac{M}{2}}}{|\boldsymbol{H}|^{\frac{1}{2}}} \tag{8}$$

where $\boldsymbol{H}$ is the matrix of second derivatives of $\log[\mathcal{L}(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta}) f_\xi(\boldsymbol{\xi})]$ with respect to $\boldsymbol{\xi}$ (see below).

1.6 Derivatives of $\log[\mathcal{L}(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta}) f_\xi(\boldsymbol{\xi})]$

We write $\log[\mathcal{L}(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta}) f_\xi(\boldsymbol{\xi})]$ as $l(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta}) + \log[f_\xi(\boldsymbol{\xi})]$ and calculatethe derivatives of $l(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta})$ and $\log[f_\xi(\boldsymbol{\xi})]$.

1.6.1 Derivatives of $l(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta})$

Since $\frac{\partial l(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta})}{\partial \xi_j} = \frac{\partial l(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta})}{\partial \log[\lambda_j]} \frac{\partial \log[\lambda_j]}{\partial \xi_j} = \frac{\partial l(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta})}{\partial \log[\lambda_j]}$, it is convenient to work in terms of $\frac{\partial l(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta})}{\partial \log[\lambda_j]}$ below, and for brevity we write $l(\boldsymbol{\xi},\boldsymbol{\beta},\boldsymbol{\theta})$ as l. Now differentiate with respect to $\log[\lambda_j]$ (i.e., $\log[\lambda(\boldsymbol{s}_j)]$ for $j = 1,\ldots,M$), noting that $\frac{\partial \lambda_j}{\partial \log[\lambda_j]} = \lambda_j$:

$$\begin{aligned}\frac{\partial l}{\partial \log[\lambda_j]} &= -\alpha_j\lambda_j p_{\cdot\cdot}(\boldsymbol{s}_j) + \sum_{i=1}^{n} \frac{\alpha_j\lambda_j P_i(\boldsymbol{s}_j)}{\sum_{j^*=1}^{M} \alpha_{j^*}\lambda_{j^*} P_i(\boldsymbol{s}_{j^*})} \\ &= -\alpha_j\lambda_j p_{\cdot\cdot}(\boldsymbol{s}_j) + \alpha_j\lambda_j \left\{ \sum_{i=1}^{n} \frac{P_i(\boldsymbol{s}_j)}{\sum_{j^*=1}^{M} \alpha_{j^*}\lambda_{j^*} P_i(\boldsymbol{s}_{j^*})} \right\}\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \log[\lambda_j]^2} &= -\alpha_j\lambda_j p_{\cdot\cdot}(\boldsymbol{s}_j) + \alpha_j\lambda_j \left\{ \frac{P_i(\boldsymbol{s}_j)}{\sum_{j^*=1}^{M} \alpha_{j^*}\lambda_{j^*} P_i(\boldsymbol{s}_{j^*})} \right\} \\ &\quad + \alpha_j\lambda_j \left\{ -\alpha_j\lambda_j \sum_{i=1}^{n} \frac{P_i(\boldsymbol{s}_j)^2}{\left[\sum_{j^*=1}^{M} \alpha_{j^*}\lambda_{j^*} P_i(\boldsymbol{s}_{j^*})\right]^2} \right\}\end{aligned} \tag{10}$$

$$\frac{\partial^2 l}{\partial \log[\lambda_j] \partial \log[\lambda_k]} = -\alpha_j\lambda_j\alpha_k\lambda_k \sum_{i=1}^{n} \frac{P_i(\boldsymbol{s}_j)P_i(\boldsymbol{s}_k)}{\left[\sum_{j^*=1}^{M} \alpha_{j^*}\lambda_{j^*} P_i(\boldsymbol{s}_{j^*})\right]^2} \tag{11}$$

1.6.2 Derivatives of $\log[f_\xi(\boldsymbol{\xi})]$

$$\frac{\partial \log[f_\xi]}{\partial \boldsymbol{\xi}} = \boldsymbol{\Sigma}_\xi^{-1}\boldsymbol{\xi} \tag{12}$$

$$\frac{\partial^2 \log[f_\xi]}{\partial \boldsymbol{\xi} \partial \boldsymbol{\xi}^t} = \boldsymbol{\Sigma}_\xi^{-1}. \tag{13}$$

1.6.3 Elements of $\boldsymbol{H}$

$$\boldsymbol{H}_{jk} = \frac{\partial^2 l}{\partial \log[\lambda_j] \partial \log[\lambda_k]} + \frac{\partial^2 f_\xi(\boldsymbol{\xi})}{\partial \xi_j \partial \xi_k} \tag{14}$$