Ecological Archives **E089-074-A1**

Devin S. Johnson, Joshua M. London, Mary-Anne Lea, and John W. Durban. 2008. Continuous-time correlated random walk model for animal telemetry data. Ecology 89:1208–1215.

Appendix B: Details of the Kalman Filter Recursions for CTCRW.

The Kalman filter recursions are an invaluable tool when modeling time series with an SSM. Because we are considering independent processes of movement in each coordinate direction, we formulate the filtering and estimation in terms of two one-dimensional series of observations. The Kalman filter recursions for the CTCRW model are given by

$$\begin{array}{ll} v_{ci} = y_{ci} - \mathbf{Z}_i'\mathbf{a}_{ci}, & F_{ci} = \mathbf{Z}_i'\mathbf{P}_{ci}\mathbf{Z}_i + H_{ci}, \\ \mathbf{K}_{ci} = \mathbf{T}_i\mathbf{P}_{ci}\mathbf{Z}_i/F_{ci}, & \mathbf{L}_{ci} = \mathbf{T}_i - \mathbf{K}_{ci}\mathbf{Z}_i', \\ \mathbf{a}_{c,i+1} = \mathbf{T}_i\mathbf{a}_{ci} + \mathbf{K}_{ci}v_{ci}, & \mathbf{P}_{c,i+1} = \mathbf{T}_i\mathbf{P}_{ci}\mathbf{L}_{ci}' + \mathbf{Q}_{ci}, \end{array} \tag{B.1}$$

for $i = 1, \ldots, n$ and $c = 1, 2$. Vague initial conditions are often used to begin the recursions (e.g., using $\mathbf{a}_{c1} = [y_{c1}, 0, 0]$ and a "large" prediction covariance matrix $\mathbf{P}_{c1}$). Assuming the series of observations is sufficiently long this choice will have little influence. See de Jong and Penzer (1998) for an alternative approach.

Due to the fact ϵ_i and $\boldsymbol{\eta}_i$ are Gaussian the log-likelihood is

$$L(\boldsymbol{\theta}) = -n \log 2\pi - \frac{1}{2}\sum_{c=1}^{2}\sum_{i=1}^{n}\left(\log F_{ci} + v_{ci}^2/F_{ci}\right), \tag{B.2}$$

where $\boldsymbol{\theta}$ is a vector of all model parameters. The log-likelihood can be optimized to provide parameter values for state estimation. One must be careful to note that the index runs only over times for which a location is observed. The residual v_{ci} is undefined if an observation has not been made.

Estimation of the state processes involves running the Kalman filter forward and then running a set of smoothing recursions backwards through the series (Durbin and Koopman, 2001, pg. 73) to obtain the state estimates $\widehat{\boldsymbol{\alpha}}_{ci}$ and variance estimates $\widehat{\mathbf{V}}_{ci}$. The smoothing recursions for each coordinate are given by

$$\begin{array}{ll} \mathbf{r}_{c,i-1} = \mathbf{Z}_i v_{ci}/F_{ci} + \mathbf{L}_{ci}'\mathbf{r}_{ci} & \mathbf{N}_{c,i-1} = \mathbf{Z}_i\mathbf{Z}_i'/F_{ci} + \mathbf{L}_{ci}'\mathbf{N}_{ci}\mathbf{L}_{ci} \\ \widehat{\boldsymbol{\alpha}}_{ci} = \mathbf{a}_{ci} + \mathbf{P}_{ci}\mathbf{r}_{c,i-1} & \widehat{\mathbf{V}}_{ci} = \mathbf{P}_{ci} - \mathbf{P}_{ci}\mathbf{N}_{c,i-1}\mathbf{P}_{ci} \end{array} \tag{B.3}$$

for $i = n, \ldots, 1$ and $c = 1, 2$, where $\mathbf{r}_{cn} = \mathbf{0}$, $\mathbf{N}_{cn} = \mathbf{0}$. The combined filter and smoothing recursions are known as the *Kalman filter and smoother* (Durbin and Koopman, 2001).

Now, let $\hat{\boldsymbol{\alpha}}_i = [\boldsymbol{\alpha}_{1i}, \boldsymbol{\alpha}_{2i}]'$ and $\widehat{\mathbf{V}}_i = \text{blockdiag}\{\widehat{\mathbf{V}}_{1i}, \widehat{\mathbf{V}}_{2i}\}$. Once the state has been estimated, the delta method can be used for estimating standard errors of a nonlinear function, such as speed $S(t_i)$ at time t_i (see *Northern fur seal pup migration* section). Then using the *deltamethod*() function of the *msm* library of R, for example, standard errors can be estimated via

`deltamethod(` $f(), \hat{\boldsymbol{\alpha}}_i, \widehat{\mathbf{V}}_i$`),`

where $f()$ is the nonlinear function (see `http://finzi.psych.upenn.edu/R/library/msm/html/deltamethod.html` for documentation)

One valuable benefit of using the KFS is that it handles missing observations easily. It was stated that a benefit of the continuous time model is that one does not have to worry about missing data if a scheduled relocation is missed. It can be useful, however, to include times in the data set for which locations were not obtained. If one would like an estimated value of $\boldsymbol{\alpha}_i$ for at a time t_i when no location was observed, all that must be done is to supplement the data set with missing $\mathbf{y}_i = [\mathrm{NA}, \mathrm{NA}]'$ values at that time. The KFS automatically filters and smooths to provide location and variance estimates $\widehat{\boldsymbol{\alpha}}_i$ and $\widehat{\mathbf{V}}_i$. This is accomplished by setting $v_{ci} = 0$, $\mathbf{K}_{ci} = \mathbf{0}$, and $\mathbf{N}_{c,i-1} = \mathbf{T}_i \mathbf{N}_{ci} \mathbf{T}_i$ for observation time t_i. See the section on *Harbor seal movement* and the following appendix for an example.

Literature Cited

de Jong, P., and J. Penzer. 1998. Diagnosing shocks in time series. Journal of the American Statistical Association **93**:796–806.

Durbin, J., and S. Koopman. 2001. Time Series Analysis by State Space Methods. Oxford University Press, Oxford, UK.