Group 1: Investigating Identifiability Using R

This worksheet is about investigating identifiability using the Hessian method and Likelihood profile, which can both be executed in R.

The following files are needed:
- hessianmethod.R
- profilemethod.R
- eg2lik.R
- eg3lik.R

A: Hessian Method

Open the R file hessianmethod.R and run the commands listed. This will execute the Hessian method on the introductory Lapwing example with parameters ϕ_1, ϕ_a, ρ and the reparameterised Lapwing example with parameters, $\phi_a, \beta = \phi_1 \rho$.

Most of the instructions are given in the comments, but note the following:
- The first function hessianmethod evaluates the Hessian method and calls the function hessian, which is based on a similar function from Morgan (2009).
- statespacelik and statespacelik2 are functions for evaluating the likelihood using the Kalman Filter. The variance parameter is fixed, and arbitrary initialisation values are used. Whilst these would need exploration for inference they are sufficient for investigating identifiability of ϕ_1, ϕ_a, ρ.

B. Profile Method

Open the R file profilemethod.R and run the commands listed. This will create likelihood profiles for the introductory Lapwing example with parameters ϕ_1, ϕ_a, ρ and the reparameterised Lapwing example with parameters, $\phi_a, \beta = \phi_1 \rho$.

Most of the instructions are given in the comments, but note the following:
- The first function profileplot plots log-likelihood profiles for each parameter in a model, it calls profilelik to find the log-likelihood where one (or more) parameter is fixed.
- statespacelik and statespacelik2 are function for evaluating the likelihood using the Kalman Filter. The variance parameter is fixed, and arbitrary initialisation values are used. Whilst these would need exploration for inference they are sufficient for investigating identifiability of ϕ_1, ϕ_a, ρ.

Further Examples

Now look at the two further examples below.

Example 2: Immigration Model

Abadi *et al* (2010) examine an integrated population model for a population of Little Owls, where the population is affected by immigration. The integrated population model consists of census data and capture-recapture data (there is also productivity data which is not considered here). The model for the census data is:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} 0 & 0.5\rho\phi_1 \\ \phi_a + imm & \phi_a + imm \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{a,t} \end{bmatrix}$$

$$y_t = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + \eta_t$$

where $x_{1,t}$ is the number of 1-year old females, $x_{2,t}$ is the number of females older than a year, y_t is the total number of females. The parameters are: ϕ_j juvenile survival probability, ϕ_a adult survival, ρ productivity and imm the immigration rate. (η_t and ϵ_i are error processes.) The capture-recapture data consists of separate data for

juveniles and adults. A CJS model is used with parameters ϕ_j, ϕ_a and p, where p is the recapture probability. (Note here we ignore the data on males, and only include the female data – this does not change the identifiability results).

The likelihood function is given in eg2lik.R. Note that we would not recommend using the Kalman filter approximation in this case for model fitting, due to the small census numbers.

Is the model with census data alone identifiable? Is the integrated model with census and capture-recapture data identifiable? Which parameters can be estimated in both cases?

Example 3: Salmon State-Space Model

This example is from Newman *et al* (2014). The number of juvenile Salmon alive in year *t* in a certain river can be modelled as

$$N_t|N_{t-1} \sim Po(\alpha N_{t-1} \exp(-\beta N_{t-1})), \qquad \alpha > 0, \beta > 0.$$

The parameter α models survival and fecundity, the parameter β models density dependence and the parameter $n_0 = N_0$. The observation process is

$$y_t|N_t \sim lognormal\left(\log(N_t) - \frac{\sigma^2}{2}, \sigma^2\right).$$

σ^2 is assumed to be a fixed known value.

The likelihood used is a normal approximation to this model using a non-linear extension of the Kalman filter (see Section 4.4.2 of Newman et al, 2014). Code for the log-likelihood is given in eg3lik.R.

Is the model identifiable? What if the parameter α was time dependent?

References

Abadi, F., Gimenez, O.**,** Arlettaz, R. and Schaub, M. (2010). An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence. *Ecology*, **91**, 7-14.

Morgan, B. J. T. (2009), *Applied Stochastic Modelling*. Chapman and Hall.

Newman, K.B., Buckland, S.T., Morgan, B.J.T., King, R., Borchers, D.L. and Cole, D.J. and Besbeas, P. and Gimenez, O. and Thomas, L. (2014) Modelling Population Dynamics: model formulation, fitting and assessment using state-space methods. Springer.