MA888 Ecological Statistics
`Kalm`: Integrated Population Modelling

R.S.McCrea and B.J.T.Morgan

Spring Term 2012

Introduction

Program `Kalm` was written by P.T.Besbeas and is a piece of software capable of fitting integrated models to ring-recovery and census data. The notes presented here are adapted from some workshop notes written by P.T.Besbeas. Program `Kalm` is written in the MATLAB computing language, however because the software is conversational no prior knowledge of MATLAB is required.

The software is stored in the MA888 moodle folder. Open MATLAB through Central Software. Once MATLAB is open, navigate to the Program Kalm folder and follow the instructions in these notes.

1 Example: Simulated Ring-recovery and Census Data

This first example is on a simulated data set and is supposed to give you a gentle introduction to the software.

1. An m-file called 'example.m' is in the kalm directory. Open the file and confirm that it contains: `Ringed = [1000;1000;1000;1000];`
   ```
   Mobs=[35 7 0 1
   0 23 11 3
   0 0 45 13
   0 0 0 33];
   Census = [500;700;800;600;900];
   ```
2. In the command window of MATLAB type `kalm1_0` to start the software.
3. Enter `1` to select 'exact analysis'.
4. Enter `1` to select 'new session'.

5. You now need to tell the software that you want to analyse your example data set. To do this type `example` to read the data into the program.

6. Enter a filename to store the session output - accepting the default `example.res` is fine.

7. Enter `1` to select 'Fit a Model'.

8. You now have to choose the parameter dependencies for each of your parameters. There are four parameters in total: juvenile and adult survival probabilities, recovery probability and the productivity parameter. The model notation follows the Freeman-Morgan notation introduced for ring-recovery data, except we have an additional productivity parameter definition to add at the end. To begin with try fitting a constant model for all parameters, i.e. `c/c/c/c`.

9. Do NOT select yes for the default starting values (there is a bug in the program!). Enter appropriate starting values for each of your parameters in turn.

10. Verify that the parameter estimates are:

    - $\phi^1 = 0.284$
    - $\phi^a = 0.296$
    - $\lambda = 0.048$
    - $p = 10.5$
    - $\sigma = 131$

11. View the parameter correlations and comment.

12. Write down the maximised log likelihood value and verify the value for AIC using the defining formula.

13. View the matrix of fitted recoveries. Why does the repeating structure exist?

14. View the pattern of residuals and use it as a goodness-of-fit diagnostic.

15. View the smoothed estimates of population numbers and note the population breakdown.

16. Plot the observed and fitted census numbers.

17. Repeat the model fitting of the constant model and assess the use of different starting values. Are the starting values important for this application?

18. Enter `0` to quit `Kalm`.

2 Example: Heron Application

This example allows you to analyse a set of real data and allows you to experiment with using time-varying covariates. The file `heron.m` is in your working directory. The file contains ring-recovery and census data on the heron and also monthly temperatures which may be used as covariates for one or more of the parameters.

1. Start `kalm1_0` as before and read in data `heron`.

2. To begin with, fit the constant model to the data. Remember not to use the default starting values.

3. Verify that the parameter estimates are:

   - $\phi^1 = 0.380$
   - $\phi^a = 0.620$
   - $\lambda = 0.094$
   - $p = 1.67$
   - $\sigma = 252$

4. Perform diagnostic steps from the previous example.

5. The monthly temperature data can be used as covariates for one or more of the parameters. To fit the model where ϕ^1 and ϕ^a depend on December temperature, and λ and p are constant, we need to define two covariate variables which contain appropriate December temperatures for the ring recovery and census data respectively. In `heron.m` these variables have been called `Decrr` and `Decc` respectively. Enter `v(Decrr,Decc)/v(Decrr(2:11),Decc)/c/c` to fit this model.

6. Verify that the parameter estimates on the logistic scale are:

   - $\phi^1(const = -0.46, regr = 0.06)$
   - $\phi^a(const = 0.53, regr = 0.01)$

   and view the individual ϕ^1 and ϕ^a values.

7. Average these values and compare them with the estimates of ϕ^1 and ϕ^a from the constant model.

8. Compare the estimates and standard errors of λ, p and σ with those from the constant model.

9. Perform the diagnostic steps from the previous example.

10. By comparing AIC values we find that the best (weather) covariate for ϕ^1 is December while that for ϕ^a is February. A different type of covariate is a 'year-covariate'. Year-covariates allow parameters to vary smoothly with time and are typically very useful for modelling λ and p. To create year-covariates for the heron ring-recovery and census data we need to edit the file `heron.m`, adding the following lines:

```
Yrrr=linspace(0,1,11)';

Yrc=linspace(0,1,16)';
```

    (Note that you need to exit `kalm` before opening the file.) These lines create two column vectors of 11 and 16 equally spaced points between 0 and 1 respectively.

11. Enter `v(Decrr,Decc)/v(Febrr(2:11),Febc)/v(Yrrr)/v(Yrc)` to fit this model.

12. Verify that your parameter estimates are:

    - $\phi^1(const = -0.46, regr = 0.04)$
    - $\phi^a(const = 0.54, regr = 0.04)$
    - $\lambda(const = -2.21, regr = 0.47)$
    - $p(const = 0.42, regr = -0.10)$
    - $\sigma = 188.8$

13. In the models above, the parameter ϕ^a was either constant or varied with time. It is sometimes useful to fit models where ϕ^a varies with age, up to age n. We can use `Kalm` to fit such models. Fit the model $c/a2/c/c$.

14. Compare the AIC value of this model with the AIC values of the previous models and discuss.

3 Example: Approximate Analysis

These examples illustrate the use of the multivariate normal approximation to the MRR likelihood and assesses its performance.

3.1 Simulated Example Revisited

The parameter estimates and variance-covariance of ϕ^1 and ϕ^a in logistic space, when the constant model $c/c/c$ is fitted to the ring-recovery data of `example.m` are:

$$\mu = \begin{pmatrix} -0.9819 \\ -1.3645 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 0.0384 & 0.0279 \\ 0.0279 & 0.3087 \end{pmatrix}$$

1. Create an m-file called `aexample.m` containing the data:
   ```
   Mu=[-0.9819;-1.3645];
   Sigma=[0.0284 0.0279
   0.0279 0.3087];
   Census=[500; 700; 800; 600; 900];
   ```
   in your working directory.

2. Start `kalm` and select a 'new session' within an 'approximate analysis'.

3. Read in the `aexample` data.

4. Fit a constant model (remembering now you only need to define three parameters).

5. Verify that your parameter estimates are:

   - $\phi^1 = 0.282$
   - $\phi^a = 0.297$
   - $p = 10.5$
   - $\sigma = 131$

6. Compare these parameter estimates to the estimates from the exact analysis completed previously.

3.2 Heron Application Revisited

The file `aheron.m` in your working directory contains the estimates and variance-covariances of the survival parameters from fitting the model

$$\phi^1(December)/\phi^a(February)/\lambda(year)$$

to the ring-recovery data alone. The file also contains the heron census data and covariates for analysing these data.

1. Read in the data `aheron` and start a new session within an approximate analysis.

2. Fit model $v(Decc)/v(Febc)/v(Yrc)$ to the data.

3. Verify that the parameter estimates are:

   - $\phi^1(const = -0.47, regr = 0.04)$
   - $\phi^a(const = 0.52, regr = 0.04)$

- $p(const = 0.45, regr = -0.10)$
- $\sigma = 189$

4. Compare with the corresponding estimates from the exact analysis.

5. Create a file `aheron2.m` in your working directory containing the data:
```
Mu=[-0.5201; 0.1823; 0.5286];
Sigma = [0.0092 0.0045 0.0074
0.0045 0.0262 0.0136
0.0074 0.0136 0.0523];
```
   and the heron census from `heron.m`.

6. Fit the model $c/a2/c$.

7. Compare the parameter estimates with the exact analysis.

4 Cormorant Application

We now return to the Cormorant example that has run through the entire course. We previously analysed the ring-recovery data alone using `Eagle` and we found that the best model for the ring-recovery data alone was $T/C/T$. We now use `Kalm` to fit models to the ring-recovery and census data.

1. Start `Kalm1_0` as before and read in data `corms`.

2. To begin with, fit the constant model to the data. Remember not to use the default starting values.

3. Complete the following table below by fitting the defined models and noting their AIC values.

4. From this table, which model do you think is the best for the combined ring-recovery and census data.

5. We note that some of these fitted models have observation error estimates of zero. This can be a warning that we have over-fitted the model, and have too many parameters in the model for the amount of census data we have. Boundary estimates of observation error are a complex problem and are a current area of work within our department.

Model	np	AIC
$c/c/c/c$		
$c/c/t/c$		
$t/c/c/c$		