SQU Journal For Science, 9 (2004) 53-66

Recursive Estimation of a Discrete-Time Lotka-Volterra Predator-Prey Model

Lakhdar Aggoun

Department of Mathematics and Statistics, Sultan Qaboos University,
P. O. Box 36, Alkhod 123, Sultanate of Oman,
Email: laggoun@squ.edu.om.

التقدير الشرطي المتكرر لنموذج لوتكه- فلتره للصائدة والمصطادة

الأخضر عقون

خلاصة : هذا المقال يعالج مسألة مهمة في الإحصاء البيني وهي مسألة حساب التعداد الحيواني لنوعين من الحيوانات وهي الصائدة و المصطادة. من كل نوعية تؤخذ عينة عشوائية وتميز بإشارة خاصة ثم يطلق سراحها داخل بقية القطيع. وبعد فترة تؤخذ عينة عشوائية من القطيعين ويحصي منها الأفراد المميزين بالإشارة الخاصة ومن هذه المعلومات تستعمل طرق القياس لتقدير التوزيع الشرطي المتكرر لمجموعة الحيوانات الصائدة والمصطادة.

ABSTRACT: In this paper, using hidden Markov models, we estimate the number of individuals in a two-species (predator-prey) animal population using partial information provided by the so-called capture-recapture technique. Random samples of individuals are captured, tagged in some way and released. After some time other random samples are taken and the marked individuals are observed. Using this information, we estimate (recursively) the sizes of the two populations. Also, using the Expectation Maximization (EM) algorithm, the parameters of the model are updated.

KEYWORDS: Hidden markov models, predator-prey, capture-recapture, EM algorithm.
AMS subject classification: 60J27, 93E11.

1. Introduction

One of the first models to incorporate interactions between predators and preys was proposed in 1925 by the American biophysicist Alfred Lotka and the Italian mathematician Vito Volterra .
Vito Volterra (1860-1940) was a famous Italian mathematician who retired from a distinguished career in pure mathematics in the early 1920s. His son-in-law, Humberto D'Ancona, was a biologist who studied the populations of various species of fish in the Adriatic. Volterra developed a series of models for interactions of two or more species.

Alfred J. Lotka (1880-1949) was an American mathematical biologist (and later actuary) who formulated many of the same models as Volterra, independently and at about the same time. His primary example of a predator-prey system comprised a plant population and an herbivorous animal dependent on that plant for food.

Here we consider a discrete-time stochastic version of the Lotka-Volterra model. Hidden Markov models (Elliott *et al*. 1995) have been used extensively in many areas of science and technology. In this paper we are extending the use of this powerful tools (see Aggoun and Elliott 1998 for a single species model) to estimate the hidden number of individuals in a multi-species animal population using partial information provided by the so-called capture-recapture technique (see Seber 1982, for instance).

Two random samples of individuals one from the predator population and one from the prey population are captured, tagged or marked in some way, and then released. After allowing time for the marked and unmarked to mix sufficiently, two second simple random samples from both populations are taken and the marked ones are observed.

At epoch ℓ write Z_ℓ for the prey population size, $\tilde{z}_\ell$ for the number of marked and released preys, $\tilde{\tilde{z}}_k = \sum_{\ell=1}^{k} \tilde{z}_\ell$ for the total number of captured and marked preys up to time k, R_ℓ for the sample size, z_ℓ for the number of available marked individuals for sampling and y_ℓ^Z for the number of captured (or recaptured) marked individuals.

Similarly write X_ℓ for the predator population size, $\tilde{x}_\ell$ for the number of marked and released predators, $\tilde{\tilde{x}}_k = \sum_{\ell=1}^{k} \tilde{x}_\ell$ for the total number of captured and marked predators up to time k, F_ℓ for the sample size, x_ℓ for the number of available marked predators for sampling and y_ℓ^X for the number of captured (or recaptured) marked predators.

2. Model Assumptions and Recursive Estimation

All random variables are defined initially on a probability space (Ω, F, P). All the filtrations defined here are assumed to be complete.

Write $G_k = \sigma(Z_\ell, z_\ell, y_\ell^Z, R_\ell, X_\ell, x_\ell, y_\ell^X, F_\ell \ell \le k)$, and $Y_k = \sigma(y_\ell^Z, y_\ell^X, \ell \le k)$.

X_k and Z_k represent the number of predators and preys, respectively, that are alive at time period k, then a (discrete time) Lotka-Volterra type model is:

$$\begin{aligned} X_{k+1} &= X_k + aZ_k X_k - bX_k + \sigma_1 v_{k+1} \\ Z_{k+1} &= Z_k + cZ_k - dZ_k X_k + \sigma_2 w_{k+1}, \end{aligned} \tag{1}$$

where the parameters are defined by:

a is the efficiency of turning predated preys into predators.
b is the natural death rate of predators in the absence of food (preys),
c is the natural growth rate of preys in the absence of predation,
d is the death rate per encounter of preys due to predation,
v and w are sequences of independent random variables with some (either discrete or continuous with finite supports) densities ϕ_k and ϑ_k respectively and σ_1, σ_2 are some positive real numbers. The random

variables v and w indicate other sources of variations in the populations like death caused by old age, diseases, births etc.

It is assumed here that x_k and z_k are random variables with binomial distributions with parameters $(\tilde{\tilde{x}}_k, p_X)$ and $(\tilde{\tilde{z}}_k, p_Z)$ respectively.

The observed random variables y_k^Z, y_k^X are assumed to have joint distribution:

$$P(y_k^X = m, y_k^Z = n \mid G_{k-1}, z_k, Z_k, R_k, x_k, X_k, F_k)$$
$$= \binom{F_k}{m}\binom{R_k}{n}\left(\frac{x_k}{X_k}\right)^m\left(1-\frac{x_k}{X_k}\right)^{F_k-m}\left(\frac{z_k}{Z_k}\right)^n\left(1-\frac{z_k}{Z_k}\right)^{R_k-n} \tag{5}$$

Let $\lambda_0 = 1$. For $\ell \geq 1$ and for suitable density functions χ, v write $\lambda_\ell = \lambda_\ell^X \lambda_\ell^Z$, where

$$\lambda_\ell^X = \frac{\sigma_1 \chi_\ell(X_\ell)}{\phi_\ell(v_\ell)} \frac{1}{2^{F_\ell + \tilde{\tilde{x}}_\ell}} p_X^{-x_\ell}(1-p_X)^{-\tilde{\tilde{x}}_\ell + x_\ell}\left(\frac{x_\ell}{X_\ell}\right)^{-y_\ell^X}\left(1-\frac{x_\ell}{X_\ell}\right)^{y_\ell^X - F_\ell} \tag{6}$$

$$\lambda_\ell^Z = \frac{\sigma_2 v_\ell(Z_\ell)}{\vartheta_\ell(w_\ell)} \frac{1}{2^{R_\ell + \tilde{\tilde{z}}_\ell}} p_Z^{-z_\ell}(1-p_Z)^{-\tilde{\tilde{z}}_\ell + z_\ell}\left(\frac{z_\ell}{Z_\ell}\right)^{-y_\ell^Z}\left(1-\frac{z_\ell}{Z_\ell}\right)^{y_\ell^Z - R_\ell} \tag{7}$$

and $\Lambda_k = \prod_{\ell=0}^{k} \lambda_\ell$.

Lemma 1. The process Λ_k is a G-martingale.

Proof. $E[\Lambda_k \mid G_{k-1}] = \Lambda_{k-1} E[\lambda_k^X \lambda_k^Z \mid G_{k-1}]$, so we must show that $E[\lambda_k^X \lambda_k^Z \mid G_{k-1}] = 1$. However using repeated conditioning it is enough to show that $E[\lambda_k^Z \mid G_{k-1}] = 1$, say.

$$E[\lambda_k^Z \mid G_{k-1}] = E\left[\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)} \frac{1}{2^{R_k + \tilde{\tilde{z}}_k}}\left(\frac{z_k}{Z_k}\right)^{-y_k^Z}\left(1-\frac{z_k}{Z_k}\right)^{y_k^Z - R_k}\right.$$
$$\left.\times p_Z^{-z_k}(1-p_Z)^{-\tilde{\tilde{z}}_k + z_k} \mid G_{k-1}\right]$$
$$= \frac{1}{2^{R_k + \tilde{\tilde{z}}_k}} E\left[\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)} p_Z^{-z_k}(1-p_Z)^{-\tilde{\tilde{z}}_k + z_k} E\left[\left(\frac{z_k}{Z_k}\right)^{-y_k^Z}\right.\right.$$
$$\left.\left.\times\left(1-\frac{z_k}{Z_k}\right)^{y_k^Z - R_k} \mid G_{k-1}, Z_k, z_k, R_k\right] \mid G_{k-1}\right]$$
$$= \frac{1}{2^{R_k + \tilde{\tilde{z}}_k}} E\left[\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)} p_Z^{-z_k}(1-p_Z)^{-\tilde{\tilde{z}}_k + z_k} \sum_{m=0}^{R_k}\binom{R_k}{m} \mid G_{k-1}\right]$$

$$= \frac{1}{2^{\tilde{\tilde{z}}_k}} E[\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)} p_Z^{-z_k}(1-p_Z)^{-\tilde{\tilde{z}}_k+z_k} \mid G_{k-1}]$$

$$= \frac{1}{2^{\tilde{\tilde{z}}_k}} E[\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)} \sum_{i=0}^{\tilde{\tilde{z}}_k} \binom{\tilde{\tilde{z}}_k}{i} \mid G_{k-1}]$$

$$= E[\frac{\sigma_2 v_k(Z_{k-1}+\sigma_2 w_k)}{\vartheta_k(w_k)} \mid G_{k-1}]$$

$$= \int \frac{\sigma_2 v_k(Z_{k-1}+\sigma_2 w)}{\vartheta_k(w)} \vartheta_k(w) dv = \int v_k(u) du = 1$$

A new probability measure Q can be defined by setting $E\left[\frac{dQ}{dP} \mid G_k\right] = \Lambda_k$. The point here is that:

Lemma 2. Under the new probability measure Q, X_k, x_k, y_k^X Z_k, z_k and y_k^Z are sequences of independent random variables which are independent of each other. Further, X_k has density χ_k, x_k has distribution $bin(\tilde{\tilde{x}}_k, \frac{1}{2})$, y_k^x has distribution $bin(F_k, \frac{1}{2})$, Z_k has density v_k, z_k has distribution $bin(\tilde{\tilde{z}}_k, \frac{1}{2})$ and y_k^Z has distribution $bin(R_k, \frac{1}{2})$.

Proof. We shall check the claim for the three processes Z_k, z_k and y_k^Z. For any integrable real-valued functions f, g and h and using a version of Bayes' theorem (see Elliott *et al*. 1995) we can write:

$$E_Q[f(Z_k)g(z_k)h(y_k^Z) \mid G_{k-1}] = \frac{E[f(Z_k)g(z_k)h(y_k^Z)\Lambda_k \mid G_{k-1}]}{E[\Lambda_k \mid G_{k-1}]}$$

$$= E[f(Z_k)g(z_k)h(y_k^Z)\lambda_k \mid G_{k-1}]$$

$$= E[f(Z_k)g(z_k)h(y_k^Z)\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)} \frac{1}{2^{\tilde{\tilde{n}}_k}} p_Z^{-z_k}(1-p_Z)^{-\tilde{\tilde{n}}_k+z_k}$$

$$E[h(y_k^Z)\frac{1}{2^{R_k}}(\frac{z_k}{Z_k})^{-y_k^Z}(1-\frac{z_k}{Z_k})^{y_k^Z-R_k} \mid G_{k-1}, Z_k, z_k, R_k] \mid G_{k-1}]$$

$$= E[f(Z_k)g(z_k)h(y_k^Z)\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)} \frac{1}{2^{\tilde{\tilde{n}}_k}} p_Z^{-z_k}(1-p_Z)^{-\tilde{\tilde{n}}_k+z_k}$$

$$[\sum_{m=0}^{R_k} h(m)(\frac{z_k}{Z_k})^m (1-\frac{z_k}{Z_k})^{-m+R_k} \binom{R_k}{m} \frac{1}{2^{R_k}} (\frac{z_k}{Z_k})^{-m}(1-\frac{z_k}{Z_k})^{m-R_k}] \mid G_{k-1}]$$

$$= E[f(Z_k)g(z_k)h(y_k^Z)\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)}\frac{1}{2^{\tilde{\tilde{n}}_k}} p_Z^{-z_k}(1-p_Z)^{-\tilde{\tilde{n}}_k+z_k}$$

$$\times[\sum_{m=0}^{R_k} h(m)\binom{R_k}{m}\frac{1}{2^{R_k}}] \mid G_{k-1}]$$

$$= E_Q[h(y_k^Z)]E[f(Z_k)g(z_k)\frac{\sigma_2 v_k(Z_k)}{\vartheta_k(w_k)}\frac{1}{2^{\tilde{\tilde{n}}_k}} p_Z^{-z_k}(1-p_Z)^{-\tilde{\tilde{n}}_k+z_k} \mid G_{k-1}]$$

$$= E_Q[h(y_k^Z)]E_Q[g(z_k)]E[f(Z_{k-1}+cZ_{k-1}-dZ_{k-1}X_k+\sigma_2 w_k)$$

$$\times\frac{\sigma_2 v_k(Z_{k-1}+cZ_{k-1}-dZ_{k-1}X_k+\sigma_2 w_k)}{\vartheta_k(w_k)} \mid G_{k-1}]$$

$$= E_Q[h(y_k^Z)]E_Q[g(z_k)]$$

$$\times\int f(Z_{k-1}+cZ_{k-1}-dZ_{k-1}X_k+\sigma_2 w)$$

$$\times\sigma_2 v_k(Z_{k-1}+cZ_{k-1}-dZ_{k-1}X_k+\sigma_2 w)dw$$

$$= E_Q[h(y_k^Z)]E_Q[g(z_k)]\int f(u)v_k(u)du$$

$$= E_Q[h(y_k^Z)]E_Q[g(z_k)]E_Q[f(Z_k)],$$

where $\binom{R_k}{m} = \frac{R_k!}{m!(R_k-m)!}$.

That is, under Q the three processes are independent sequences of random variables with the desired distributions.
Using this fact we derive a recursive equation for the unnormalized conditional distribution of Z_k and X_k given Y_k.
For any measurable test function f consider:

$$E[f(X_k, Z_k) \mid Y_k] = \frac{E_Q[f(X_k, Z_k)\Lambda_k^{-1} \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]}. \tag{8}$$

The denominator of (8) being a normalizing factor we focus only on the expectation under Q in the numerator. Write

$$E_Q[f(X_k, Z_k)\Lambda_k^{-1} \mid Y_k] = \int f(x,z)q_k(x,z)dxdz. \tag{9}$$

Theorem 1. The unnormalized conditional joint probability density function of the populations' sizes given by the dynamics in (1) and (2) follows the recursion:

$$q_k(x,z) = A_k(x,z)\int B_k(u,v,x,z)q_{k-1}(u,v)dudv \tag{10}$$

where

$$A_k(x,z) = \frac{2^{R_k+F_k}}{\sigma_1\sigma_2}\sum_{j=0}^{\tilde{\tilde{x}}_k}\sum_{i=0}^{\tilde{\tilde{z}}_k}\binom{\tilde{\tilde{z}}_k}{i}\binom{\tilde{\tilde{x}}_k}{j} p_Z^i(1-p_Z)^{\tilde{\tilde{z}}_k-i}p_X^j(1-p_X)^{\tilde{\tilde{x}}_k-j}$$

$$\times(\frac{j}{x})^{y_k^X}(1-\frac{j}{x})^{F_k-y_k^X}(\frac{i}{z})^{y_k^Z}(1-\frac{i}{z})^{R_k-y_k^Z} \tag{12}$$

$$B_k(u,v,x,z) = \vartheta_k(\frac{z-v-cv+dvu}{\sigma_2})\phi_k(\frac{x-u-avu+bu}{\sigma_1}) \tag{13}$$

(Note we take $0^0 = 1$.)

Proof. In view of Lemma 2 the left hand side of (9) is:

$$= E_Q[f(X_k,Z_k)\Lambda_{k-1}^{-1}\lambda_k^{-1}\,|\,Y_k]$$

$$= 2^{\tilde{\tilde{z}}_k+R_k+\tilde{\tilde{x}}_k+F_k}\frac{1}{2^{\tilde{\tilde{x}}_k}}\frac{1}{2^{\tilde{\tilde{z}}_k}}$$

$$\times E_Q[\sum_{j=0}^{\tilde{\tilde{x}}_k}\sum_{i=0}^{\tilde{\tilde{z}}_k}\binom{\tilde{\tilde{z}}_k}{i}\binom{\tilde{\tilde{x}}_k}{j}p_Z^i(1-p_Z)^{\tilde{\tilde{z}}_k-i}p_X^j(1-p_X)^{\tilde{\tilde{x}}_k-j}$$

$$\times\int f(x,z)(\frac{j}{x})^{y_k^X}(1-\frac{j}{x})^{F_k-y_k^X}(\frac{i}{z})^{y_k^Z}(1-\frac{i}{z})^{R_k-y_k^Z}$$

$$\frac{\vartheta_k(\frac{z-Z_{k-1}-cZ_{k-1}+dZ_{k-1}X_{k-1}}{\sigma_2})}{\sigma_2 v_k(z)}v_k(z)$$

$$\times\frac{\phi_k(\frac{x-X_{k-1}-aZ_{k-1}X_{k-1}+bX_{k-1}}{\sigma_1})}{\sigma_1\chi_k(x)}x_k(x)$$

$$\times p_Z^i(1-p_Z)^{\tilde{\tilde{z}}_k-i}p_X^j(1-p_X)^{\tilde{\tilde{x}}_k-j}dxdz\,\Lambda_{k-1}^{-1}\,|\,Y_k]$$

$$= \frac{2^{R_k+F_k}}{\sigma_1\sigma_2}\sum_{j=0}^{\tilde{\tilde{x}}_k}\sum_{i=0}^{\tilde{\tilde{z}}_k}\binom{\tilde{\tilde{z}}_k}{i}\binom{\tilde{\tilde{x}}_k}{j}p_Z^i(1-p_Z)^{\tilde{\tilde{z}}_k-i}p_X^j(1-p_X)^{\tilde{\tilde{x}}_k-j}$$

$$\times\int f(x,z)(\frac{j}{x})^{y_k^X}(1-\frac{j}{x})^{F_k-y_k^X}(\frac{i}{z})^{y_k^Z}(1-\frac{i}{z})^{R_k-y_k^Z}$$

$$\int\vartheta_k(\frac{z-v-cv+dvu}{\sigma_2})\phi_k(\frac{x-u-avu+bu}{\sigma_1})q_{k-1}(u\,y)dudvdxdz$$

Comparing this last expression with the right hand side of (9) gives the result.

Remark 1. The normalized conditional joint density function of (X_k, Z_k) is simply $\dfrac{q_k(x,z)}{\int q_k(u,v)dudv}$.

The initial (normalized) probability density of (X_0, Z_0), prior to sampling, is $\pi_0(.)$, so $q_0(x,z) = \pi_0(x,z)$.

$$q_1(x,z) = \frac{2^{R_1+F_1}}{\sigma_1\sigma_2}\sum_{j=0}^{\tilde{\tilde{x}}_1}\sum_{i=0}^{\tilde{\tilde{z}}_1}\binom{\tilde{\tilde{z}}_1}{i}\binom{\tilde{\tilde{x}}_1}{j} p_Z^i (1-p_Z)^{\tilde{\tilde{z}}_1 - i} p_X^j (1-p_X)^{\tilde{\tilde{x}}_1 - j}$$
$$\times (\frac{j}{x})^{y_1^X} (1 - \frac{j}{x})^{F_1 - y_1^X} (\frac{i}{z})^{y_1^Z} (1 - \frac{i}{z})^{R_1 - y_1^Z}$$
$$\int \vartheta_1(\frac{z - v - cv + dvu}{\sigma_2})\phi_1(\frac{x - u - avu + bu}{\sigma_1})\pi_0(u.v)dudv.$$

And further estimates follow from (10).

If the distribution of (X_0, Z_0) is a delta function concentrated at (A, B) say

$$q_1(x,z) = \frac{2^{R_1+F_1}}{\sigma_1\sigma_2}\sum_{j=0}^{\tilde{\tilde{x}}_1}\sum_{i=0}^{\tilde{\tilde{z}}_1}\binom{\tilde{\tilde{z}}_1}{i}\binom{\tilde{\tilde{x}}_1}{j} p_Z^i (1-p_Z)^{\tilde{\tilde{z}}_1 - i} p_X^j (1-p_X)^{\tilde{\tilde{x}}_1 - j}$$
$$\times (\frac{j}{x})^{y_1^X} (1 - \frac{j}{x})^{F_1 - y_1^X} (\frac{i}{z})^{y_1^Z} (1 - \frac{i}{z})^{R_1 - y_1^Z}$$
$$\vartheta_1(\frac{z - B - cB + dBA}{\sigma_2})\phi_1(\frac{x - A - aBA + bA}{\sigma_1}).$$

3. Parameter Revision

Here we shall assume, for simplicity, that the random variables v and w in our model are standard normal (means 0 and standard deviations 1).

The EM algorithm, (Baum and Petrie 1966, Dempster *et al*. 1977) is a widely used iterative numerical method for computing maximum likelihood parameter estimates of partially observed models such as linear Gaussian state space models. For such models, direct computation of the MLE is difficult. The EM algorithm has the appealing property that successive iterations yield parameter estimates with non-decreasing values of the likelihood function.

Suppose that we have observations $y_1, \ldots, y_K$ available, where K is a fixed positive integer. Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability measures on (Ω, F), all absolutely continuous with respect to a fixed probability measure P_0. The log-likelihood function for computing an estimate of the parameter θ based on the information available in Y_K is $L_K(\theta) = E_0\left[\log \frac{dP_\theta}{dP_0} \mid Y_K\right]$, and the maximum likelihood estimate (MLE) is defined by $\hat{\theta} \in \mathrm{argmax}_{\theta\in\Theta} L_K(\theta)$.

Let $\hat{\theta}_0$ be the initial parameter estimate. The EM algorithm generates a sequence of parameter estimates $\{\theta_j\}$, $j \geq 1$, as follows. Each iteration of the algorithm consists of two steps:

Step 1. (E-step). Set $\tilde{\theta} = \hat{\theta}_j$ and compute $Q(\theta, \tilde{\theta})$, where

$$Q(\theta, \tilde{\theta}) = E_{\tilde{\theta}}\left[\log \frac{dP_\theta}{dP_{\tilde{\theta}}} \mid Y_K\right].$$

Step 2. (M-step). Find $\hat{\theta}_{j+1} \in \operatorname{argmax}_{\theta \in \Theta} Q(\theta, \theta_j)$.

Using Jensen's inequality it can be shown (see Theorem 1 in Dempster *et al*. 1977) that the sequence of model estimates $\{\hat{\theta}_j,\ j \geq 1\}$ from the EM algorithm are such that the sequence of likelihoods $\{L_K(\hat{\theta}_j)\}$, $j \geq 1$ is monotonically increasing with equality if and only if $\hat{\theta}_{j+1} = \hat{\theta}_j$.

Sufficient conditions for convergence of the EM algorithm are given in Wu (1983). We briefly summarize them here:

1. The parameter space Θ is a subset of some finite dimensional Euclidean space R^r.
2. $\{\theta \in \Theta : L_K(\theta) \geq L_K(\hat{\theta}_0)\}$ is compact for any $L_K(\hat{\theta}_0) > -\infty$
3. L_K is continuous in Θ and differentiable in the interior of Θ.

(As a consequence of (1), (2) and (iii), clearly $L_K(\hat{\theta}_j)$ is bounded from above).

4. The function $Q(\theta, \hat{\theta}_j)$ is continuous in both θ and $\hat{\theta}_j$.

Then by Theorem 2 in Wu (1983), the limit of the sequence EM estimates $\{\hat{\theta}_j\}$ has a stationary point $\overline{\theta}$ of L_K. Also $\{L_K(\hat{\theta}_j)\}$ converges monotonically to $\overline{L}_t = L_t(\overline{\theta})$ for some stationary point $\overline{\theta}$. To make sure that $\overline{L}_t$ is a maximum value of the likelihood, it is necessary to try different initial values $\hat{\theta}_0$.

The model in (1) and (2) is determined by the parameters a, b, c and d which need to be updated as new information is obtained. These parameters are estimated using the expectation maximization (EM) algorithm.

Maximum likelihood estimation of the parameters via the EM algorithm requires computation of the filtered estimates of quantities such as

$$T_k^{(I)} = \sum_{\ell=1}^{k} X_{\ell-1}^2 Z_{\ell-1}^2,$$
$$T_k^{(II)} = \sum_{\ell=1}^{k} X_{\ell-1}^2 Z_{\ell-1},$$
$$T_k^{(III)} = \sum_{\ell=1}^{k} X_\ell X_{\ell-1} Z_{\ell-1},$$
$$T_k^{(IV)} = \sum_{\ell=1}^{k} X_{\ell-1}^2,$$
$$T_k^{(V)} = \sum_{\ell=1}^{k} X_\ell X_{\ell-1},$$
$$S_k^{(I)} = \sum_{\ell=1}^{k} X_{\ell-1} Z_{\ell-1}^2,$$

$$S_k^{(II)} = \sum_{\ell=1}^{k} X_{\ell-1} Z_\ell Z_{\ell-1},$$
$$S_k^{(III)} = \sum_{\ell=1}^{k} Z_\ell^2,$$
$$S_k^{(IV)} = \sum_{\ell=1}^{k} Z_\ell Z_{\ell-1},$$
$$U_k = \sum_{\ell=1}^{k} x_k \text{ and } V_k = \sum_{\ell=1}^{k} z_k.$$

For $M = I,\ldots,V$ write

$$E[T_k^{(M)} I(X_k \in dx, Z_k \in dz) \mid Y_k] = \frac{E_Q[\Lambda_k^{-1} T_k^{(M)} I(X_k \in dx, Z_k \in dz) \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]},$$
$$\beta_k^{(M)}(x,z) = E_Q[\Lambda_k^{-1} T_k^{(M)} I(X_k \in dx, Z_k \in dz) \mid Y_k],$$
$$E[T_k^{(M)} \mid Y_k] = \frac{E_Q[\Lambda_k^{-1} T_k^{(M)} \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]} = \frac{\int \beta_k^{(M)}(x,z)dxdz}{\int q_k(x,z)dxdz} \triangleq \hat{T}_k^{(M)}.$$

For $N = I,\ldots,IV$.

$$E[S_k^{(N)} I(X_k \in dx, Z_k \in dz) \mid Y_k] = \frac{E_Q[\Lambda_k^{-1} S_k^{(N)} I(X_k \in dx, Z_k \in dz) \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]},$$
$$\eta_k^{(N)}(x,z) = E_Q[\Lambda_k^{-1} S_k^{(N)} I(X_k \in dx, Z_k \in dz) \mid Y_k],$$
$$E[S_k^{(N)} \mid Y_k] = \frac{E_Q[\Lambda_k^{-1} S_k^{(N)} \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]} = \frac{\int \eta_k^{(N)}(x,z)dxdz}{\int q_k(x,z)dxdz} \triangleq \hat{S}_k^{(N)}.$$

First we compute ML estimates of the parameters $\theta = (a,b,c,d,\sigma_1,\sigma_2)$ given the history in Y_k. Now the expression for $Q(\theta,\tilde{\theta})$ is derived.

To update the set of parameters from $\tilde{\theta}$ to θ, the following density is introduced $\frac{dP_\theta}{dP_{\tilde{\theta}}}|_{G_k} = \prod_{\ell=1}^{k} \gamma_\ell$,

$$\gamma_\ell = \frac{\tilde{\sigma}_1 \phi(\frac{X_k - X_{k-1} - aZ_{k-1}X_{k-1} + bX_{k-1}}{\sigma_1})}{\sigma_1 \phi(\frac{X_k - X_{k-1} - \tilde{a}Z_{k-1}X_{k-1} + \tilde{b}X_{k-1}}{\tilde{\sigma}_1})} \frac{\tilde{\sigma}_2 \vartheta(\frac{Z_k - Z_{k-1} - cZ_{k-1} + dZ_{k-1}X_{k-1}}{\sigma_2})}{\sigma_2 \vartheta(\frac{Z_k - Z_{k-1} - \tilde{c}Z_{k-1} + \tilde{d}Z_{k-1}X_{k-1}}{\tilde{\sigma}_2})}.$$

Now

$$E_{\tilde{\theta}}\left[\log \frac{dP_\theta}{dP_{\tilde{\theta}}}|_{G_k} \mid Y_k\right] = -k\log\sigma_1 - k\log\sigma_2 - \frac{1}{2}E_{\tilde{\theta}}\left[\sum_{\ell=1}^{k}(\frac{X_k - X_{k-1} - aZ_{k-1}X_{k-1} + bX_{k-1}}{\sigma_1})^2 \mid Y_k\right]$$
$$-\frac{1}{2}E_{\tilde{\theta}}\left[\sum_{\ell=1}^{k}(\frac{Z_k - Z_{k-1} + dZ_{k-1}X_{k-1} - cZ_{k-1}}{\sigma_2})^2 \mid Y_k\right] + R(\tilde{\theta}) = Q(\theta,\tilde{\theta}),$$

where $R(\tilde{\theta})$ does not involve θ.

To implement the M-step set the derivatives $\frac{\partial Q}{\partial \theta} = 0$. This yields

$$a = \frac{\hat{T}_k^{(III)} + \hat{T}_k^{(II)}(b-1)}{\hat{T}_k^{(I)}},$$

$$b - 1 = \frac{\hat{T}_k^{(III)}\hat{T}_k^{(II)} - \hat{T}_k^{(V)}\hat{T}_k^{(I)}}{\hat{T}_k^{(IV)}\hat{T}_k^{(I)} - (\hat{T}_k^{(II)})^2},$$

$$c = \frac{\hat{S}_k^{(IV)} - \hat{S}_k^{(III)} + d\hat{S}_k^{(I)}}{\hat{S}_k^{(III)}},$$

$$d = \frac{2\hat{S}_k^{(III)}\hat{S}_k^{(I)} - \hat{S}_k^{(IV)}\hat{S}_k^{(I)} + \hat{S}_k^{(II)}}{(\hat{S}_k^{(I)})^2 - \hat{T}_k^{(I)}\hat{S}_k^{(III)}}$$

$$\sigma_1^2 = \frac{1}{k} E_{\tilde{\theta}}\left[\sum_{\ell=1}^{k} (X_k - X_{k-1} - aZ_{k-1}X_{k-1} + bX_{k-1})^2 \mid Y_k\right],$$

$$\sigma_2^2 = \frac{1}{k} E_{\tilde{\theta}}\left[\sum_{\ell=1}^{k} (Z_k - Z_{k-1} + dZ_{k-1}X_{k-1} - cZ_{k-1})^2 \mid Y_k\right].$$

For any "test" function g, write

$$E_Q[\Lambda_k^{-1} T_k^{(M)} g(X_k, Z_k) \mid Y_k] = \int \beta_k^{(M)}(x,z) g(x,z)\,dxdz, \quad M = I,\ldots,V.$$

$$E_Q[\Lambda_k^{-1} S_k^{(N)} g(X_k, Z_k) \mid Y_k] = \int \eta_k^{(N)}(x,z) g(x,z)\,dxdz, \quad N = I,\ldots,IV.$$

Theorem 1. For $k \geq 1$, the unnormalized densities $\beta_k^{(M)}(x,z)$, $\eta_k^{(N)}(x,z)$, $M = I,\ldots,V$, $N = I,\ldots,IV$ are given by the following recursions.

$$\beta_k^{(I)}(x,z) = A_k(x,z)[\int B_k(u,v,x,z)\beta_{k-1}^{(I)}(u,v)\,dudv$$
$$+ \int B_k(u,v,x,z)u^2v^2 q_{k-1}(u,v)\,dudv\,]$$

$$\beta_k^{(II)}(x,z) = A_k(x,z)[\int B_k(u,v,x,z)\beta_{k-1}^{(I)}(u,v)\,dudv$$
$$+ \int B_k(u,v,x,z)u^2 v q_{k-1}(u,v)\,dudv\,],$$

$$\beta_k^{(III)}(x,z) = A_k(x,z)[\int B_k(u,v,x,z)\beta_{k-1}^{(I)}(u,v)\,dudv$$

$$+x\int B_k(u,v,x,z)uvq_{k-1}(u,v)dudv\,],$$

$$\beta_k^{(IV)}(x,z)=A_k(x,z)[\int B_k(u,v,x,z)\beta_{k-1}^{(I)}(u,v)dudv$$
$$+\int B_k(u,v,x,z)u^2q_{k-1}(u,v)dudv\,]$$

$$\beta_k^{(V)}(x,z)=A_k(x,z)[\int B_k(u,v,x,z)\beta_{k-1}^{(I)}(u,v)dudv$$
$$+x\int B_k(u,v,x,z)uq_{k-1}(u,v)dudv\,],$$

$$\eta_k^{(I)}(x,z)=A_k(x,z)[\int B_k(u,v,x,z)\eta_{k-1}^{(I)}(u,v)dudv$$
$$+\int B_k(u,v,x,z)uv^2q_{k-1}(u,v)dudv\,]$$

$$\eta_k^{(II)}(x,z)=A_k(x,z)[\int B_k(u,v,x,z)\eta_{k-1}^{(II)}(u,v)dudv$$
$$+z\int B_k(u,v,x,z)uvq_{k-1}(u,v)dudv\,]$$

$$\eta_k^{(III)}(x,z)=A_k(x,z)[\int B_k(u,v,x,z)\eta_{k-1}^{(III)}(u,v)dudv$$
$$+z^2q_k(x,z)]$$

$$\eta_k^{(IV)}(x,z)=A_k(x,z)[\int B_k(u,v,x,z)\eta_{k-1}^{(IV)}(u,v)dudv$$
$$+z\int B_k(u,v,x,z)vq_{k-1}(u,v)dudv\,]$$

where $A_k(x,z)$ and $B_k(u,v,x,z)$ are given in (12) and (13) and q_k is given recursively in (10).

Proof. First note that $T_k^{(I)}=T_{k-1}^{(I)}+X_{k-1}^2Z_{k-1}^2$. Therefore

$$=E_Q[\Lambda_k^{-1}T_k^{(I)}g(X_k,Z_k)\mid Y_k]$$
$$=E_Q[\Lambda_k^{-1}T_{k-1}^{(I)}g(X_k,Z_k)\mid Y_k]+E_Q[\Lambda_k^{-1}X_{k-1}^2Z_{k-1}^2g(X_k,Z_k)\mid Y_k].$$

The first expectation is simply

$$E_Q[\Lambda_k^{-1}T_{k-1}^{(I)}g(X_k,Z_k)\,|\,Y_k]$$
$$=\frac{2^{R_k+F_k}}{\sigma_1\sigma_2}\sum_{j=0}^{\approx x_k}\sum_{i=0}^{\approx z_k}\binom{\approx z_k}{i}\binom{\approx x_k}{j}p_Z^i(1-p_Z)^{\approx z_k-i}p_X^j(1-p_X)^{\approx x_k-j}$$
$$\times\int g(x,z)(\frac{j}{x})^{y_k^X}(1-\frac{j}{x})^{F_k-y_k^X}(\frac{i}{z})^{y_k^Z}(1-\frac{i}{z})^{R_k-y_k^Z}$$

$$\int \vartheta_k \left(\frac{z - v - cv + dvu}{\sigma_2}\right) \phi_k \left(\frac{x - u - avu + bu}{\sigma_1}\right)$$
$$\times \beta_{k-1}^{(1)}(u, v) du dv dx dz .$$

And

$$E_Q[\Lambda_k^{-1} X_{k-1}^2 Z_{k-1}^2 g(X_k, Z_k) \mid Y_k]$$
$$= \frac{2^{R_k + F_k}}{\sigma_1 \sigma_2} \sum_{j=0}^{\tilde{\tilde{x}}_k} \sum_{i=0}^{\tilde{\tilde{z}}_k} \binom{\tilde{\tilde{z}}_k}{i} \binom{\tilde{\tilde{x}}_k}{j} p_Z^i (1 - p_Z)^{\tilde{\tilde{z}}_k - i} p_X^j (1 - p_X)^{\tilde{\tilde{x}}_k - j}$$
$$\times \int g(x,z) \left(\frac{j}{x}\right)^{y_k^X} \left(1 - \frac{j}{x}\right)^{F_k - y_k^X} \left(\frac{i}{z}\right)^{y_k^Z} \left(1 - \frac{i}{z}\right)^{R_k - y_k^Z}$$
$$\int \vartheta_k \left(\frac{z - v - cv + dvu}{\sigma_2}\right) \phi_k \left(\frac{x - u - avu + bu}{\sigma_1}\right)$$
$$\times u^2 v^2 q_{k-1}(u, v) du dv dx dz$$

Using (12) and (13) yields the result.

Write

$$E[U_k I(X_k \in dx, Z_k \in dz) \mid Y_k] = \frac{E_Q[\Lambda_k^{-1} U_k I(X_k \in dx, Z_k \in dz) \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]},$$
$$\zeta_k(x,z) = E_Q[\Lambda_k^{-1} U_k I(X_k \in dx, Z_k \in dz) \mid Y_k],$$
$$E[U_k \mid Y_k] = \frac{E_Q[\Lambda_k^{-1} U_k \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]} = \frac{\int \zeta_k(x,z) dx dz}{\int q_k(x,z) dx dz} \triangleq \widehat{U}_k ,$$
$$E[V_k I(X_k \in dx, Z_k \in dz) \mid Y_k] = \frac{E_Q[\Lambda_k^{-1} V_k I(X_k \in dx, Z_k \in dz) \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]},$$
$$\xi_k(x,z) = E_Q[\Lambda_k^{-1} V_k I(X_k \in dx, Z_k \in dz) \mid Y_k],$$
$$E[V_k \mid Y_k] = \frac{E_Q[\Lambda_k^{-1} V_k \mid Y_k]}{E_Q[\Lambda_k^{-1} \mid Y_k]} = \frac{\int \xi_k(x,z) dx dz}{\int q_k(x,z) dx dz} \triangleq \widehat{V}_k .$$

To update the parameter from $\tilde{p}_X$ to p_X, the following density is introduced

$$\frac{dP_{p_X}}{dP_{p_X}} \Big|_{G_k} = \prod_{\ell=1}^{k} \left(\frac{p_X}{\tilde{p}_X}\right)^{x_k} \left(\frac{1 - p_X}{1 - \tilde{p}_X}\right)^{\tilde{\tilde{x}}_k - x_k} .$$

Now

$$E_{p_X}\left[\log\frac{dP_{p_X}}{dP_{p_X}}|_{G_k}|Y_k\right]=\widehat{U}_k\log p_X+\sum_{\ell=1}^{k}\tilde{\tilde{x}}_\ell\log(1-p_X)+\widehat{U}_k\log(1-p_X)$$
$$+R(\tilde{p}_X)=Q(p_X,\tilde{p}_X),$$

where $R(\tilde{p}_X)$ does not involve p_X.

To implement the M-step set the derivatives $\frac{\partial Q}{\partial p_X}=0$. This yields

$$p_X=\frac{\widehat{U}_k}{\sum_{\ell=1}^{k}\tilde{\tilde{x}}_\ell}.$$

A similar argument yields:

$$p_Z=\frac{\widehat{V}_k}{\sum_{\ell=1}^{k}\tilde{\tilde{z}}_\ell}.$$

For any "test" function g, write

$$E_Q[\Lambda_k^{-1}U_k g(X_k,Z_k)\mid Y_k]=\int\zeta_k(x,z)g(x,z)dxdz.$$
$$E_Q[\Lambda_k^{-1}V_k g(X_k,Z_k)\mid Y_k]=\int\xi_k(x,z)g(x,z)dxdz.$$

Theorem 2. For $k\geq 1$, the unnormalized densities $\zeta_k(x,z)$ and $\xi_k(x,z)$ are given by the following recursions.

$$\zeta_k(x,z)=A_k(x,z)\int B_k(u,v,x,z)\zeta_{k-1}(u,v)dudv$$
$$+\frac{\tilde{\tilde{x}}_k}{2}q_k(x,z)$$
$$\xi_k(x,z)=A_k(x,z)\int B_k(u,v,x,z)\xi_{k-1}(u,v)dudv$$
$$+\frac{\tilde{\tilde{z}}_k}{2}q_k(x,z)$$

Proof. First note that $U_k=U_{k-1}+x_k$. Therefore

$$=E_Q[\Lambda_k^{-1}U_k g(X_k,Z_k)\mid Y_k]$$
$$=E_Q[\Lambda_k^{-1}U_{k-1}g(X_k,Z_k)\mid Y_k]+E_Q[\Lambda_k^{-1}x_k g(X_k,Z_k)\mid Y_k].$$

The first expectation is simply

$$
\begin{aligned}
&E_Q[\Lambda_k^{-1} U_{k-1} g(X_k, Z_k) \mid Y_k] \\
&= \frac{2^{R_k+F_k}}{\sigma_1 \sigma_2} \sum_{j=0}^{\approx{x}_k} \sum_{i=0}^{\approx{z}_k} \binom{\approx{z}_k}{i} \binom{\approx{x}_k}{j} p_Z^i (1-p_Z)^{\approx{z}_k - i} p_X^j (1-p_X)^{\approx{x}_k - j} \\
&\times \int g(x,z) \left(\frac{j}{x}\right)^{y_k^X} \left(1-\frac{j}{x}\right)^{F_k - y_k^X} \left(\frac{i}{z}\right)^{y_k^Z} \left(1-\frac{i}{z}\right)^{R_k - y_k^Z} \\
&\int \vartheta_k\left(\frac{z - v - cv + dvu}{\sigma_2}\right) \phi_k\left(\frac{x - u - avu + bu}{\sigma_1}\right) \\
&\times \zeta_{k-1}(u,v)\, du\, dv\, dx\, dz.
\end{aligned}
$$

from which the result follows.

Remark. The above theorems do not require v_ℓ and w_ℓ to be Gaussian.

4. References

AGGOUN, L. and ELLIOTT, R.J. 1998. Recursive Estimation in Capture-Recapture Methods, *Journal of Science and Technology, Sultan Qaboos University* **3:** 67–75.

BAUM, L.E. and PETRIE, T. 1966. Statistical Inference for Probabilistic Functions of Finite State Markov Chains, *Annals of the Institute of Statistical Mathematics* **37:** 1554-1563.

DEMPSTER, A.P., LAIRD, N.M. and RUBIN, D.B. 1977. *Maximum Likelihood from Incomplete Data via the EM Algorithm*, Jour. of the Royal Statistical Society, B, **39:** 1-38.

ELLIOTT, R.J., AGGOUN, L. and MOORE, J.B. 1995. *Hidden Markov Models : Estimation and Control*. Applications of Mathematics. Vol. 29 Springer-Verlag, Berlin-Heidelberg-New York.

LOTKA, A., J. 1925. *Elements of physical biology*. Baltimore: Williams & Wilkins Co.

SEBER, G.A.F. 1982. The *Estimation of Animal Abundance and Related Parameters*. Second Edition. Edward Arnold.

VOLTERRA, V. 1926. *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. Mem. R. Accad. Naz. dei Lincei. Ser. VI, vol. 2.

WU, C.F.J. 1983. On the Convergence Properties of the EM algorithm, *The Annals of Statistics* 11, pp. 95-103.

Received 25 February 2004
Accepted 26 September 2004