Brian Dennis, José Miguel Ponciano, Mark L. Taper. 2010. Replicated sampling increases efficiency in monitoring biological populations. ***Ecology*** **91:610-620.**

Ecological Archives **E091-044**

Appendix A. Multivariate-normal likelihood function for replicated sampling.

Here we construct the multivariate-normal likelihood function for replicated sampling under the Gompertz state-space (GSS) model, for the stationary case. The likelihood function allows maximum-likelihood (ML) estimation of model parameters through numerical maximization. In addition, we obtain a multivariate-normal likelihood for use in obtaining restricted maximum-likelihood (REML) estimates with transformed observations.

We assume the sampling process is replicated p_t times at sampling time t, producing observations Y_{1t}, Y_{2t}, ..., $Y_{p_t t}$. Denote by $\mathbf{Y}_t$ the $p_t \times 1$ column vector [Y_{1t}, Y_{2t}, ..., $Y_{p_t t}$]′ of the observations (as random variables) at time t, and denote by $\mathbf{y}_t$ the $p_t \times 1$ column vector [y_{1t}, y_{2t}, ..., $y_{p_t t}$]′ of the recorded outcomes (data values) of the random variables in the vector $\mathbf{Y}_t$ at time t. We write **j** for a column vector of ones, **o** for a column vector of zeros, **J** for a matrix of ones, **O** for a matrix of zeros, and **I** for an identity matrix, with the sizes determined by context (or clarified with subscripts, when necessary) so as to be consistent with the matrix operations. The GSS model consists of the underlying population process joined with the multivariate sampling process:

$$X_t = a + cX_{t-1} + E_t$$

$$\mathbf{Y}_t = \mathbf{j}X_t + \mathbf{F}_t$$

where $E_t \sim$ normal(0, σ^2) and $\mathbf{F}_t \sim$ MVN($\mathbf{o}_t$, $\tau^2 \mathbf{I}$), with $\mathbf{F}_t$ assumed independent of E_t and X_t, and no autocorrelation of the noise processes E_t and $\mathbf{F}_t$. Also, denote by $\mathbf{X}$ the $(q+1)\times 1$ column vector $[X_0, X_1, \ldots, X_q]'$, and denote by $\mathbf{Y}$ the $r \times 1$ column vector ($r = p_0 + p_1 + \cdots + p_q$) formed by stacking the vectors $\mathbf{Y}_0, \mathbf{Y}_1, \ldots, \mathbf{Y}_q$. Similarly denote by $\mathbf{F}$ the $r \times 1$ column vector formed by stacking the vectors $\mathbf{F}_0, \mathbf{F}_1, \ldots, \mathbf{F}_q$.

ML estimation

The essential idea is that $\mathbf{X}$ has a multivariate-normal distribution, and $\mathbf{Y}$ is the sum of $\mathbf{F}$ and a linear transformation of $\mathbf{X}$. First, a well-known property of the stationary AR(1) process X_t is that

$$\mathbf{X} \sim \text{MVN}(\,\mathbf{j}a/(1-c)\,,\, \boldsymbol{\Sigma}\,)$$

with all the main diagonal elements of the variance--covariance matrix $\boldsymbol{\Sigma}$ equal to the stationary variance $\mathrm{V}(X_t) = \sigma^2/(1-c^2)$, and the other elements giving the stationary covariances $\mathrm{CV}(X_t\,,\, X_{t+s}) = |c|^s \sigma^2/(1-c^2)$ (see for instance, Dennis et al. 2006, Harvey 1993). Second, let a matrix $\mathbf{C}$ be defined by stacking column vectors of ones and zeros in the following manner:

$$\mathbf{C} = \begin{bmatrix} \mathbf{j}_0 & \mathbf{o}_0 & \cdots & \mathbf{o}_0 \\ \mathbf{o}_1 & \mathbf{j}_1 & \cdots & \mathbf{o}_1 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{o}_q & \mathbf{o}_q & \cdots & \mathbf{j}_q \end{bmatrix}.$$

Here $\mathbf{j}_t$ and $\mathbf{o}_t$ are $p_t \times 1$ (the size of $\mathbf{y}_t$). One can see that $\mathbf{Y}$ is a linear transformation of $\mathbf{X}$:

$$\mathbf{Y} = \mathbf{C}\mathbf{X} + \mathbf{F}\,.$$

Therefore,

$$\mathbf{Y} \sim \mathrm{MVN}(\ \mathbf{Cj}\,a/(1-c)\ ,\ \mathbf{C\Sigma C'} + \tau^2\,\mathbf{I}\ )\ .$$

The log-likelihood function for a vector of data $\mathbf{y}$ thus is the log-probability density for a multivariate-normal distribution:

$$\ln L(a, c, \sigma^2, \tau^2) \ = \ -\frac{r}{2}\ln(2\pi) \ - \ \frac{1}{2}\ln|\mathbf{V}| - \ \frac{1}{2}\,(\mathbf{y}-\boldsymbol{\mu})'\,\mathbf{V}^{-1}\,(\mathbf{y}-\boldsymbol{\mu})\ ,$$

where $\mathbf{V} \ = \ \mathbf{C\Sigma C'} + \ \tau^2\,\mathbf{I}$, and $\boldsymbol{\mu}$ is a $r \times 1$ vector with all elements equal to $a/(1-c)$.

REML estimation

A REML transformation of the observations in replicated sampling can be defined as follows. Multiply each element in the vector $\mathbf{y}_t$, (t = 1, 2,…, q) by p_{t-1} (the size of the previous vector), and then subtract the sum of the elements in the previous vector $\mathbf{y}_{t-1}$, that is,

$$w_{i\,t} \ = \ p_{t-1}y_{i\,t} - \left(y_{1\ t-1} + y_{2\ t-1} + \ldots + y_{p_{t-1}\ t-1}\right).$$

Because all the observations have the same mean (the stationary mean $a/(1-c)$), each $w_{i\,t}$ arises from a distribution with mean zero. Let $\mathbf{w}_t = [\,w_{1\,t}\,,\ w_{2\,t}\,,\ \ldots,\ w_{p_t\,t}\,]'$. Then $\mathbf{w}_t$ can be represented by

$$\mathbf{w}_t \ = p_{t-1}\,\mathbf{y}_t - \mathbf{j}_t\,\mathbf{j}_{t-1}'\,\mathbf{y}_{t-1}\ .$$

Furthermore, let $\mathbf{w}$ be the vectors $\mathbf{w}_1$, $\mathbf{w}_2$, …, $\mathbf{w}_q$ stacked into a column vector, and denote by $\mathbf{W}$ the random vector version (of which $\mathbf{w}$ is a particular realization). Using the transformation matrix given by

$$\mathbf{D} \ = \begin{bmatrix} -\mathbf{J}_{10} & p_0\mathbf{I}_{11} & \mathbf{O}_{12}\cdots & \mathbf{O}_{1q} \\ \mathbf{O}_{20} & -\mathbf{J}_{21} & p_1\mathbf{I}_{22}\cdots & \mathbf{O}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O}_{q0} & \mathbf{O}_{q1} & \cdots & p_{q-1}\mathbf{I}_{qq} \end{bmatrix}$$

where $\mathbf{J}_{st}$, $\mathbf{O}_{st}$ and $\mathbf{I}_{st}$ are $p_s \times p_t$ matrices, we establish that $\mathbf{W} = \mathbf{DY}$ is a $(r - p_0) \times 1$ vector that has a multivariate-normal distribution with a mean vector of zero and a variance--covariance matrix given by $\mathbf{\Phi} = \mathbf{D}[\mathbf{C\Sigma C'} + \tau^2 \mathbf{I}]\,\mathbf{D'} = \mathbf{DVD'}$. The restricted log-likelihood function for $\mathbf{w}$ is then:

$$\ln L(c, \sigma^2, \tau^2) = -\frac{r - p_0}{2} \ln(2\pi) - \frac{1}{2} \ln|\mathbf{\Phi}| - \frac{1}{2} \mathbf{w}' \mathbf{\Phi}^{-1} \mathbf{w} .$$

The parameter a does not appear in the restricted log-likelihood. The restricted log-likelihood is maximized numerically over the values of the parameters c, σ^2, and τ^2. The parameter a is then estimated as

$$a = (1 - c)\frac{\mathbf{j'V^{-1}y}}{\mathbf{j'V^{-1}j}}$$

where everything on the right-hand side of the equation is evaluated at the REML estimates for c, σ^2, and τ^2.

LITERATURE CITED

Dennis, B., J. M. Ponciano, S. R. Lele, M. L. Taper, and D. F. Staples. 2006. Estimating density dependence, process noise, and observation error. Ecological Monographs 76:323–341.

Harvey, A. C. 1993. Time series models. Second edition. MIT Press, Cambridge, Massachusetts, USA.