

Stability properties for MAR(1) models

Mark D. Scheuerell

U.S. Geological Survey
Washington Cooperative Fish and Wildlife Research Unit
School of Aquatic and Fishery Sciences
University of Washington
Seattle, WA USA
scheuerl@uw.edu

Background

There is growing interest in the use of first-order vector autoregressive models in ecology where they are often referred to as multivariate autoregressive, or MAR(1), models¹. In particular, MAR(1) models have been used to estimate interaction strengths among various members in a food web and evaluate overall community stability. The general form of a MAR(1) model is

$$\mathbf{x}_t = \mathbf{m} + \mathbf{B}(\mathbf{x}_{t-1} - \mathbf{m}) + \mathbf{w}_t, \quad (1)$$

where \mathbf{x}_t is an $n \times 1$ vector of state variates at time t , \mathbf{m} is an $n \times 1$ vector containing the means of the rows of the $x \times t$ matrix \mathbf{x} , \mathbf{B} is an $n \times n$ transition matrix, and \mathbf{w}_t is an $n \times 1$ vector of process errors. Here I assume the values in \mathbf{w}_t are drawn from a multivariate normal distribution, such that $\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})$.

In ecological applications of MAR(1) models, the $x_{i,t}$ values within \mathbf{x}_t represent the log-density (or biomass) of a species or functional group at time t . Each of the b_{ij} values within \mathbf{B} represents the effect of species j on species i , such that the values along the diagonal of \mathbf{B} (i.e., where $i = j$) give strength of density-dependence. In addition, for computational purposes the data are often standardized to have a mean of zero ($\mathbf{x}_i - \bar{x}_i$), such that \mathbf{m} in Eqn (1) is then set to zero. In that case, the MAR(1) model becomes

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{w}_t \quad (2)$$

This MAR(1) model can also be used as part of a state-space model, wherein the observed data \mathbf{y} are an imperfect sample of the true realizations \mathbf{x} , such that

$$\mathbf{y}_t = \mathbf{x}_t + \mathbf{a} + \mathbf{v}_t, \quad (3)$$

\mathbf{a} is an $n \times 1$ vector of offsets (levels) to account for non-zero means in \mathbf{y} , and \mathbf{v}_t is an $n \times 1$ vector of observation errors. When combined, Eqns (1) and (3) form a full state-space model:

¹Ives *et al.* (2003) *Ecol Monogr* 73:301–330

$$\begin{aligned}\mathbf{x}_t &= \mathbf{B}\mathbf{x}_{t-1} + \mathbf{w}_t \\ \mathbf{y}_t &= \mathbf{x}_t + \mathbf{a} + \mathbf{v}_t\end{aligned}\tag{4}$$

Following Holmes *et al.*², I will refer to (4) as a MARSS model, where SS stands for state-space, and we drop the parenthetic lag-1 identifier for convenience.

Variance of the stationary distribution

Here I will consider only stationary models wherein all of the eigenvalues of \mathbf{B} lie outside the unit circle. Under this assumption, the probability distribution of \mathbf{x}_t as $t \rightarrow \infty$ is called the *stationary distribution*. Let's define $\mathbf{\Sigma}$ to be the covariance matrix of the stationary distribution, so that it follows from Eqn (2) that

$$\mathbf{\Sigma} = \mathbf{B}\mathbf{\Sigma}\mathbf{B}^\top + \mathbf{Q}.\tag{5}$$

Unfortunately, there is no closed-form solution for $\mathbf{\Sigma}$ when written in this form. However, we can use the vec operator to derive an explicit solution for $\mathbf{\Sigma}$, as it converts an $i \times j$ matrix into an $(ij) \times 1$ column vector. For example, if

$$\mathbf{M} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix},$$

then

$$\text{vec}(\mathbf{M}) = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

Thus, if \mathbf{I} is an $n \times n$ identity matrix, and we define $\tilde{\mathbf{I}} = \mathbf{I} \otimes \mathbf{I}$ and $\tilde{\mathbf{B}} = \mathbf{B} \otimes \mathbf{B}$, then

$$\text{vec}(\mathbf{\Sigma}) = (\tilde{\mathbf{I}} - \tilde{\mathbf{B}})^{-1} \text{vec}(\mathbf{Q}).\tag{6}$$

Importance of species interactions to stability

Among the many ways of classifying stability, I am interested in the extent to which community interactions, relative to environmental forcing, contribute to the overall variance of the stationary distribution. It is helpful to first consider a model for a single species given by

$$x_t = bx_{t-1} + w_t\tag{7}$$

where b is the strength of density dependence, the stationary variance of x_t is σ , and the variance of w_t is q . Analogous to Eqn (5), the variance of both sides of Eqn (7) is given by

²Holmes *et al.* (2012) *The R Journal* 4:11-19

$$\sigma = b^2\sigma + q. \quad (8)$$

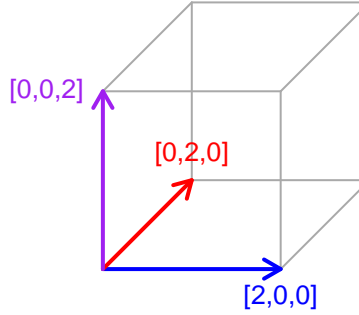
We can rearrange the terms in Eqn (8) to express the proportion of σ owing to the one species' interaction with itself, π_b , as

$$\pi_b = \frac{\sigma - q}{\sigma} = b^2. \quad (9)$$

Using matrix determinants

The size or magnitude of a scalar is just the value itself, but in a multidimensional community we need a way to measure the “volume” of the matrices describing the various interactions among the members (\mathbf{B}) and the (co)variances of the distributions ($\mathbf{\Sigma}, \mathbf{Q}$). The determinant of a matrix provides a convenient way to do so³.

For example, consider a cube where the length of each of its sides equals 2 units. The volume of this cube is $2 \times 2 \times 2 = 8$ units³. We can also define this cube with three vectors, $[x \ y \ z]$, each of which defines one of the cube's vertices.



We can combine these three vectors into a square matrix where each of the columns defines a vertex, such that

$$\mathbf{K} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

The determinant of \mathbf{K} gives the volume of this cube. The determinant of a 2×2 matrix is

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

and the determinant of a 3×3 matrix is given by

³Ives *et al.* (2003) *Ecol Monogr* 73:301–330

$$\begin{aligned}\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} &= a \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \det \begin{bmatrix} d & e \\ g & h \end{bmatrix} \\ &= a(ei - fh) + b(fg - di) + c(dh - eg)\end{aligned}$$

Thus, the volume of the cube defined by \mathbf{K} is

$$\begin{aligned}\det(\mathbf{K}) &= 2(2 \cdot 2 - 0 \cdot 0) + 0(0 \cdot 0 - 0 \cdot 2) + 0(0 \cdot 0 - 2 \cdot 0) \\ &= 2(4 - 0) + 0(0 - 0) + 0(0 - 0) \\ &= 8\end{aligned}$$

Looking back to (5), we see that

$$\begin{aligned}\Sigma &= \mathbf{B}\Sigma\mathbf{B}^\top + \mathbf{Q}, \\ \Sigma - \mathbf{Q} &= \mathbf{B}\Sigma\mathbf{B}^\top.\end{aligned}\tag{10}$$

Therefore, the volume of the difference $\Sigma - \mathbf{Q}$ provides a measure of how much species interactions contribute to the variance of the stationary distribution. Taking determinants of both sides, we get

$$\begin{aligned}\det(\Sigma - \mathbf{Q}) &= \det(\mathbf{B}\Sigma\mathbf{B}^\top) \\ &= \det(\mathbf{B})\det(\Sigma)\det(\mathbf{B}^\top) \\ &= \det(\mathbf{B})\det(\Sigma)\det(\mathbf{B}).\end{aligned}$$

By rearranging terms the proportion of the volume of Σ attributable to species interactions, $\pi_{\mathbf{B}}$, is then

$$\pi_{\mathbf{B}} = \frac{\det(\Sigma - \mathbf{Q})}{\det(\Sigma)} = \det(\mathbf{B})^2.\tag{11}$$

The effects of covariates

Often we would also like to infer the possible effects of covariates on the various members of the community. In that case, the MAR(1) model

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{C}\mathbf{c}_t + \mathbf{w}_t,\tag{12}$$

where \mathbf{C} is an $n \times p$ matrix of covariate effects, and \mathbf{c}_t is a $p \times 1$ vector of covariates at time t .

Although the variance of the stationary distribution will be affected by any covariate effects, the method described above for estimating the relative effects of species interactions is unaffected. That is, if Ξ is the variance-covariance matrix of the covariates in \mathbf{c} , then analogous to (10) the variance of the stationary distribution will be

$$\Sigma = \mathbf{B}\Sigma\mathbf{B}^\top + \mathbf{C}\Xi\mathbf{C}^\top + \mathbf{Q},\tag{13}$$

and hence

$$\begin{aligned}\boldsymbol{\Sigma} - \mathbf{C}\boldsymbol{\Xi}\mathbf{C}^\top - \mathbf{Q} &= \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^\top \\ \det(\boldsymbol{\Sigma} - \mathbf{C}\boldsymbol{\Xi}\mathbf{C}^\top - \mathbf{Q}) &= \det(\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^\top) \\ \frac{\det(\boldsymbol{\Sigma} - \mathbf{C}\boldsymbol{\Xi}\mathbf{C}^\top - \mathbf{Q})}{\det(\boldsymbol{\Sigma})} &= \det(\mathbf{B})^2.\end{aligned}$$

Sensitivity of $\pi_{\mathbf{B}}$

We can formally examine the sensitivity of $\pi_{\mathbf{B}}$ to each of the elements in \mathbf{B} via the following relationship:

$$\begin{aligned}\left[\frac{\partial \pi_{\mathbf{B}}}{\partial \mathbf{B}}\right]_{ij} &= \left[\frac{\partial \det(\mathbf{B})^2}{\partial \mathbf{B}}\right]_{ij} \\ &= [2 \det(\mathbf{B})^2 (\mathbf{B}^{-1})^\top]_{ij}.\end{aligned}$$

Thus, we can ask whether the proportion $\pi_{\mathbf{B}}$ is most sensitive to intra- or inter-species interactions, and which ones.

Long-term changes in abundance

Ives *et al.*⁴ provide a means for assessing the expected long-term change in the density (biomass) of species i , L_i , within a community of n total members, owing to the effect of some covariate j . Specifically,

$$L_i = \frac{\det(\mathbf{B}_1 \dots \mathbf{B}_{i-1} \mathbf{C}_j \mathbf{B}_{i+1} \dots \mathbf{B}_n)}{\det(\mathbf{B}_1 \dots \mathbf{B}_n)}, \quad (14)$$

where \mathbf{B}_i is a $n \times 1$ column vector containing the estimated effects of species i on all of the species (including itself), and \mathbf{C}_i is a $n \times 1$ column vector containing the estimated effects of covariate j on each of the species.

Reactivity

We can also calculate the reactivity of a community following an external perturbation, which measures the displacement of the community away from its stationary distribution. There are two methods to do so.

The first follows Ives *et al.*⁵, where reactivity is based upon estimates of the process covariance \mathbf{Q} and stationary covariance $\boldsymbol{\Sigma}$. Specifically, they define reactivity ν as

⁴Ives *et al.* (1999) *Ecology* 80:1405–1421

⁵Ives *et al.* (2003) *Ecol Monogr* 73:301–330

$$\nu = -\frac{\text{tr}(\mathbf{Q})}{\text{tr}(\mathbf{\Sigma})} \quad (15)$$

$$\leq 1 - \lambda_{max}(\mathbf{B}^\top \mathbf{B}). \quad (16)$$

The second method comes from Neubert *et al.*⁶, wherein reactivity is

$$\nu = \log \|\mathbf{B}\|_2 \quad (17)$$

$$= \log \sqrt{\lambda_{max}(\mathbf{B}^\top \mathbf{B})} \quad (18)$$

$$= \log \sigma_{max}(\mathbf{B}), \quad (19)$$

and $\|\cdot\|_2$ is the spectral norm, $\lambda_{max}(\cdot)$ is the maximum eigenvalue, and $\sigma_{max}(\cdot)$ is the largest singular value. This method only requires estimates of the species interactions in \mathbf{B} .

⁶Neubert *et al.* (2009) *Ecology* 90:2683–2688