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), models[1]. 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

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

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 $\mathbf{w}_t \sim \mathrm{MVN}(\mathbf{0}, \mathbf{Q})$.

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.

Often we would also like to infer the possible effects of covariates on the various members of the community. In addition, for computational purposes the data are often standardized to have a mean of zero ($\mathbf{x}_i - \bar{\mathbf{x}}_i$), such that $\mathbf{m}$ in Eqn (1) is then set to zero. In that case, the MAR(1) model becomes

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{C}\mathbf{c}_t + \mathbf{w}_t, \tag{2}$$

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.

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

$$\mathbf{y}_t = \mathbf{x}_t + \mathbf{a} + \mathbf{v}_t, \tag{3}$$

---

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

$\mathbf{a}$ 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:

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{C}\mathbf{c}_t + \mathbf{w}_t$$
$$\mathbf{y}_t = \mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \tag{4}$$

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

We will restrict this treatment to stationary models wherein all of the eigenvalues of $\mathbf{B}$ lie within the unit circle. In that case, the probability distribution of $\mathbf{x}_t$ as $t \to \infty$ is called the stationary distribution. Thus, because $\lim_{t \to \infty} t - 1 = t$, under assumptions of stationarity we can write (1) as

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_t + \mathbf{w}_t. \tag{5}$$

From this, it follows that

$$\mathrm{cov}(\mathbf{x}_t) = \mathbf{B}\mathrm{cov}(\mathbf{x}_t)\mathbf{B}^\top + \mathrm{cov}(\mathbf{w}_t). \tag{6}$$

If we define $\boldsymbol{\Sigma} = \mathrm{cov}(\mathbf{x}_t)$, then

$$\boldsymbol{\Sigma} = \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^\top + \mathbf{Q}. \tag{7}$$

Unfortunately, there is no closed-form solution for $\boldsymbol{\Sigma}$ when written in this form. However, we can use the vec operator to derive an explicit solution for $\boldsymbol{\Sigma}$, 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

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

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

$$\mathrm{vec}(\boldsymbol{\Sigma}) = (\mathcal{I} - \mathcal{B})^{-1}\mathrm{vec}(\mathbf{Q}). \tag{8}$$

---

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

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. In a stable system, any perturbation affecting one or more of the community members does not amplify as it moves throughout the community as a whole, such that the variances in log-densities over time would be driven almost entirely by random environmental variation. Given that same magnitude of random environmental variation, less stable systems are characterized by greater variances in the temporal dynamics of their constituents.

Here we will use determinants to measure the "volume" of a matrix. Looking back to (7), we see that

$$\boldsymbol{\Sigma} = \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^{\top} + \mathbf{Q},$$
$$\boldsymbol{\Sigma} - \mathbf{Q} = \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^{\top}. \tag{9}$$

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

$$\det(\boldsymbol{\Sigma} - \mathbf{Q}) = \det(\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^{\top})$$
$$= \det(\mathbf{B})\det(\boldsymbol{\Sigma})\det(\mathbf{B}^{\top})$$
$$= \det(\mathbf{B})\det(\boldsymbol{\Sigma})\det(\mathbf{B}).$$

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

$$\pi_{\mathbf{B}} = \frac{\det(\boldsymbol{\Sigma} - \mathbf{Q})}{\det(\boldsymbol{\Sigma})} = \det(\mathbf{B})^2. \tag{10}$$

The effects of covariates

Although the variance of the stationary distrbution will be affected by any covariate effects, the method described above for estimating the relative effects of species interactions is unaffected. That is, if $\boldsymbol{\Xi}$ is the variance-covariance matrix of the covariates in $\mathbf{c}$, then analogous to (9) the variance of the stationary distribution will be

$$\boldsymbol{\Sigma} = \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^{\top} + \mathbf{C}\boldsymbol{\Xi}\mathbf{C}^{\top} + \mathbf{Q}, \tag{11}$$

and hence

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

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:

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

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.*[3] 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,

$$L_i = \frac{\det(\mathbf{B}_1 \ldots \mathbf{B}_{i-1} \mathbf{C}_j \mathbf{B}_{i+1} \ldots \mathbf{B}_n)}{\det(\mathbf{B}_1 \ldots \mathbf{B}_n)}, \tag{12}$$

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.*[4], where reactivity is based upon estimates of the process covariance $\mathbf{Q}$ and stationary covariance $\mathbf{\Sigma}$. Specifically, they define reactivity ν as

$$\nu = -\frac{\text{tr}(\mathbf{Q})}{\text{tr}(\mathbf{\Sigma})} \tag{13}$$

$$\leq 1 - \lambda_{max}(\mathbf{B}^{\top}\mathbf{B}). \tag{14}$$

The second method comes from Neubert *et al.*[5], wherein reactivity is

---

[3]Ives AR *et al.* (1999) *Ecology* 80:1405–1421
[4]Ives AR *et al.* (2003) *Ecol Monogr* 73:301–330
[5]Neubert MG *et al.* (2009) *Ecology* 90:2683–2688

$$\nu = \log ||\mathbf{B}||_2 \tag{15}$$

$$= \log \sqrt{\lambda_{max}(\mathbf{B}^\top \mathbf{B})} \tag{16}$$

$$= \log \sigma_{max}(\mathbf{B}), \tag{17}$$

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 $\mathbf{B}$.