Multivariate Autoregressive Modeling for the Environmental Sciences

E. E. Holmes, M. D. Scheuerell, and E. J. Ward

2022-02-20

Contents

Preface

This is material that was developed as part of a course we teach at the University of Washington on applied time series analysis for fisheries and environmental data. You can find our lectures on our course website ATSA.

Book package

The book uses a number of R packages and a variety of fisheries data sets. The packages and data sets can be installed by installing our **atsalibrary** package which is hosted on GitHub:

```
library(devtools)
devtools::install_github("atsa-es/atsalibrary")
```

Authors

The authors are United States federal research scientists. This work was conducted as part of our jobs at the Northwest Fisheries Science Center (NWFSC), a research center for NOAA Fisheries, and the United States Geological Survey, which are United States federal government agencies. E. Holmes and E. Ward are affiliate faculty and M. Scheuerell is an associate professor at the University of Washington.

Links to more code and publications can be found on our academic websites:

- http://faculty.washington.edu/eeholmes
- http://faculty.washington.edu/scheuerl
- http://faculty.washington.edu/warde

Citation

Holmes, E. E., M. D. Scheuerell, and E. J. Ward. Multivariate Autoregressive Modeling for the Environmental Sciences. NOAA Fisheries, Northwest Fisheries Science Center, 2725 Montlake Blvd E., Seattle, WA 98112. Contacts eeholmes@uw.edu, eward@uw.edu, and scheuerl@uw.edu.

Chapter 1

Template

Start with one paragraph overview.

A script with all the R code in the chapter can be downloaded here. The Rmd for this chapter can be downloaded here.

Data and packages

All the data used in the chapter are in the **MARSS** package. Install the package, if needed, and load to run the code in the chapter.

```
library(MARSS)
```

1.1 Introduction

Introdution. Here's how to do an equation with numbering.

$$\begin{gathered}\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \text{ where } \mathbf{v}_t \sim \text{MVN}(0, \mathbf{R}) \\ \mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t \text{ where } \mathbf{w}_t \sim \text{MVN}(0, \mathbf{Q})\end{gathered} \tag{1.1}$$

1.2 New section

New section. Example of an equation with matrix.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}_t = \begin{bmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \\ z_{31} & z_{32} & z_{33} \\ z_{41} & z_{42} & z_{43} \\ z_{51} & z_{52} & z_{53} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_t + \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}_t . \tag{1.2}$$

and the process model would look like

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{t-1} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}_t \tag{1.3}$$

The observation errors would be

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}_t \sim \text{MVN} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{12} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{13} & r_{23} & r_{33} & r_{34} & r_{35} \\ r_{14} & r_{24} & r_{34} & r_{44} & r_{45} \\ r_{15} & r_{25} & r_{35} & r_{45} & r_{55} \end{bmatrix} \right) \tag{1.4}$$

And the process errors would be

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}_t \sim \text{MVN} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \right). \tag{1.5}$$

1.3 Section 3

Another section.

1.4 Discussion

For your homework this week, we will continue to investigate common trends in the Lake Washington plankton data.

1. Fit other DFA models to the phytoplankton data with varying numbers of trends from 1-4 (we fit a 3-trend model above). Do not include any covariates in these models. Using `R="diagonal and unequal"` for the observation errors, which of the DFA models has the most support from the data?

   Plot the model states and loadings as in Section 1.3. Describe the general patterns in the states and the ways the different taxa load onto those trends.

2. How does the best model from Question 1 compare to a DFA model with the same number of trends, but with `R="unconstrained"`?

   Plot the model states and loadings as in Section 1.2. Describe the general patterns in the states and the ways the different taxa load onto those trends.

   Also plot the the model fits as in Section 1.3. Do they reasonable? Are there any particular problems or outliers?

3. Fit a DFA model that includes temperature as a covariate and 3 trends (as in Section **??**), but with`R="unconstrained"`? How does this model compare to the model with `R="diagonal and unequal"`? How does it compare to the model in Question 2?

   Plot the model states and loadings as in Section 1.2. Describe the general patterns in the states and the ways the different taxa load onto those trends.

Chapter 2

Univariate models

Start with one paragraph overview.

A script with all the R code in the chapter can be downloaded here. The Rmd for this chapter can be downloaded here.

Data and packages

All the data used in the chapter are in the **MARSS** package. Install the package, if needed, and load to run the code in the chapter.

```
library(MARSS)
```

2.1 Introduction

Introdution. Here's how to do an equation with numbering.

$$\begin{gathered}\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \text{ where } \mathbf{v}_t \sim \text{MVN}(0, \mathbf{R}) \\ \mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t \text{ where } \mathbf{w}_t \sim \text{MVN}(0, \mathbf{Q})\end{gathered} \tag{2.1}$$

2.2 New section

New section. Example of an equation with matrix.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}_t = \begin{bmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \\ z_{31} & z_{32} & z_{33} \\ z_{41} & z_{42} & z_{43} \\ z_{51} & z_{52} & z_{53} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_t + \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}_t . \tag{2.2}$$

and the process model would look like

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{t-1} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}_t \tag{2.3}$$

The observation errors would be

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}_t \sim \text{MVN}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{12} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{13} & r_{23} & r_{33} & r_{34} & r_{35} \\ r_{14} & r_{24} & r_{34} & r_{44} & r_{45} \\ r_{15} & r_{25} & r_{35} & r_{45} & r_{55} \end{bmatrix} \right) \tag{2.4}$$

And the process errors would be

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}_t \sim \text{MVN}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \right). \tag{2.5}$$

And here is a plot.

```
plot(1:100)
```

```
plot(1:10)
```

2.3 Section 3

Another section with a plot.

```
plot(1:10)
```

2.4 Discussion

Here is a discussion.