Intro to ARIMA models

FISH 550 – Applied Time Series Analysis   Download Rmd

Eli Holmes

6 Apr 2023

Big picture

Let's imagine that we can describe our data as a combination of the mean trend (m_t) and error.

$$x_t = m_t + e_t$$

Fisheries biologists and ecologists often want to know m_t. *In fact, that is often our main and sole goal.*

But let's say we don't care about m_t. Our only goal is to predict x_{t+1}.

How could we do this?

Approach 1 which we will use in the rest of the course

1. We write a model for m_t and estimate that model from the data.
2. Once we have an estimated model for m_t, we have an estimate of e_t.
3. We can model that error (think AR and MA)
4. With 2 and 3, we can predict the future x_t from x_{t-1}.

Approach 2 Box-Jenkins

The Box-Jenkins approach (ARIMA models) is **totally** different.

1. Keep differencing the data to until you get a new transformed stationary time series $\Delta^d x_t$
2. Any stationary time series can be modeled as a ARMA process (Wold Decomposition). So now fit a ARMA model to $\Delta^d x_t$
3. Using the estimated ARMA, predict $\Delta^d x_{t+1}$
4. Using $\Delta^d x_{t+1}$, x_t, x_{t-1}, x_{t-2}, etc, you can compute x_{t+1}
5. That's the prediction!

What's $\Delta^d x_{t+1}$

$$\Delta^1 x_t = x_t - x_{t-1}$$
$$\Delta^2 x_t = \Delta^1 x_t - \Delta^1 x_{t-1}$$
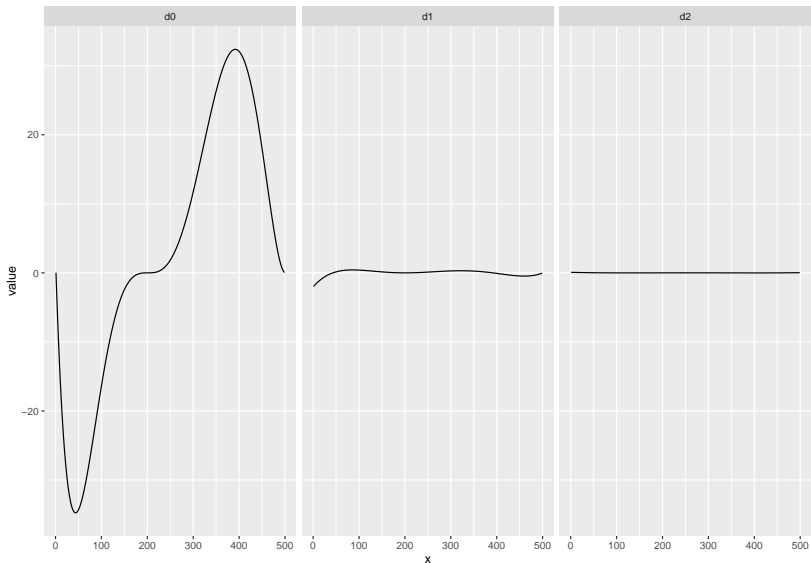$$\Delta^3 x_t = \Delta^2 x_t - \Delta^2 x_{t-1}$$

In Box-Jenkin's approach m_t is gone

In this approach to predicting x_{t+1}, we remove m_t from our data using differencing.

We don't have to worry about a model for m_t because we have removed it!!

How is that possible?

You can remove any wiggly trend with enough differencing.

A biological interpretation is hard for ARIMA models

The error structure of $\Delta^d x_{t+1}$ is NOT the same as e_t.

$$\Delta^d x_t = \phi_1 \Delta^d x_{t-1} + \phi_2 \Delta^d x_{t-2} + \cdots + z_t$$

z_t is the error of the differences. And the ϕ in the AR part are for the differences not the original x_t.

But remember, the objective was to predict x_{t+1} not to fit a model with a biological interpretation.

Box-Jenkins method for fitting ARIMA model

1. Make data stationary by differencing the data
2. Fit a ARMA model to the differenced data
3. Estimate the ARMA parameters
4. Predict x_{t+1}
5. Assess the residuals for problems

Other approaches for non-stationary time series data

ARIMA models are one approach for fitting data that have underlying trends.

Other approaches

- ▶ Regression (we won't cover this)
- ▶ Dynamic Linear Regression (we will cover this)
- ▶ Stochastic level models (we will do a lot of variants of this in class)
- ▶ ARMAX models: $x_t = bx_{t-1} + \beta \text{covariates} + \text{error}$ (we will do some of this)