Lecture 16-17

ARIMA models
Textbook Sections: 6.1-6.4

Motivation

We know that the difference operator $\nabla X_t = X_t - X_{t-1}$ can be used to remove the trend from a process $\{X_t\}$ that is not mean-stationary. If the trend is a polynomial of order d, then applying the difference operator d times would remove the trend perfectly. In general, even if the trend is not exactly polynomial, applying the difference operator (possibly several times) would do a decent job at removing the trend.

The ARIMA framework (also known as the Box-Jenkins framework after its developers) combines the ARMA framework we have covered with the difference operator. Starting with data that has a trend, but is otherwise stationary, we can use differences to remove the trend, then fit an ARMA model to the differenced data.

The framework can also be extended to include seasonal differences, which can be used to remove the seasonal component as well. A seasonal ARIMA model could be used with data that exhibits both a atrend and a seasonal component. We will cover the topic briefly, but will mostly restrict our work to non-seasonal ARIMA models.

Non-seasonal ARIMA models

If $\{X_t\}$ is a zero-mean $ARIMA(p, d, q)$ process, then

$$Y_t = \nabla^d X_t,$$

$$Y_t - \phi_1 Y_{t-1} - \cdots - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q},$$

where $\{\varepsilon_t\}$ is white noise.

Using the backshift operator $B(X_t) = X_{t-1}$, we express the above expressions as

$$Y_t = \nabla^d X_t = (1 - B)^d X_t,$$

$$\phi(B) Y_t = \theta(B) \varepsilon_t,$$

where $\phi(z)$ and $\theta(z)$ are the AR and MA polynomials, respectively.

Finally, we can rewrite the process without using Y_t as

$$(1 - B)^d \phi(B) X_t = \theta(B) \varepsilon_t.$$

When to Fit ARIMA?

In general, it is a good habit to inspect the differenced data along with the raw data, and to keep ARIMA models in mind as an option. Here are three signs that an ARIMA model may be appropriate.

- Analysis of the differenced sample data may reveal that the differenced data is more stationary, or even suggests an $ARMA(p,q)$ structure.
- A *slowly* decaying sample ACF plot is indicative of a trend. It also indicates that differencing may be helpful.
- Unit root tests offer a more systematic approach for detecting if differencing is needed.

Rob Hyndman outlines a more detailed procedure in his textbook.

Unit-Root tests

Recall that we usually work with $ARMA(p,q)$ models that satisfy the following:

- roots of the AR polynomial $\phi(z)$ are outside the unit circle (for stationarity and causality),
- roots of the MA polynomial $\theta(z)$ are outside the unit circle (for invertibility).

Now let's take another look at the $ARIMA(p,d,q)$ process from before:

$$(1-B)^d \phi(B) X_t = \theta(B)\varepsilon_t.$$

If the process is treated as and $ARMA(p,q)$ (instead of $ARIMA(p,d,q)$), then the AR polynomial of this $ARMA(p,q)$ process is

$$\phi^*(z) = (1-z)^d \phi(z).$$

Notice that it has a root $z=1$ of multiplicity d.

This is an important result: if an $ARIMA(p,d,q)$ is (incorrectly) treated as $ARMA(p,q)$, then the AR polynomial will have a unit root of multiplicity d.

Conversely, it can be shown that if the AR polynomial of an $ARMA(p,q)$ process has a root $z=1$, then the AR polynomial of the differenced process will not have this root. In general, **applying ∇ d times will remove the root $z=1$ of multiplicity d.**

This is the motivation behind **unit root tests**. These are hypothesis tests based on the behavior of AR polynomials and their roots. They are used as a formal way of determining whether using the difference operator will make the data stationary.

The R package `tseries` contains functions `adf.test()` and `kpss.test()`, which perform the above two tests.

The R package `forecast` contains the function `ndiffs()`, which uses the ADF or KPSS test (specified by user) to determine the smallest number of differences necessary to make given data stationary.

Seasonal Differences

Regular differencing $\nabla X_t = X_t - X_{t-1}$ can be effective in removing trend.

Seasonal differencing $\nabla_m X_t = X_t - X_{t-m}$ can be used for removing seasonal components as well.

We won't focus on these models, but it's not hard to find resources that explain the concept. The functions in the `forecast` package can be used to fit seasonal models as well.

Forecasting

Point forecasts can be obtained similarly to the ARMA case. We first need to expand the $ARIMA(p, d, q)$ expression, and solve for X_t. Then we need to adjust all the indices. If we're interested in forecasting X_{n+h}, then t should be replaced by $n + h$ in every index. Finally, we plug in observed values, forecasts, and/or estimates of ε_t.

Prediction intervals have complicated expressions, but can be computed with software. Exact descriptions can be found in theoretical texts, such as *Time Series Theories and Methods* by Brockwell and Davis. That book is the graduate-level textbook by the authors of our textbook. It's fairly dense, and is a good reference for time series theory.

A good description of the process is found here.
Keep in mind that the author uses a different notation. He uses y_t for X_t, and T for n.