Date: Tuesday, November 13, 2018.

Non-Stationary models

Idea: Transform a non-stationary time series so that the new series is stationary, and hence can be modeled by an ARMA model.

When the "transformation" involves differencing, what we're really doing is fitting an ARIMA model.

ARIMA

Autoregressive
Integrated
Moving average

ARIMA models model a non-stationary time series which after finitely many "differences" becomes stationary.

If d is a non-negative integer, then $\{Y_t\} \sim ARIMA(p,d,q)$ if:

$$\boxed{X_t = (1-B)^d Y_t \quad \text{is } ARMA(p,q)}$$

This definition implies that an $ARIMA(p,d,q)$ model can be represented as:

$$\phi(B) X_t = \theta(B) \varepsilon_t$$

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

What is differencing actually doing?

Short answer: It eliminates trend.

But, how? Notation: $\nabla Y_t = (1-B) Y_t = Y_t - Y_{t-1}$

$$\begin{aligned} \nabla^2 Y_t &= (1-B)^2 Y_t \\ &= (1-B)(Y_t - Y_{t-1}) \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2}. \end{aligned}$$

Remark: If $\{Y_t\}$ exhibits polynomial trend of the form $m_t = \sum_{i=0}^{d} \alpha_i t^i$, then $\nabla^d Y_t = (1-B)^d Y_t$ will not have any trend.

Example:

$Y_t = c + bt + \varepsilon_t$, where c, b are constants.

$$\nabla Y_t = (1-B)(c+bt+\varepsilon_t)$$

$$= (c - Bc) + (bt - Bbt) + (\varepsilon_t - B\varepsilon_t)$$

$$= (c - c) + (bt - b(t-1)) + (\varepsilon_t - \varepsilon_{t-1})$$

because it is a constant.

$$= bt - bt + b + \varepsilon_t - \varepsilon_{t-1}$$

$$= b + \varepsilon_t - \varepsilon_{t-1}$$

$$= b + \nabla \varepsilon_t.$$

Example:

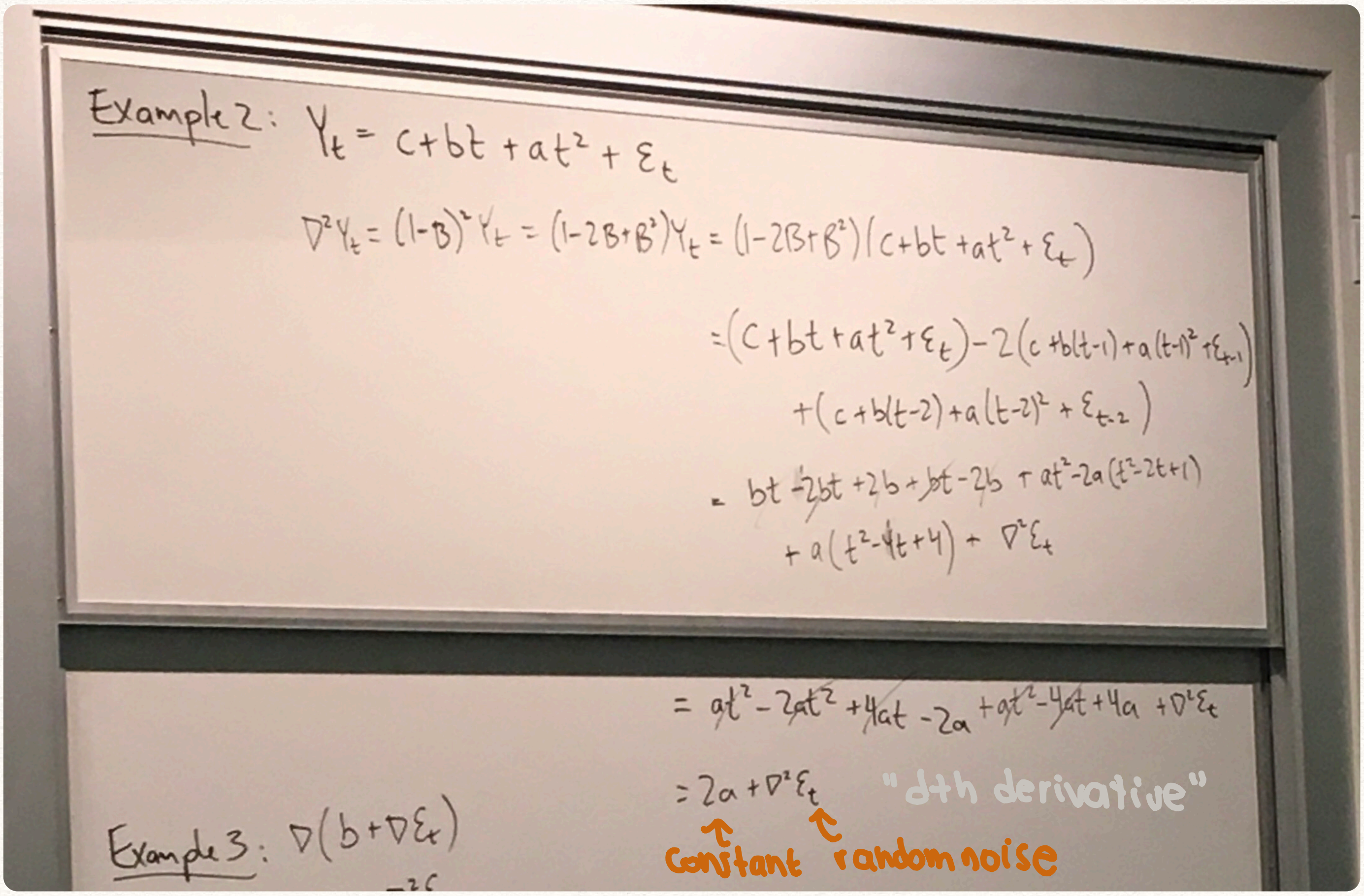
$Y_t = c + bt + at^2 + \varepsilon_t$

$$\nabla^2 Y_t = (1-B)^2 Y_t$$

$$= (1 - 2B + B^2) Y_t$$

$$= (1 - 2B + B^2)(c + bt + at^2 + \varepsilon_t)$$

$$= (c+bt+at^2+\varepsilon_t) - 2B(c+bt+at^2+\varepsilon_t) + B^2(c+bt+at^2+\varepsilon_t)$$



Example: (Be careful with over differenciating)

$$
\begin{aligned}
\nabla(b+\nabla\varepsilon_t) &= \nabla b + \nabla^2\varepsilon_t \\
&= \nabla^2\varepsilon_t \\
&= \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}
\end{aligned}
$$

↑ you incorporate more errors (more noise). Therefore, more difficult to model.

[BREAK]

* How do I choose d? In other words, how many times must I difference my time series before the result is stationary:

Graphically: Plot of the time series.
↳ It shouldn't exhibit trend.
(It could still have seasonality).

OR

ACF plot of the time series
↳ Rapid decay (as opposed to slow decay).

Formally: using "unit" root tests" like the augmented Dickey-Fuller test.

R function: ndiffs

The Augmented Dickey-Fuller test

(often referred to as the ADF test)

H_0: Time series is <u>not</u> stationary. ← The generating function roots lie on or inside the unit circle

vs.

H_A: Time series is stationary. ← The generating function roots lie outside the unit circle.

Idea: Fit an AR(p) model to the data and obtain the estimates $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \ldots, \hat{\phi}_p$ and determine whether these are consistent with what is required by the stationarity conditions.

The choice of p may be determined by the user, but automated implementations of this test by default choose as large as p as can be handled by the sample size.

Example: AR(1): $Y_t = \phi Y_{t-1} + \varepsilon_t$.
Stationarity condition: $|\phi| < 1$
→ Is $\hat{\phi}$ consistent with this condition?

Using a test like this, we choose d by iteratively performing the test. This iteration determines the number of differences necessary to make your resultant time series stationary.