25 Spring 439/639 TSA: Lecture 10

Dr Sergey Kushnarev

Table of contents

More on ARIMA(p, d, q)

Last time, we introduced ARIMA(p, d, q), a model for non-stationary time series. If $Y_t \sim \text{ARIMA}(p, d, q)$, taking difference d times gives a stationary time series $W_t = \nabla^d Y_t \sim \text{ARMA}(p, q)$.

Review the previous example

Let's take another look at the random walk + noise example from last lecture.

$$Y_t = X_t + \eta_t = \sum_{j=1}^{t} e_j + \eta_t, \quad \eta_t \sim \text{iid}(0, \sigma_\eta^2), \quad e_t \sim \text{iid}(0, \sigma_e^2)$$

where (X_t) is a random walk, (η_t) is a sequence of noise, and (η_t) is independent of (e_t). Taking the difference gives

$$W_t = \nabla Y_t = e_t + \eta_t - \eta_{t-1}.$$

We can verify (W_t) is stationary:

$$\mathbb{E}W_t = \mathbb{E}\left[e_t + \eta_t - \eta_{t-1}\right] = 0,$$

$$\text{Var}(W_t) = \text{Var}\left(e_t + \eta_t - \eta_{t-1}\right) = \sigma_e^2 + 2\sigma_\eta^2 \qquad \text{(by independence)},$$

$$\gamma_1 = \text{Cov}(W_t, W_{t-1}) = \text{Cov}\left(e_t + \eta_t - \eta_{t-1},\, e_{t-1} + \eta_{t-1} - \eta_{t-2}\right) = -\text{Cov}(\eta_{t-1}, \eta_{t-1}) = -\sigma_\eta^2,$$

$$\gamma_k = \text{Cov}(W_t, W_{t-k}) = \text{Cov}\left(e_t + \eta_t - \eta_{t-1},\, e_{t-k} + \eta_{t-k} - \eta_{t-k-1}\right) = 0, \quad \text{for all } k \geq 2.$$

So (W_t) is stationary. Then by the reasoning from last time, there exist an uncorrelated stationary process $(\tilde{e}_t)$ (think of $(\tilde{e}_t) \sim \text{iid}\left(0, \tilde{\sigma}_e^2\right)$) and a constant $\tilde{\theta}$ such that

$$W_t = \tilde{e}_t - \tilde{\theta}\,\tilde{e}_{t-1} \sim \text{MA}(1) \implies Y_t \sim \text{IMA}(1, 1) = \text{ARIMA}(0, 1, 1).$$

AR and MA polynomial for ARIMA(p, d, q)

If $Y_t \sim \text{ARIMA}(p, d, q)$, then $W_t = \nabla^d Y_t \sim \text{ARMA}(p, q)$. So it can be characterized by the AR polynomial $\Phi(x)$ and MA polynomial $\Theta(x)$:

$$\Phi(B) W_t = \Theta(B) e_t.$$

Note that $W_t = (1 - B)^d Y_t$, we have

$$\Phi(B)\,(1-B)^d\,Y_t = \Theta(B)\,e_t.$$

So $\Phi^*(x) = \Phi(x)\ (1-x)^d$ can be seen as an AR polynomial for Y_t. **Assume** (W_t) **is causal**, then $\Phi^*(x)$ has $p + d$ roots, with $z = 1$ repeated d times and the other p roots (i.e. the roots of $\Phi(x)$) are all outside the unit disk.

Overdifferencing

In reality, usually $d = 1$ or $d = 2$. If d is too large, this is called overdifferencing, and it has the following issues:

- Leads to more complicated than necessary models.
- Leads to non-invertible models.

For example, consider the random walk $Y_t = \sum_{i=1}^t e_i$. (Y_t) is non-stationary. Take the difference:

$$W_t = \nabla Y_t = Y_t - Y_{t-1} = \sum_{i=1}^{t} e_i - \sum_{i=1}^{t-1} e_i = e_t.$$

So W_t can by modeled by an MA(0), which is stationary (and invertible). If we take the difference one more time:

$$Z_t = \nabla^2 Y_t = W_t - W_{t-1} = e_t - e_{t-1} \sim \text{MA}(1)$$

Although it can still be modeled by an ARMA model MA(1), but it is more complicated than MA(0), and this MA(1) above is not invertible (since $|\theta| = 1$).

ARFIMA(p, d, q)

This part will not be tested, just for your information. The FI in ARFIMA stands for fractionally integrated (recall that the letter I in ARIMA stands for integrated).

For a real number $d \in (0, 0.5)$, we can define the operator ∇^d by a series

$$\nabla^d := (1-B)^d = \sum_{i=0}^{\infty} b_i\, B^i$$

where the coefficients $\{b_i\}$ are determined by d.

It can be used in modelling long-range dependencies. It has slow decaying ACFs (polynomially, not exponentially).

GLP-like ("GLP") representation of ARIMA(p, d, q)

As we discussed earlier,

$$Y_t \sim \text{ARIMA}(p, d, q) \implies Y_t \sim \text{a non-stationary ARMA}(p + d, q)$$

since

$$\Phi(B)\ \nabla^d Y_t = \Theta(B) e_t \implies \Phi(B)\,(1-B)^d\,Y_t = \Theta(B)\,e_t.$$

For a non-stationary time series, we cannot get a GLP representation (because GLP is stationary.) But we will try to derive a similar form of GLP for ARIMA(p, d, q) processes, to get a sense of the ACF behavior.

The corresponding non-stationary ARMA$(p+d,q)$ for ARIMA(p,d,q)

Example: Suppose $Y_t \sim \text{ARIMA}(p,1,q)$. Let $W_t = \nabla Y_t$, then $W_t \sim \text{ARMA}(p,q)$, so

$$W_t - \phi_1 W_{t-1} - \cdots - \phi_p W_{t-p} = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}.$$

$$(Y_t - Y_{t-1}) - \phi_1(Y_{t-1} - Y_{t-2}) - \phi_2(Y_{t-2} - Y_{t-3}) - \cdots - \phi_p(Y_{t-p} - Y_{t-p-1}) = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}.$$

$$Y_t - (1+\phi_1)Y_{t-1} - (\phi_2 - \phi_1)Y_{t-2} - \cdots - (\phi_p - \phi_{p-1})Y_{t-p} + \phi_p Y_{t-p-1} = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}.$$

As we already know, the last equation above is an ARMA$(p+1,q)$, but not stationary. Indeed, its AR polynomial is

$$\Phi^*(x) = 1 - (1+\phi_1)x - (\phi_2-\phi_1)x^2 - \cdots - (\phi_p - \phi_{p-1})x^p + \phi_p x^{p+1} = (1-x)\left(1 - \phi_1 x - \phi_2 x^2 - \cdots - \phi_p x^p\right)$$

which has a root $z = 1$.

"GLP" representation of ARIMA$(0,1,1)$

As we mentioned before, we cannot really derive a GLP representation for a non-stationary process. Some steps in the following analysis are not rigorous. Keep in mind that the big idea is to get a sense of the ACF behavior through an analogous way of GLP.

Suppose $Y_t \sim \text{ARIMA}(0,1,1)$, with $\nabla Y_t = e_t - \theta e_{t-1}$. Then $Y_t - Y_{t-1} = e_t - \theta\, e_{t-1}$. So

$$\begin{aligned} Y_t &= Y_{t-1} + e_t - \theta\, e_{t-1} = Y_{t-2} + e_{t-1} - \theta\, e_{t-2} + e_t - \theta\, e_{t-1} \\ &= Y_{t-2} + e_t + (1-\theta)\, e_{t-1} - \theta\, e_{t-2} \\ &= \cdots = Y_{t-m} + e_t + (1-\theta)\, e_{t-1} + \ldots + (1-\theta)\, e_{t-m+1} - \theta\, e_{t-m} \\ &\approx e_t + \sum_{j=1}^{\infty} (1-\theta) e_{t-j} \end{aligned}$$

where the last step is not rigorous, but can be thought as: $Y_{t-m} \to 0$ as $m \to \infty$ (assuming the process started at zero).

The last line above $e_t + \sum_{j=1}^{\infty}(1-\theta)e_{t-j}$ looks like a GLP, but it is not a GLP because $\sum_{j=1}^{\infty} |1-\theta|$ diverges (assuming $\theta \neq 1$), the condition $\sum |\psi_j| < \infty$ fails.

One can also show that :

$$\text{Var}(Y_t) = \left[1 + \theta^2 + (1-\theta)^2(t+m)\right] \sigma_e^2 \qquad \text{(which grows linearly in } t\text{)},$$

and the following results (not rigorous) for moderate k and large t

$$\rho_{t,t-k} = \frac{[1 - \theta + \theta^2 + (1-\theta)^2(t+m-k)]\sigma_e^2}{\sqrt{\text{Var}(Y_t)\,\text{Var}(Y_{t-k})}} \approx \frac{(1-\theta)^2(t+m)\,\sigma_e^2}{\sqrt{(1-\theta)^2(t+m)\cdot(1-\theta)^2(t+m-k)}\,\sigma_e^2} \approx 1.$$

So for an IMA$(1,1)$ process, the ACF $\rho_{t,t-k} \approx 1$ for moderate k and large t, which behaves similar to a random walk (for large t). If we plot the time series, it will exhibit wandering behavior (RW-like).

"GLP" representation of ARIMA$(1,1,0)$

Suppose $Y_t \sim \text{ARIMA}(1,1,0)$, with $Y_t - Y_{t-1} - \phi(Y_{t-1} - Y_{t-2}) = e_t$. So

$$Y_t = (1+\phi)Y_{t-1} - \phi Y_{t-2} + e_t$$

which is a non-stationary AR(2).

As before, we suppose there is a "GLP" representation $Y_t = \psi_0 e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \cdots$. Plug into the non-stationary AR(2) above:

$$(\psi_0 e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \cdots) = (1+\phi)\,(\psi_0 e_{t-1} + \psi_1 e_{t-2} + \psi_2 e_{t-3} + \cdots) - \phi\,(\psi_0 e_{t-2} + \psi_1 e_{t-3} + \psi_2 e_{t-4} + \cdots) + e_t$$

Comparing the coefficients of e_{t-k}:

$$\begin{cases} \psi_0 = 1 \\ \psi_1 = (1+\phi)\psi_0 \\ \psi_k = (1+\phi)\psi_{k-1} - \phi\psi_{k-2} \quad \text{for } k \geq 2 \end{cases}$$

which gives

$$\psi_k = 1 + \phi + \cdots + \phi^k = \frac{1-\phi^{k+1}}{1-\phi}, \quad \text{for any } k \geq 0.$$

Exercise: verify this result.

As we expected, this "GLP" is not a GLP, since the condition $\sum |\psi_j| < \infty$ fails.

Transformations of time series

Suppose a time series (Y_t) satisfies $\mathbb{E}Y_t = \mu_t$ and $\mathrm{Var}(Y_t) \approx \mu_t^2 \cdot \sigma^2$ (the latter implies $\mathrm{SD}(Y_t) \approx \mu_t \cdot \sigma$). Also assume (Y_t) is positive. A useful transformation for this type of time series is taking the logarithm:

$$\widetilde{Y}_t = \log Y_t.$$

This transformation has some nice property. First, using Taylor Series, we have the approximation $\log y \approx y_0 + \log'(y_0) \cdot (y - y_0)$. Replace y with Y_t, and let $y_0 = \mu_t$:

$$\log Y_t \approx \mu_t + \frac{1}{\mu_t}(Y_t - \mu_t).$$

Since μ_t is a non-random constant,

$$\mathrm{Var}(\log Y_t) \approx \mathrm{Var}\left[\frac{1}{\mu_t}(Y_t - \mu_t)\right] = \frac{1}{\mu_t^2}\,\mathrm{Var}(Y_t) = \frac{1}{\mu_t^2} \cdot \mu_t^2 \cdot \sigma^2 = \sigma^2 = \text{constant}.$$

So the variance of $\log Y_t$ is approximately the constant σ^2. For this reason, we call this transformation $\widetilde{Y}_t = \log Y_t$ (under the setting above) **Variance Stabilizing Transformation**.