Notes on ARIMA

14 декабря 2024 г. 19:01

Notes:

1\. Covariance showing us in which direction two(or more in common case) random values moving together. Or more simply how are they related. If cov is positive than they are positively related, if negative than negtively(with one rising, the other one falls)

2. Correlation = Cov/stderr1*stderr2 - shows us magnitude of the relation. If far from zero than strongly related, if close than not really related.
3. Autocorrelation is just a correlation but the value with itself in some lag in time
4. ACF accounts all effects of t-k to t. On the other hand, PACF only cares about direct effect of t-k to t. And that may occur useful as t-k may have been a very strong direct predictor of t.
5. If the PACF doesn't show any strong autocorrelation from lag k and higher than we should stick with AR(k)
6. PACF on lag k shows us coefficient i.e. impact of this lag in time on our current value.
7. So if PACF cuts off at lag 1 and other lags are no more than error, than we can consider coefficient approximately one(cause it's amount of correlation will be very close to one). This frequently occur for example in stock prices, especially of some big indexes like S&P500 or DowJones
8. Why differencing works? With each unit of x we move on some value of y. If we now remove this y transformation, we will see only remained noise which should wiggle around some mean value and not going upwards or downwards
9. the mu coefficient is basically the mean around which differenced function wiggles
10. In Moving average model you take in account errors, not values. Say you take errors with lag 2. Then formula will be y = u + phi1*errt-1 + phi2*errt-2 + errt
11. ACF is used to estimate order of MA and PACF is used to estimate order of AR.
12. The estimation of differencing parameter is actually pretty easy. We are trying d=0(assuming we already have stationary sequence) and looking at standard error. If nonstationary, than stderr should be very high. Also ACF should be quite decent and show a lot of correlation throughout all data points. Next trying d=1. ACF should now cut off much earlier and stderr should also drop down significantly. We continue this process until suddenly with new d our ACF changes sign and/or stderr rises. Then pick the previous d as it was an optimal choice