Formula Sheet - Final Exam

Jeff Nguyen

April 28, 2021

Properties of a Good ARIMA Model:

- Parsimonious: don't select large p and q (nothing more than order of 2).
- Has coefficients that are stationary and are invertible: recall stationary ($|\alpha| < 1$) and invertibility ($|\beta| < 1$).
- Fits data well (measures through AIC, SBC/BIC and select the best one).
- Has residuals that are white noise (No spikes in ACF and PACF. If residuals of ARIMA are not white noise, it means you have not filtered out information from your variable. The point of ARIMA is to filter out the information from variable Y).
- Has coefficients that do not change over sample period: If split sample to different time periods, run 2 ARIMA models and test coefficients are statistifcally the same or not.
- Has good out-of-sample forecast.

Stationary: Means $AR(p)$ to be stationary.
Invertibility: Means $MA(q)$ to be stationary.

Models of Volatility:

Many financial time series go through periods of tranquility and periods of volatility. Volatility and its measures are some of the most important concepts in finance.

Historical Volatility:

Measures the variance or standard deviation of returns over some period and uses it as measure of forecast of future volatility. It's used in option valuation and option pricing.

$$\sigma^2_{historical} = \frac{\sum(r - \overline{r})}{n-1} \tag{1}$$

Characteristics of Financial Data:

- Non-stationary (price). Most of the time, returns will be stationary.

- Leptokurtosis: sharp peaks, fat tails.
- Volatility clustering: Property of financial data to go through period of tranquility and volatility pooling/clustering.
- Leverage effect: Volatility rise more with price falls than price rises.
- Has ARIMA coefficients (p and q) that do not change over sample period (F-test of split periods).
- Has plenty noises as well as signals.
- Volatility is mean reverting.

Conditional Heteroscedacity: When forecasting, we are concerned more with conditional mean and conditional variance rather than historical mean and historical variance (long-run variance).
Heteroscedacity in Regression: happens when the variance of the errors depends on an indenpendent variable $Var[\epsilon_i] = cX_i^2$. For financial data, heteroscedacity happens when variance of errors depends on a dependent variable (volatility pooling).

Exponentially Weighted Moving Average Model (EWMA):

$$\begin{aligned}
\sigma_t^2 =& (1-\lambda)\sum \lambda_i (r_{t-i} - \overline{r})^2 \\
\text{where:}& \\
\sigma_t^2 =& \text{estimate variance for the period t, also the forecast of future volatility for all period} \\
\overline{r} =& \text{average returns} \\
\lambda =& \text{decay factor, usually .94}
\end{aligned} \tag{2}$$

Reading ACF and PACF

Non-stationary:

- If PACF has a single spike at $r_0 \approx 1$ and the rest is insignificant and ACF mostly significant and gradually decreasing $\Rightarrow$ non-stationary model.

$MA(\infty)$ and $AR(\infty)$

Several spikes in ACF (significant coefficients) $\Rightarrow MA(\infty) \Rightarrow AR(1)$.
Several spikes in PACF (significant coefficients) $\Rightarrow AR(\infty) \Rightarrow MA(1)$.

(A)DF Test

- Failed to reject the null hypothesis: there is a unit root $\Rightarrow$ non-stationary.
- Reject the null hypothesis: stationary.

Test for White Noise

Box-Pierce Test

Null hypothesis: Residuals are White Noise.

$$Q = T \sum r^2 \tag{3}$$

Box-Ljung Test

Null hypothesis: Residuals are White Noise.

$$Q = T(T+2) \sum (T-K)^{-r^2} \tag{4}$$

Recall AR(2)

$$\begin{aligned}
Y_t &= \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} \\
\text{Regressing } \epsilon: & \\
b_i &= \alpha_1 b_{i-1} + \alpha_2 b_{i-2} \\
\text{We have:} & \\
\epsilon_{Y_{t+1}} &= \sum_{i=0}^{1} b_i \epsilon_{t-1} \\
&= b_0 \epsilon_t + b_1 \epsilon_{t-1} \\
&= \epsilon_t + \alpha_1 \epsilon_{t-1} \\
\text{Thus:} & \\
Var[Y_{t+1}] &= (1 + \alpha_1^2)\sigma_\epsilon^2
\end{aligned} \tag{5}$$

Similarly:

$$\begin{aligned}
\epsilon_{Y_{t+2}} &= \sum_{i=0}^{2} b_i \epsilon_{t-1} \\
&= b_0 \epsilon_t + b_1 \epsilon_{t-1} + b_2 \epsilon_{t-2} \\
\text{Recall that: } b_i = \alpha_1 b_{i-1} + \alpha_2 b_{i-2} & \\
\text{We have:} \epsilon_{Y_{t+2}} &= \epsilon_t + \alpha_1 \epsilon_{t-1} + (\alpha_1 b_1 + \alpha_2 b_0)\epsilon_{t-2} \\
&= \epsilon_t + \alpha_1 \epsilon_{t-1} + (\alpha_1^2 + \alpha_2)\epsilon_{t-2} \\
\text{Thus:} & \\
Var[Y_{t+2}] &= \left[1 + \alpha_1^2 + (\alpha_1^2 + \alpha_2)^2\right]\sigma_\epsilon^2
\end{aligned} \tag{6}$$