Lecture 9A

Conditional Heteroscedastic Time Series Models - forecasting for $GARCH$ models

Recall that, from **Lecture 8**, that an $ARMA$ model in Finance has the representation given by

$$Y_t = \underbrace{\sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}}_{\text{Expected Return}} + \underbrace{\epsilon_t.}_{\text{Source of Volatility}}$$

Though the fundamental model for the asset return Y_t retains an $ARMA$ structure, where $Y_t = ExpectedReturn + Error$ (stated earlier above), because of the changing volatility (the conditional standard deviation of the error), the errors form a a white noise sequence with the structure given by

$$\epsilon_t = \sigma_t a_t,$$

where

$$\sigma_t^2 = \phi_0 \ + \ \sum_{i=1}^{m} \phi_i \epsilon_{t-i}^2 + \ \sum_{j=1}^{n} \gamma_j \sigma_{t-j}^2,$$

$$\phi_0 > 0,\ \phi_i \geq 0,\ \gamma_j \geq 0 \ \ \forall\, 1 \leq i \leq m, 1 \leq j \leq n,$$

and $\{a_t\}$ is a sequence of i.i.d. standard normal variables, a_t independent of $\mathcal{F}_{t-1}$ and the error ϵ_t.

The above model for the error sequence $\{\epsilon_t\}$ is called a $GARCH(m, n)$. model.

Also, recall that, we then have that $\{\epsilon_t^2\}$ as an $ARMA$ structure given by

$$\epsilon_t^2 = \underbrace{\phi_0 \ + \ \sum_{i=1}^{\max(m,n)} (\phi_i + \gamma_i)\epsilon_{t-i}^2}_{AR} \ + \ \underbrace{\eta_t - \sum_{j=1}^{n} \gamma_j \eta_{t-j}}_{MA}. \qquad (*)$$

The forecasting of the return on the stock (Y_t) is done by the usual procedure we have for a regular $ARMA$ model - with or without the changing volatility. The only difference between the forecasts of the return based purely on an $ARMA$ *only* and an $ARMA$ with $GARCH$ structure - that is, an $ARMA$ with constant volatility and an $ARMA$ with changing volatility - is that the confidence intervals of the return forecasts in the former case will involve the constant volatility σ, whereas in the latter case, the forecast of the return at a future time t will involve the corresponding forecast of the volatility σ_t.

Now, let $\{Y_t\}$ be the stationary $ARMA-GARCH$ model and let $\{Y_1, Y_2, ..., Y_T\}$ be the data set. Also, let $\{e_1, e_2, ..., e_T\}$ be the corresponding set of the residuals. In the existing procedure in the theory of $Time\ Series$, the future volatility (the future σ_t) is forecasted in the following way:

Procedure 1:

Here, the one-step-ahead forecast of σ^2_{T+1}, denated by $\hat{\sigma^2_T}(1)$ is given by

$$\hat{\sigma^2_T}(1) = \hat{\phi}_0 + \sum_{i=1}^{m} \hat{\phi}_i e^2_{(T+1)-i} + \sum_{j=1}^{n} \hat{\gamma}_j \sigma^2_{(T+1)-j}.$$

The two-step-ahead forecast is given by

$$\hat{\sigma^2_T}(2) = \hat{\phi}_0 + \sum_{i=2}^{m} \hat{\phi}_i e^2_{(T+2)-i} + \hat{\gamma}_1 \hat{\sigma^2_T}(1) + \sum_{j=2}^{n} \hat{\gamma}_j \sigma^2_{(T+2)-j},$$

and the k-step-ahead forecast by

$$\hat{\sigma^2_T}(k) = \hat{\phi}_0 + \sum_{i=k}^{m} \hat{\phi}_i e^2_{(T+k)-i} + \sum_{l=1}^{k-1} \hat{\gamma}_l \hat{\sigma^2_T}(l) + \sum_{j=k}^{n} \hat{\gamma}_j \sigma^2_{(T+k)-j}, \quad \text{if} \quad k < min(m, n).$$

The above gives the way to compute the future volatilities for each time point.

Note: The above forecasts, in no way, add anything new to the $ARMA$ forecasts of the future returns Y_t's.

Now, we will outline a simple, heuristic alternative method of forecasting the future volatilities, but not rigorously researched to be included in the literature of $Time\ Series$.

Procedure 2:

We know that for $GARCH(m, n)$ model, the earlier equation **(*)** gives an $ARMA$ structure to ϵ^2_t's. So, based on the data set of the squared residuals $\{e^2_1, e^2_2, ..., e^2_T\}$ as a proxy data set for the original $\{\epsilon^2_1, \epsilon^2_2, ..., \epsilon^2_T\}$, we can forecast ϵ^2_{T+k}'s for different

k's and use them appropriately in the forecasts of future σ^2_{T+k}'s. But, more importantly, we can forecast the *future* errors ϵ_{T+k}'s - except for their signs (negative or positive). Here is where we use the fact that $\epsilon_t = \sigma_t a_t$, and that $\{a_t\}$ is a sequence of i.i.d. standard normal variables and for each t, a_t is independent of $\mathcal{F}_{t-1}$ and the error ϵ_t. So, corresponding to each forecasted $\epsilon^2_{T+1}, \epsilon^2_{T+2}, ..., \epsilon^2_{T+k}$, we generate k independent standard normal values representing $a_{T+1}, a_{T+2}, ..., a_{T+k}$. Now, each of these generated k a_t's will be either positive or negative. We now assign the sign of a_{T+1} to the square root of the forecasted e^2_{T+1} and call it the *forecasted value* of ϵ_{T+1}, denoted by $e_T(1)$,and simply add it to the earlier $ARMA$ forecast of Y_{T+1} to get a new or modified forecast of Y_{T+1}. Then we apply this approach to get a new set of modified $ARMA$ forecasts of the returns Y_t's. That is, the new modified forecasts of Y_t's are defined by

$$\hat{Y_T}^M(1) = \hat{Y_T}^O(1) + e_T(1),$$

$$\hat{Y_T}^M(2) = \hat{Y_T}^O(2) + e_T(2),$$

and so on.

Assignment 2 - due by midnight of November 21, 2021 (Sunday)

Consider a financial data set (for example, stock price or index price etc.), say $\{P_t\}$, of sufficient size - at least a thousand observations.

1. Check for the stationarity of $\{P_t\}$ (generally the price process will *not* be stationary). If not stationary, check for the stationarity of the log return series $\{Y_t\}$ where $Y_t = \log(\frac{P_t}{P_{t-1}})$ - generally, $\{Y_t\}$ will be stationary. From this stationary data keep, at the most, the last 7 data points for testing and the rest for model development as required in question 2 below.

2. For the stationary $\{Y_t\}$ fit the appropriate $ARMA$ model and get its residuals $\{e_t\}$.

3. Perform the Ljung - Box test first on $\{e_t\}$, and then on $\{e^2_t\}$. If the null hypothesis is accepted for $\{e_t\}$ and is rejected for $\{e^2_t\}$, fit an appropriate $GARCH$ model and get the forecasts of Y_t's based on both Procedure 1 ($ARMA$ forecasts) and Procedure 2 (modified forecasts which involve addition of forecasted errors). Compute the MSE of your forecasts.

4. For all the above tasks, get the PDF files of the plots of P_t, Y_t, stationarity results (for the price and the log return series), $ARMA$ model results for the log returns, Ljung - Box test results (for both e_t and e^2_t sets), single graph of both the forecasts (original $ARMA$ and modified form) with actual test data values, and the MSE.