Time Series Modelling and Forecasting with Applications in R

Prof. Sudeep Bapat

Shailesh J. Mehta School of Management

Indian Institute of Technology Bombay

Week 10

Lecture 48: ARCH LM Test and GARCH Models

Hello all, welcome to this course on time series modeling and forecasting using R. Now, we are almost midway through this week, and again, as probably all of you know, the broad idea we are discussing this week is how to capture volatility and what exactly are the different models to capture the volatility, which is somewhat persistent in an underlying time series. And again, if you vaguely remember, just to briefly recap what we covered in the last lecture, we talked about one very important model called ARCH, or in short, ARCH, and the full form is autoregressive conditional heteroscedasticity, OK? Now, again, as all of you might know, the term heteroscedasticity means changing variance. So, the entire idea is, rather than focusing more on how to capture the mean aspect of the time series, we will try to capture the variance or the volatility in the underlying time series.

And again, as a brief reminder, in the last session, we talked about a generalized ARCH model, let us say with a certain order m, and further, we wrote down a particular ARCH model, let us say with an order of 1, and there we wrote down its model structure and a couple of properties about that, and so on and so forth. Now, this is again a new session this week, and the very first thing we will do this week is learn how to test for the ARCH effect. So, again, if you remember vaguely, we discussed even in the last lecture that before you try to model using any volatility models, we have to first ensure that the underlying residuals contain some ARCH effect. So, the ARCH effect means that the underlying residuals have changing variance properties. I mean, again, if the underlying residuals do not have changing variance properties, then it is of no use to focus on the variance aspect, right?

So, to ensure that the underlying residuals, which come from the mean equation, have some ARCH effect or rather exhibit the changing variance property, we have to formally test that, OK? And this is exactly what we will do now. So, the very first test we will discuss is called the LM test, and again, this is a very famous test. So, the full form is the Lagrange multiplier test, or in short, the LM test, or in short, you might see it referred to as the ARCH test in some textbooks. So, the whole idea is that you have to implement this LM test to confirm that the underlying residuals you get from the mean equation contain the ARCH effect or, in other words, exhibit some changing variance pattern.

Now, again, one small point to mention before we start with anything new here is how do you model the mean structure? So, we tend to model the mean equation by all the usual models we have studied so far. So, let us say ARMA, AR, MA, ARIMA, etc. Now, again, in the last session, we sort of walked through the rough steps as to how you sort of model the volatility of a time series. So, the very first thing is to sort of model the mean equation.

So, using an ARIMA model and so on and so forth, then you sort of extract the residuals from that mean modeling, and if the residuals show some changing variance behavior or changing variance aspect, then we say that you have to apply some ARCH or GARCH or its extensions on that. Okay. All right. So now, now again, how do you test that? So, the first test, again, just to summarize, is called the LM test.

So, the Lagrange multiplier, or in short, LM test, also called Engle's ARCH test. So, by the way, this was developed by Engle. It is called Engle's ARCH test. It is a commonly used test for detecting the ARCH effects. And exactly, we will study what this test involves and then what the following steps are.

So, the test involves the following steps. So, the first thing, as we talked about a short while back, is estimating the mean equation. And for that, you have to fit a model—let us say ARMA, ARIMA, or something like that—to the return series to get the residuals. So, the very first step is: how do you estimate the mean equation? So, let us assume that you estimate the mean equation using an ARMA process.

So, you model the ARMA and then extract the residuals. So, you model the ARMA for the return series and then get the residuals. And then what? So, the second step is you square the residuals. So, once you obtain the residuals from the mean equation, you square them.

So, square the residuals. So, square the residuals from the mean equation to get a proxy for the variance at each point in time. So, in short, this E_t square—when you take the square of the residual—this would be kind of equivalent, not entirely of course, to the variance of the underlying time series. And then again, we will show why exactly this is true. So, the moment you square the residuals, it serves as a proxy for the variance at each point in time.

So, this is something like. So, again, we are assuming that the variance of y_t is sort of changing along with time. So, you can write down something like sigma_t square, and this e_t square is kind of similar to the behavior of the variance term, which is sigma_t square. So, in short, the square of the residual proves to be a very important ingredient when you try to model the changing variance of a time series. Okay, then what?

So, once you extract the residuals, then you sort of fit a very simple regression model. And how exactly? So, you regress the squared residuals on a set of its lags to determine if the variance at a given time depends on past residuals or not. Again, just to repeat, once you extract the residuals and get hold of them, then you create a very simple regression model. And how exactly?

So, you regress the squared residuals on a set of lags to determine if the variance at any given time depends on some past residuals or not. And lastly, test for significance. So, the null hypothesis is that there are no ARCH effects. So, again, just to summarize, if you write down H_0, then I can simply write down something like ARCH effect is not there, whereas the alternative hypothesis would be something like the ARCH effect would be there, okay. So, once you sort of perform this test, you have to actually test it out or rather test for its significance.

So, the null hypothesis or H naught is that there are no ARCH effects, that is, no conditional heteroscedasticity exists. So, if the test statistic is significant, then ARCH effects are present. So, if you are able to reject the null hypothesis, which means that if the test statistic is significant, and you are able to reject the null, then we can sort of safely conclude that there exist some ARCH effects in the underlying residuals, ok. So, in general, this LM test is not really difficult.

So, what you do is you try to fit the mean equation using, let us say, an ARMA model or ARIMA model, and so on and so forth. And once you fit the mean equation to the return series, you extract the residuals, then you square the residuals because squaring the residuals serves as a proxy to the underlying variance at each point in time. And once you square the residuals, you create a very simple regression model where, of course, the dependent variable would be nothing but E t square, ok. So, this regression model that we are talking about here contains the dependent variable to be nothing but the squared residuals, ok. And what would be the independent variables?

Independent variables would be a set of lags to determine whether the variance at any given time depends on some past residuals or not. And lastly, you simply test for its significance. So, again, just to summarize all the steps. So, the very first step is model fitting. So, let us say, fit an ARMA model to the returns or residuals and obtain the residuals of that model, that particular ARMA model.

Then, the second step is to calculate the squared residuals, all right. Third, run a regression of the squared residuals on their lagged values, and this is, in particular, called the ARCH(1) model, OK. And lastly, you simply test the hypothesis that the coefficients of the lagged squared residuals are zero or not, OK. Alright, so now the next section is again important as to how one can pick the optimal ARCH order. OK, now again, let me just very briefly go back to the last session where we discussed a general kind of ARCH model.

Let us say a general ARCH model with order M, right. Now, again, the idea is how do you pick this order, right? I mean, how exactly do you pick the order optimally? So, would you say that the ARCH(1) model is the best, or the ARCH(2) model is the best, or the ARCH(3) model is the best, right? So, there has to be some yardstick to sort of pick this general order M. So, this is a small description here that selecting the appropriate order for an ARCH model is a very crucial step in time series modeling.

So, once you decide that, OK. So, there exists some ARCH effect in the residuals, and then you go ahead and decide that we have to put forward some ARCH model or other GARCH model on the residual data, but then the next question is how do you decide on its optimal order. So, hence, selecting the appropriate order for an ARCH model is a very crucial step in time series modeling. The order of an ARCH model, typically denoted as Q—so, let us say Q denotes the order of the ARCH model—represents the number of lagged squared residuals used to model the conditional variance, OK. So, let us say if you assume that the order of this ARCH model is Q, that letter Q determines the number of lagged squared residuals which are used in the model to model the conditional variance.

So, of course, these are the steps as to how you can pick the optimal ARCH order. So, the very first step is to perform an ARCH effect test, of course. So, before you actually go ahead and try to fix some Q, you have to first confirm the presence of the ARCH effect in the residuals of the time series model, and for this, one can again easily use the LM test.

So, confirm the presence of the ARCH effect in the residuals of the time series model, and for that, we can use the LM test, which was discussed just a short while back. Now, the next step is important: you have to always choose some initial order Q. So, let us say, choose some initial order Q and then use some information criteria, and then we have studied what exactly are some of the well-known information criteria, let us say AIC or BIC, etc.

So, use some of these information criteria, let us say AIC or BIC, and what you do is fit ARCH models of different orders and pick the best one corresponding to the least AIC. So, again, I will say that step 3 is kind of very easy. So, it is not difficult at all. So, what you do is you go ahead and try to fit different sorts of ARCH models containing different orders. So, let us say ARCH(2), ARCH(3), probably ARCH(4), ARCH(1), right.

And from this entire plethora of different ARCH models, you pick the one or rather fix the order that corresponds to the least information criteria. So, be it AIC or BIC. Okay. But lastly, you have to make sure that the fixed order or rather that particular ARCH model suits the data well. Right.

And how do you do that? So, examine the PACF of the squared residuals. Okay. So, examine the PACF or partial autocorrelation function. So, you can always have a graph of that.

Right. So, examine the PACF of the squared residuals. And when you examine the PACF of the squared residuals, peaks in the PACF indicate significant lags, which correspond to the ARCH order. Okay. So, again, I will say that step 4 is just to double-check that the order you fixed from step 3 makes sense or not, right? And then, how do you do that? Again, just to summarize, you create the PACF of the squared residuals, and wherever you see peaks in the PACF, those peaks indicate significant lags, which correspond to the ARCH order. Makes sense again.

Okay. So, now, we shift attention to a slightly advanced kind of model. So, we are shifting our attention from a very general sort of ARCH model to an extension of that, which is called the GARCH model. Right. And again, as discussed in the last section, probably, what exactly is the full form of GARCH?

So, the full form of GARCH is nothing but generalized autoregressive conditional heteroscedasticity, or rather generalized ARCH. Okay. And, of course, GARCH would be a suitable extension of the ARCH model. So, even in the last session, we saw that the ARCH model has some limitations as well. So, how do you counter those limitations?

So, of course, one can actually implement a suitable GARCH model or its extension. So, probably in the next session, we will see what all extensions are possible for a very general GARCH model as well. So, again, this is a very summary kind of slide. So, again, just to summarize the full form. So, GARCH stands for generalized autoregressive conditional heteroscedasticity.

So, now again, the idea is how do you model the changing waves, okay? And the GARCH model is nothing but an extension of the ARCH model, as discussed before as well, that adds an autoregressive component to the conditional variance, allowing for more flexibility and efficiency in modeling time series volatility, okay? So, what sort of is the extra portion in the GARCH model is that. Firstly, GARCH is nothing but an extension of the ARCH model that adds an autoregressive component to the conditional variance, allowing for more flexibility and efficiency in modeling time series volatility.

Secondly, the GARCH model captures both the past variances and the past squared residuals. So, it has two components instead. So, the GARCH model captures both the past variances as well as the past squared residuals. So, you have to account for or consider both the past variances. So, its own lags and the past squared residuals.

So, lags of E_t square as well. So, again, the GARCH model captures both the past variances as well as the past squared residuals, which is particularly useful in financial time series where volatility tends to cluster. Now, again, one very useful application of this entire ARCH or GARCH modeling can be seen in financial literature. So, how do you model, let us say, volatile time series? Again, volatile time series are very widely applicable or rather commonly seen in finance literature. So, let us say, just to give you a brief example, if you are observing the stock price and suddenly the stock price jumps due to some event, let us say it has a lot of fluctuations in between.

So again, we can safely assume that the overall variance is not constant. So, how do you model that? And hence, one can propose a suitable ARCH model or a GARCH model, etc. So now, we'll briefly study what the model structure is and how it exactly differs from an ARCH model. Now, first, you should understand that a GARCH model contains two orders.

So, P comma Q, OK. So, we will specify a GARCH model with orders P comma Q. The model specifies the conditional variance sigma t squared as. So, this is the equation for capturing the conditional variance. So, sigma t squared equals omega. Plus summation i going from 1 to q. So, the second order is q alpha i epsilon t minus i squared plus summation j going from 1 to p beta j and then sigma t minus j squared.

And then again, this is a brief summary of all the notations which are involved in the previous equation. Now, again, you may pause the video and then try to sort of understand as to how this GARCH model is behaving and how exactly it is different from an ARCHM model or rather an ARCHQ model, etcetera. So, again, the very important part here is that the GARCH model contains two different and disjoint sort of summations. So, the first summation contains these error squared terms or rather squared residuals. And the second summation contains the past lags or rather the past values of the conditional variance itself.

Make sense? So, epsilon t is nothing but the residual from the mean equation. So, whatever mean equation you fit, let us say using an ARIMA model or ARMA model, whatever. So, epsilon t denotes the residuals obtained or rather extracted from that mean equation. Sigma t squared, as always, is nothing but the conditional variance at time t. Omega serves as an intercept.

So, omega which is assumed to be positive is the constant term. Alpha i's are the coefficients for past squared residuals. So, by the way this is the arch effect or rather the arch terms right. So, again if you go back to the arch equation. So, arch equation does not contain the second term that we have here right.

So, this term is sort of extra here in the garch right. So, again alpha i's are nothing but coefficients for past squared residuals which correspond to the arch terms whereas beta j's are nothing but the additional coefficients for past variances which correspond to the garch terms. So, this entire first set or the first summation correspond to the arch equation and this entire second set correspond to the garch equation. And again lastly P and Q what are they? So, P and Q are nothing, but the order of the garch and arch terms respectively ok.

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

So, P corresponds to the garch order and then Q corresponds to the arch order make sense so far hopefully ok. So, again just to summarize very quickly. So, sigma t square is dependent on some constant term plus some arch terms and then again arch terms are what? So, arch terms correspond to nothing, but the lags of the squared residual.

So, epsilon t minus i squared and multiplied by some coefficients, let us say alpha i, plus you have several GARCH terms, and then GARCH terms contain the past lags of the conditional variance itself. So, sigma t minus g squared multiplied by coefficients which are denoted by beta g. So, now we will study very briefly what some of the important key features of this GARCH model are. So, the very first feature is called volatility clustering, and again, I am very sure that we have covered this first point several times in the last couple of lectures as well. So, what do you mean by volatility clustering? It models periods of high and low volatility commonly observed in financial data, but again, in short, volatility clustering means that periods of high volatility

are followed by the same sort of period, and periods of low volatility are followed by low volatility, okay. So, this phenomenon is called volatility clustering. Secondly, mean reversion. So, what do you mean by mean reversion? So, mean reversion is implied when summation alpha i plus summation beta j is strictly less than 1, meaning volatility eventually reverts to a long-term level, okay.

So, under this condition that summation alpha i plus summation beta j, if you simply add all the coefficients right, is strictly less than 1, this means that eventually the volatility or changing variance would revert to a long-term level, okay. I will give you a very short example in this regard: let us say that even though you will see drastically changing variance initially, let us say something like that, but later on, the variance has to converge to some long-term sort of a constant, something like that. And then this tendency in the volatility is called mean reversion. And lastly, leverage effect. So, what do you mean by leverage effect?

So, standard GARCH models cannot handle asymmetries like negative shocks affecting volatility more than positive shocks. So, let us say for some reason if negative news or negative shocks affect the changing variance or volatility more than positive news or positive shocks. Then, a simple ARCH model or a simple GARCH model cannot handle such asymmetries. For that, you require all sorts of extensions, which we will study in the next lecture. But these are a couple of extensions.

So, let us say EGARCH or GJR-GARCH, right? So, extensions like these address this phenomenon, which is called the leverage effect, okay? So, these are some shortcomings when it comes to GARCH modeling or some of the key features of GARCH modeling. Now, this slide I have tried to put just to summarize the entire idea and to briefly explain the steps to fit a GARCH model. So, how exactly can one fit a suitable GARCH model to data?

So, again, the very first step is to always perform the ARCH effect test to confirm the presence of conditional changing variance or heteroscedasticity. So, the very first step is to always perform the ARCH effect test. One can use the LM test for that on the residuals. Then, the second step is to use model selection criteria like AIC or BIC to decide on P and Q. Thirdly, one can use MLE. So, let us say maximum likelihood estimation to fit the GARCH models.

Then check the diagnostics. So, what do you mean by diagnostics? So, for example, residual should have no arch effects that is residual should be completely homoscedastic right. I mean once you fit the GARCH model then you obtain the residuals of that GARCH model and then residual should not have any more arch effects obviously ok. And lastly use the fitted GARCH model to forecast the future variances and any conditional volatilities.

So, this is sort of a walk through as to how can one sort of fit a suitable GARCH model where you have to first confirm that whether ARCH effects are there or not. Then use some model selection criteria like AIC or BIC to decide on the orders P and Q. Then you have to estimate all the parameters. So, one can make use of MLE estimation here and check the diagnostics. So, check diagnostics is always there rather which is nothing but diagnostic checking we always do right after the model fitting. And lastly, forecasting.

So, use the fitted GARCH model to forecast some future variances and conditional alternatives. So, in short, if you want to specify what exactly or how exactly does a GARCH 1-1 model look like. So, I have just presented this slide here just again as a summary. So, the GARCH 1-1 model is one of the most commonly used models for volatility modeling in time series data. And what it does is that it extends the basic ARCH model by incorporating lagged conditional variances in addition to lagged square decadal.

So, again, the entire idea is the same. The only thing is, P happens to be 1 here, and Q happens to be 1 here instead. So, this is nothing but the mean equation. So, y_t equals mu plus epsilon_t, and epsilon_t also appears in the second equation, where you are trying to model the variance. So, sigma_t squared happens to be equal to omega plus alpha times epsilon_t minus 1 squared plus beta times sigma_t minus 1 squared, okay.

So, this is a very simplistic-looking sort of GARCH(1,1) model, which is again a very common model used to model real-life data. Now, firstly, what exactly are the key properties of the GARCH(1,1) model? So, again, I will say that these key properties are exactly similar to some of the key features we studied for a general GARCH(P,Q) model, right? The first is again the same. So, volatility clustering.

So, it models periods of high and low volatility, commonly observed in financial data. The second one is again the same: mean reversion. So, when does mean reversion happen? Whenever this condition is met. So, if the sum of alpha_i plus the sum of beta_j is less than 1, then the overall volatility eventually reverts to a long-term level.

And thirdly, stationarity. So, again, on top of that, if this condition is also met—that if alpha plus beta is less than 1. For example, in a GARCH(1,1) model, you only have two coefficients, right? So, you do not have summations here, right? So, again, if you go back a slide.

So, you do not see any summations here, and why is that? Because both the orders are 1, is it not? So, you do not have any summations here. So, you only have these two coefficients to focus on. So, alpha and beta. So, if it happens that alpha plus beta is less than 1, the model is stationary.

On the other hand, if alpha plus beta is really close to 1 or almost equal to 1, then the volatility exhibits some long memory or persistence. And again, I am very sure that by this time, you should know what you mean by long memory. So, we had an entire week where we discussed what you mean by long memory, persistence in the underlying time series, and all. So, these small conditions pertaining to a GARCH(1,1) model could either mean mean reversion or stationarity. And on the other hand, if alpha plus beta is very close to 1, then the changing variance exhibits some long memory or persistent properties.

Alright. So now, towards the end, we will sort of wrap up this week by discussing a few practical applications. And again, the practical applications could be seen in multiple different areas. The first one is, let us say, financial markets in, let us say, asset price modeling. So, what is the objective here?

So, the objective is to model the volatility of returns for assets like stocks, bonds, currencies, etc. A small example in this regard is predicting volatility in stock prices or cryptocurrency returns for, let us say, trading strategies. And what is the impact of that? So, it helps traders optimize some risk-return trade-offs in high-frequency trading. So, the first example could be seen in financial markets pertaining to, let us say, asset price modeling.

Again, a second example in financial markets could be in relation to option pricing. So, how exactly? So, the objective here is nothing but to compute the implied volatility used in pricing options. For example, GARCH models are integrated into the stochastic volatility models for dynamic pricing, let us say, in Black-Scholes-type frameworks. And what exactly is the impact?

So, provides a better understanding of changing variance or the skew observed in the market. Now, this is a slightly different sort of an example coming from a risk management point of view. So, value at risk. So, value at risk or in short VAR is a very very common kind of a metric to capture the extremes of the risk measures. So, the objective here is to estimate the risk of portfolio loss over a certain time horizon.

For example, banks use GARs to compute the value at risk considering time varying volatility. And what is the impact? So critical for compliance with regulatory requirements like Basel 3 records and so on and so forth. So, by the way, Basel 3 records come under the compliance framework or rather the regulatory framework. And lastly, stress testing.

So again, coming from coming under the risk management heading. So stress testing, the objective is to simulate worst case scenarios by considering periods of high volatility. One small example is analyzing portfolio behavior during financial crisis, etc., So, this sort of sums up the idea behind let us say the extension of ARCH which is called as GARCH model. Now, in the next session, so next session would be the last theoretical session this week, we will try to extend this a step ahead and then try to incorporate lot of extensions of a general GARCH model.

And then the idea behind that is that each and every extension we study in the next lecture would be suitable to model and put forward some practical applications in a different manner.

Thank you.