Stat 556 Final Project:
Volatility and Correlation Forecast with Intra-day Data

Cunwei Fan

Department of Physics, University of Illinois at Urbana-Champaign

The volatility and correlations for stocks are always of great interest to people in the finance and statistical fields. With the accessible of intra-day data, it is important to fully utilize the given information. In this project, we design a simple model to utilize the intraday data to model and forecast the volatility of several symbols and correlations among them simultaneously. Specifically, the model tries to describe the correlation among shocks and the individual stock volatility independently. We do this by employing the realized volatility into the dynamic conditional correlation (DCC) model [1] and call it the rGARCH-DCC model. The model achieves better volatility estimation on training data than simple GARCH and the correlations among the stocks are successfully described.

Contents

1 Introduction

Stock volatility and correlations are of great interests to practitioners in the financial market fields. With accurate modeling of correlations and volatility, successful hedge strategies and profitable portfolios can be designed accordingly. Based on daily data, various models have been designed to model volatility, such as GARCH and the various modifications to GARCH. However, daily returns estimation for volatility could forbid the model to be accurately estimated [2, 3]. With the accessible intraday data, people could have better estimation of volatility and the classical model, such as GARCH, performs better with the help of the more accurate estimation of volatility given by the high frequency data [3]. Another way of utilizing the realized measurement by high frequency data is to add those quantities explicitly into the model like the efforts shown in [4, 5, 6, 7]. People also tried to model the high frequency data explicitly and extract the volatility estimation or prediction from those high frequency data. The introduction and references within [8] explained this type of efforts explicitly. In this project, we focus on the first method that we try to add the realized measurements explicitly into the volatility modeling. The advantage of this type of models is that long horizon prediction can be made even though the future realized measurement is not available.

Besides volatility, the correlation among symbols in a given portfolio will also play an important role in portfolio volatility prediction. The dynamic conditional correlation (DCC) model [1] is a successful trial in modeling the correlation among shocks for correlated symbols. A good idea of combing the correlation modeling and realized measurement is given by [9]. However the model is complicated and a satisfactory estimation of the model parameters require large amount of data. In this project, we propose a simpler way. We model the correlation among shocks by the conventional DCC model but the scale (the volatility) of returns will be modeled separately. In this way, we could add the realized measurement to the individual volatility modeling while keeping the correlation among shocks. Also, the individual stock volatility estimation and forecasting will not be affected by the correlation modeling and thus the only affected estimation and prediction will be the portfolio volatility which needs the information of covariance among returns.

In this project, we investigate our model on a realistic stock dataset with AAPL, AMZN, FB, MSFT, TSLA and SP500 stock high frequency data in 2019. Our model fits well in AAPL, MSFT, TSLA and SP500 but does not converge well on FB and AMZN. However the correlation among the shocks is successfully described where our forecast of the volatility of six random portfolios consisting of the six symbols behave well. In conclusion, although our model doest not converge well on two specific symbols, it successfully incorporated the realized variance into the Garch model and successfully model the correlation among shocks of the symbols.

2 Method

Given intraday data for several stock symbols, an appropriate model and a good estimate of realized measurement should be chosen. In this section, we will introduce the concept of realized measurement and the realized correlated volatility model we proposed in this project.

2.1 Realized Volatility

The standard stochastic model for the stock price is the generalized exponential random walk model. As a result, the log return between time ticks follows the generalized random walk model:

$$d\log P_t = (\mu - \frac{\sigma^2}{2})dt + \sigma dW_t. \tag{2.1}$$

In this project, we use the subscript t to denote integer days of trading and fraction of it to denote intraday time ticks. For example, if there are s ticks in one trading day, the intraday data points will be labeled as $P_{t+i/s}$ for $i = 0, 1, \cdots, s-1$. The task of the project is to estimate daily return volatility, and as a result, we could estimate it using the intraday data as follows

$$\int_0^1 \sigma^2(t)dt \sim \sum_{i=0}^{s-1} \sigma^2(t + i/s)\Delta t \sim \sum_{i=0}^{s-1} (\log P_{t+(i+1)/s} - \log P_{t+i/s})^2. \tag{2.2}$$

In the equation, the left hand side is the volatility of the daily log return if we assume the model (2.1). The middle part is the discrete summation approximation for the integral and the right hand side the data estimation for the summation. We define the right hand side summation as the realized variance or realized volatility

$$\mathrm{RV}_{t+1}^{(s)} = \sum_{i=1}^{s} r_{t+i/s}^2 \equiv \sum_{i=1}^{s} (\log P_{t+i/s} - \log P_{t+(i-1)/s})^2. \tag{2.3}$$

Ideally, the realized volatility will be an error free estimation of the daily volatility if $s \to \infty$ because the summation approaches the integral in (2.1). However, as the s becomes larger, the interval between ticks become smaller and the non-ideal liquidity will introduce the microstructral noise into the log returns $r_{t+i/s}^2$. As a result, a consistent way to choose the appropriate s is needed to deal with the "bias variance" trade off. We follow the recipe introduced in [10] and the exact procedure on real data will be shown in detail in the experiment section.

2.2 The model

In order to estimate risk, design portfolio and trading strategy, a good forecast of both volatility and correlation among the symbols of interest is needed. Univariate model describes volatility successfully and the correlation among the shocks could be estimated by the standardized shocks. This estimation assumes constant correlation among shocks. However, the correlation among the symbols may also evolve with time and a constant estimation of correlation matrix among the symbols will not give satisfactory results, especially when the time span is long and the correlation matrix oscillates a lot. The way we handle the correlation is borrowed from DCC model [1]. The way we incorporate the realized measurement is that we model the individual realized volatility instead of the realized covariance among the symbols. Moreover, We model the realized volatility in the view that it is an observation of the hidden true volatility with random noise.

Specifically, we let r_t^i to denote the log return $\log P_t^i - \log P_{t-1}^i$ for stock i, σ_t^i to be the conditional variance of stock i, z_t^i to be the shock for stock i and C_t be the conditional correlation matrix among z_t^i. The model can be described as

$$r_t^i = \mu^i + \sigma_t^i z_t^i \tag{2.4}$$

$$\log\big(\mathrm{RV}_t^i\big) = \xi^i + \phi^i \log\big(\sigma_t^i\big)^2 + \tau_1^i z_t^i + \tau_2^i((z_t^i)^2 - 1) + u_t^i \tag{2.5}$$

$$(\sigma_{t+1}^i)^2 = \omega^i + \beta^i(\sigma_t^i)^2 + \gamma^i \mathrm{RV}_t^i \tag{2.6}$$

$$C_t = \mathrm{diag}(Q_t)^{-1/2} Q_t \mathrm{diag}(Q_t)^{-1/2} \tag{2.7}$$

$$Q_{t+1} = S(1-a-b) + a(z_t z_t^T) + bQ_t \tag{2.8}$$

$$z_t \sim \mathcal{N}(0, C_t) \tag{2.9}$$

$$u_t \sim \mathcal{N}(0, \Sigma_u). \tag{2.10}$$

The first equation describes the return using the volatility scale and the constant expected mean μ^i. The second equation describes the "measurement" mechanics: the realized variance depends on the hidden true volatility and the shocks as well as a measurement noise u_t. These u_t have zero pairwise correlation but have different scales for different stock i and thus Σ_u is a diagonal matrix. The log transformation on the measurement equation is to make the noise u_t similar to Normal distribution. The third equation is the updating equation for the volatility σ_t^i. Notice, the usual GARCH model uses return squared to update the volatility but here, we use the more accurate estimate, the realized variance. The Q_t is the sample estimate of the conditional covariance matrix of z_t and the update equation (2.8) is mean revert to the sample estimation S. The (2.7) ensures that the covariance matrix for the shocks is normalized.

2.3 Parameter Estimation

The model could be estimated using maximum log-likelihood. Note the measurement equation does not couple with the dynamic part and thus we should focus on the dynmaic part first and after get good estimate of σ_t, we could proceed to get the parameters in the measurement equation. The log likelihood of the returns r_i can be calculated from the probability density as

$$p(r_1, \ldots, r_T|\mathrm{RV}, \theta) = p(r_T|\mathcal{F}_{T-1}, \theta)p(r_{T-1}|\mathcal{F}_{T-2}, \theta)\cdots p(r_1|\mathcal{F}_0, \theta)p(\sigma_0, RV_0). \tag{2.11}$$

We first focus on the conditional likelihood estimation, where we neglect the last factor for the initialization. Given $\mathcal{F}_{t-1}$, r_t follows the multivariate distribution. If we let the D_t be the diagonal matrix with diagonal terms equal to σ_t^i, then the log likelihood of $r_t|\mathcal{F}_{t-1}$ becomes

$$\begin{aligned}\log p(r_t|\mathcal{F}_{t-1}, \theta) &= -\frac{1}{2}\left\{\log|D_tC_tD_t| + (r_t-\mu)^\dagger D_t^{-1}C_t^{-1}D_t^{-1}(r_t-\mu)\right\} \\ &= -\frac{1}{2}\left\{\Big(\log|D_t|^2 + (r_t-\mu)^\dagger D_t^{-2}(r_t-\mu)\Big) + \Big(\log C_t + z_t^\dagger C_t^{-1}z_t - z_t^\dagger z_t\Big)\right\}\end{aligned} \tag{2.12}$$

where we used the fact that $D_t(r_t - \mu) = z_t$. We also dropped the constant term $-\frac{1}{2}n\log(2\pi)$ in the log likelihood. We should notice that the log likelihood only assumes a multivariate normal among r_t with covariance matrix $D_tC_tD_t$ and mean μ. This makes it possible to compare the log likelihood directly with that of other Guassian family models to find a better model because they assume the same conditional distribution for r_t.

The structure of the log likelihood shows that

$$L = -\frac{1}{2}\sum_{t=1}^{T}\left\{\sum_{i=1}^{n} 2\log\sigma_t^i + \frac{(r_t^i-\mu^i)^2}{(\sigma_t^i)^2}\right\} - \frac{1}{2}\sum_{t=1}^{T}\left\{\log C_t + z_t^\dagger C_t^{-1}z_t - z_t^\dagger z_t\right\}. \tag{2.13}$$

The first part is the usual GARCH model part and the second part is the dynamical correlation part. The strategy of the optimization would be that we first optimize the first part and get the standardized shocks z_t and optimize the second part. It should be noticed that the covariance matrix $H = D_tC_tD_t$ has diagonal terms as

$$H_{ii} = \sum_{j,k}(D_t)_{ij}(C_t)_{jk}(D_t)_{ki} = (D_t)_{ii}(C_t)_{ii}(D_t)_{ii} = (D_t)_{ii}^2. \tag{2.14}$$

The optimization procedure can be viewed as we first estimate the diagonal terms in the covariance matrix H_t and then estimate the non-diagonal terms in H_t.

Although, the first procedure is almost the usual GARCH procedure, the added realized volatility RV_t cannot be put into the standard library of GARCH. As a result, we would use the iterative algorithm to solve the problem. It is beneficial to have a closed form gradient. Given the parameters μ^i, ω, β, γ, σ_0^i and the $(\sigma_t^i)^2$ corresponding to the parameters, we could get the gradient for the parameters as

$$\begin{aligned}\frac{\partial L}{\partial\omega^i} &= -\frac{1}{2}\sum_{t=1}^{T}\left(\frac{1}{(\sigma_t^i)^2} - \frac{(r_t^i-\mu^i)^2}{(\sigma_t^i)^4}\right)\frac{\partial(\sigma_t^i)^2}{\partial\omega^i} \\ \frac{\partial L}{\partial\beta^i} &= -\frac{1}{2}\sum_{t=1}^{T}\left(\frac{1}{(\sigma_t^i)^2} - \frac{(r_t^i-\mu^i)^2}{(\sigma_t^i)^4}\right)\frac{\partial(\sigma_t^i)^2}{\partial\beta^i} \\ \frac{\partial L}{\partial\gamma^i} &= -\frac{1}{2}\sum_{t=1}^{T}\left(\frac{1}{(\sigma_t^i)^2} - \frac{(r_t^i-\mu^i)^2}{(\sigma_t^i)^4}\right)\frac{\partial(\sigma_t^i)^2}{\partial\gamma^i}\end{aligned} \tag{2.15}$$

The derivative of σ_t^2 on $,\omega,\beta,\gamma$ can be calculated recursively from the $t=0$. For example, the derivative with respect to β is:

$$\frac{\partial(\sigma_t^i)^2}{\partial\beta^i} = (\sigma_{t-1}^i)^2 + \beta^i \frac{\partial(\sigma_{t-1}^i)^2}{\partial\beta^i} \quad \text{with } \frac{\partial(\sigma_1^i)^2}{\partial\beta^i} = (\sigma_0^i)^2. \tag{2.16}$$

The update of μ^i could be carried out by the exact solution of $\partial_\mu L = 0$ which gives

$$\mu^i = \left(\sum_{t=1}^{T} \frac{1}{(\sigma_t^i)^2}\right)^{-1} \sum_{t=1}^{T} \frac{r_t^i}{(\sigma_t^i)^2}. \tag{2.17}$$

Thus, for given parameters, we would first update μ^i and calculate the derivative for ω,β,γ and then after the gradient descent update, we update μ^i again and do the iteration until convergence.

The correlation part has similar algorithm for optimization. For simplicity, since (2.8) shows the mean reverting property, we could fix S as the sample covariance matrix and do not optimize over it. The exact derivative for the parameters a and b can be found in the appendix.

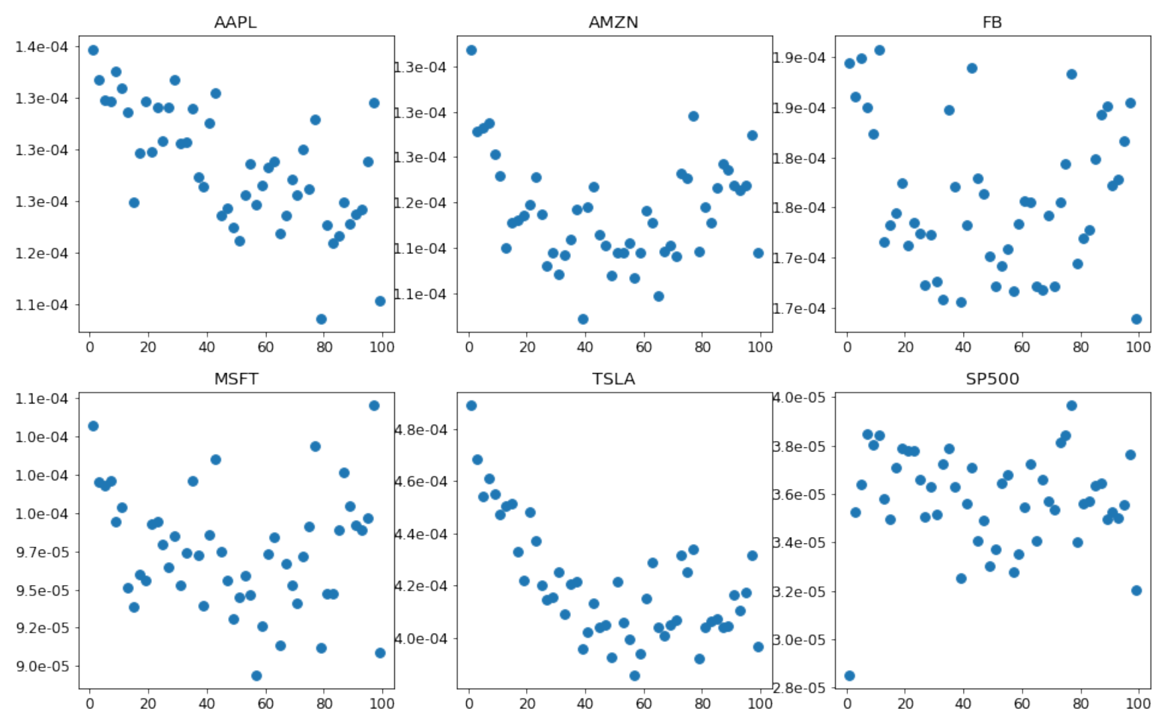


Figure 1: The average of daily realized volatility from 2019-01-02 to 2019-10-31 is calculated for various tick spacing. The figure plots the realized volatility versus the tick spacing s with s ranges from 1 to 100.

3 Experiment

The data set we want to test the model on is the stock prices for APPL, AMZN, FB, MSFT, TSLA and the index of SP500. Since the stocks are all in technology fields and most of them plays an role in SP500, we would expect non-trivial correlations are among the symbols. The data spans from January 2nd, 2019 to December 31st, 2019 and the data are all intraday data with time ticks as 1 minute [11].

The data contains open, close, highest and lowest price at each one minute interval. Also, the stock data contains information for after hours trading. We split the data into training and validation data. The training data spans from January 2nd to November 29th and the remaining is the validation data.

3.1 Construct Realized Volatility

The realized volatility should be constructed first. For simplicity, we would only use close price to calculate realized volatility. As discussed in section 2.1, the time intervals to calculate the volatility should be chosen carefully to approximate the hidden truth and also eliminate microstructral noises. We follow the following recipe: we first select a spacing s minutes from $s = 1$ to $s = 100$. Then we calculate the realized variance RV^s for various days and take the average. Finally, we plot the average realized variance with respect to s. We would find the smallest s such that the RV^s stabilizes after s.

In this data, for each spacing, we calculated the realized variance for all trading dates from 2019-01-02 to 2019-10-31 in the training set. In the Figure. 1, the average realized variance is plotted versus the spacing s. The selected spacing for each symbol is listed in the Table. 1. With fixed spacing s, in order to eliminate the noise, we choose different starting point to calculate the realized variance. Then we take average of the calculated realized variance from different starting point for each trading day. Specifically, the final realized variance for the chosen s is calculated as

$$\mathrm{RV}_{t+1} = \frac{1}{s}\sum_{j=0}^{s-1}\sum_{i=1}^{N/s}\left(\log P_{t+(is+j)\delta t} - \log P_{t+((i-1)s+j)\delta t}\right)^2 \tag{3.1}$$

where j is the various starting point and N is the number of ticks for each day. δt is the time difference between each tick it is one minute in our case. It should be understood that the divide N/s takes the integer part of the dividing result.

symbol	AAPL	AMZN	FB	MSFT	TSLA	SP500
s	40	30	10	10	40	10

Table 1: The spacing s chosen for the realized variance for each symbol

3.2 Model Fits and Results

The model is optimized using "SLSQP" algorithm which is an approximate second order iterative optimization method. It considers the bounds and constraints for the parameters. Since an initial guess is needed for the iterative method, we try multiple starting points for the optimization and choose the one with the maximum log likelihood. In order to estimate the variance of the parameter estimation, we use the idea of bootstrapping. Since we only have one time series observation, the bootstrapping we do is using sliding window. Specifically, we move a window of width L with L less than the length of the sequence T. For each window, we fit a model and collect the estimated parameters and calculate the variance from this data collection. This assumes that the model parameters are pretty stable within a certain period.

The fitted parameters are given in the table. It should be noted that for TSLA, the coefficient β is zero. This means that the volatility depends highly on the last realized volatility instead of the hidden true volatility which indicates that the response of the stock price is sensitive to shocks. Another explanation may be that the realized volatility can the hidden volatility σ_{t-1}^2 have strong correlation and the bound $\gamma + \beta < 1$ makes the optimization more like the Lasso regression and the coefficient γ is set automatically to zero by the procedure. The coefficient for FB shows that the γ is just zero and thus the estimation for the volatility of FB will be an exponential decay from initial σ_0 to the long term volatility estimated. The oscillation of realized volatility does not affect the σ_t^2 estimation. This indicates that the conditional volatility of FB does not depend on the behaviour of the market. This may be due to the some numerical instability or data is not fully cleaned. The model fits are thus not successful for FB. The mean estimation for log returns shows the log returns have significant mean for all symbols except for AMZN. This conincides the fact the stocks increase a lot for all symbols except for AMZN.

For the measurement equation, the interesting feature we see is that the τ_1 coefficient is zero for all stock. Also, the error of the estimation is small compared to the value. This indicates the fact that when the return is decreasing, the volatility tends to be large. Besides, τ_2 is statistically significant with positive value. This shows that the realized volatility tends to increase with large oscillation in the market. In addition, the power coefficient ϕ is around 1 for all stocks except FB. This shows that our assumption about that realized variance is a measurement of underlying volatility. Also, ξ seems not statistically significant since the estimated standard error are larger than the estimation. This indicates that the observation of realized variance will be a non-biased estimator of the underlying volatility. This coincides with our measurement assumption.

Realized Garch Parameters

	AAPL	AMZN	FB	MSFT	TSLA	SP500
$\omega(10^{-5})$	3.52 (0.20)	1.31 (1.50)	1.18 (5.60)	1.10 (1.60)	24.1 (5.20)	0.30 (0.07)
β	0.31 (0.03)	0.80 (0.18)	0.93 (0.35)	0.67 (0.21)	0.00 (0.00)	0.42 (0.08)
γ	0.50 (0.05)	0.11 (0.10)	0.00 (0.04)	0.25 (0.13)	0.76 (0.13)	0.50 (0.11)
$\mu(10^{-4})$	10.3 (0.75)	0.49 (1.95)	2.46 (1.49)	1.75 (0.94)	4.50 (1.76)	3.87 (0.61)

Measurement Parameters

	AAPL	AMZN	FB	MSFT	TSLA	SP500
ξ	0.22 (0.49)	7.07 (3.25)	153 (518)	2.08 (3.53)	-1.13 (1.08)	-0.79 (0.98)
ϕ	1.08 (0.05)	1.85 (0.36)	18.8 (60.2)	1.26 (0.38)	0.93 (0.14)	0.94 (0.10)
τ_1	-0.10 (0.01)	-0.09 (0.03)	-0.11 (0.01)	-0.07 (0.01)	-0.03 (0.01)	-0.17 (0.02)
τ_2	0.24 (0.00)	0.18 (0.01)	0.13 (0.00)	0.16 (0.01)	0.15 (0.00)	0.15 (0.01)
σ_u	0.65 (0.01)	0.58 (0.01)	0.49 (0.02)	0.44 (0.01)	0.57 (0.00)	0.55 (0.01)

DCC Parameters

	Estimate	Err
a	0.3095	0.0053
b	0.5918	0.0156

Table 2: Parameter Estimation for the rGARCH-DCC model. The boots-trapping estimated standard deviation for each parameter is listed in the parenthesis.

3.3 Prediction

Although the fitted model does not do well for FB, we still could do the prediction for realized variance or volatility on the last month from December 1st to December 31st in the test dataset. This could validate our model for the individual prediction. For the correlation model, we could construct various portfolio with different portions of the stocks on the test. Then, we could see how the prediction of volatility of the portfolio behaves on the test data.

In order to construct prediction confidence interval, we run the prediction multiple times and take the $0.025, 0.50, 0.975$ quantiles on the predicted realized variance and volatility. We plot the fitted σ_t^2 and prediction in Figure 2. In the plot, the orange line is the fitted volatility σ_t^2 and the blue line is the underlying true realized variance. The gray shaded region is the 95% confidence interval for the predicted RVs. In the plot, it is clear that except FB, the stocks realized volatility after December 1st is within the confidence interval. This indicates a good sign of prediction. Besides, For AAPL, MSFT and SP500, the conditional volatility follows the realized variance closely which indicates the model is sensitive to shocks. On the other hand, AMZN and TSLA has rigidity that is not observed in other symbols. This is shown in the plot that the conditional variance does not drop to low value when the realized variance does.

We can also validate our correlation estimation calculating some random portfolio volatilities. Since the log returns are small in magnitude, the portforlio log return can be simply the weighted average of the components portfolio. In the figure, we plot six random portfolios. In the plot, the orange line is the fitted σ^2 for the portfolio and the shaded area is the 97.5% and 2.5% quantiles of the predicted σ^2 of the portfolio. The blue line is the portfolio log return squared r^2. It is shown in the figure, the fitted conditional σ^2 follows trend of the return squared. Also, the log return squared in December lies in the shaded region. This indicates a good estimation of the correlation matrix among the symbols.

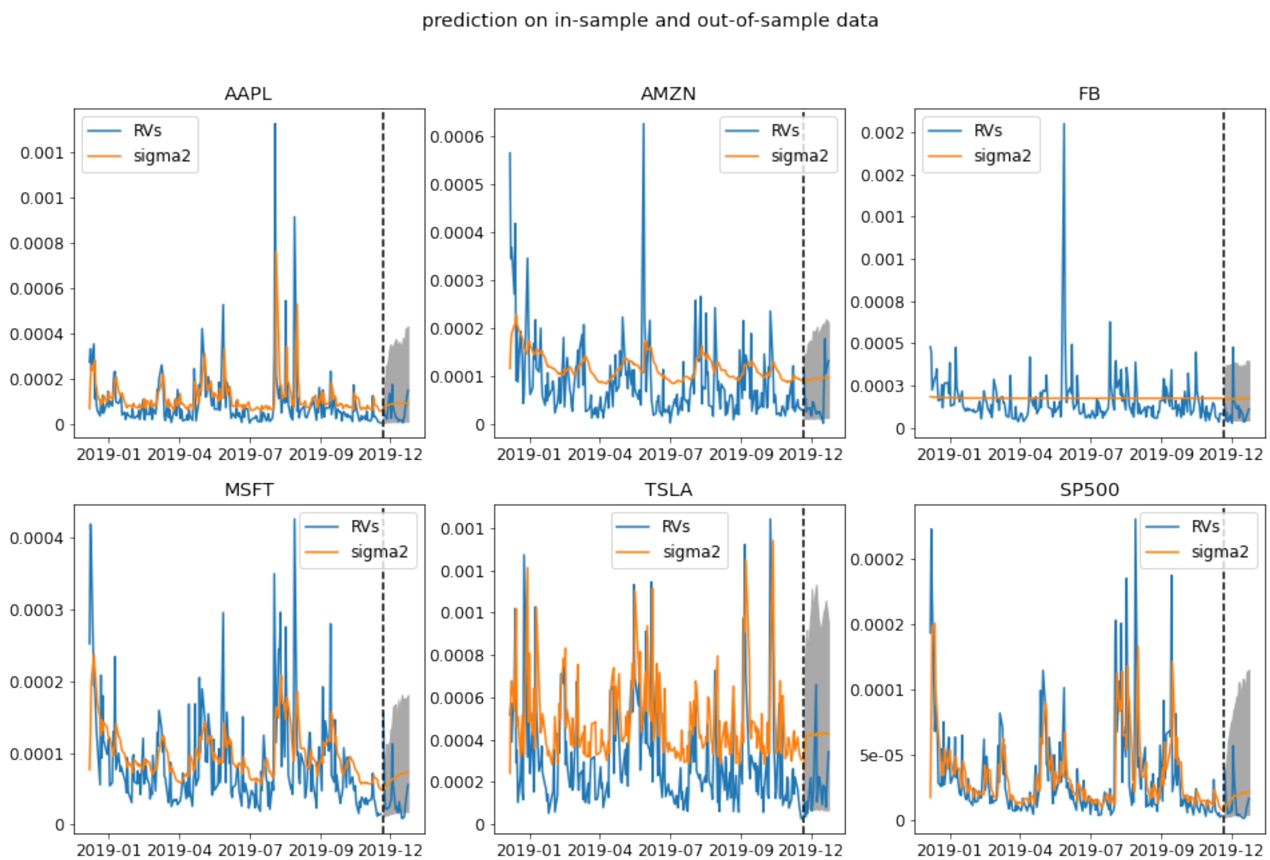


Figure 2: The model behaviour on the training and test data set. The fitted σ_t^2 on the training set is plotted as the orange line and the realized variance is plotted as the blue line. The gray area is 95% confidence interval for the last month realized variance.

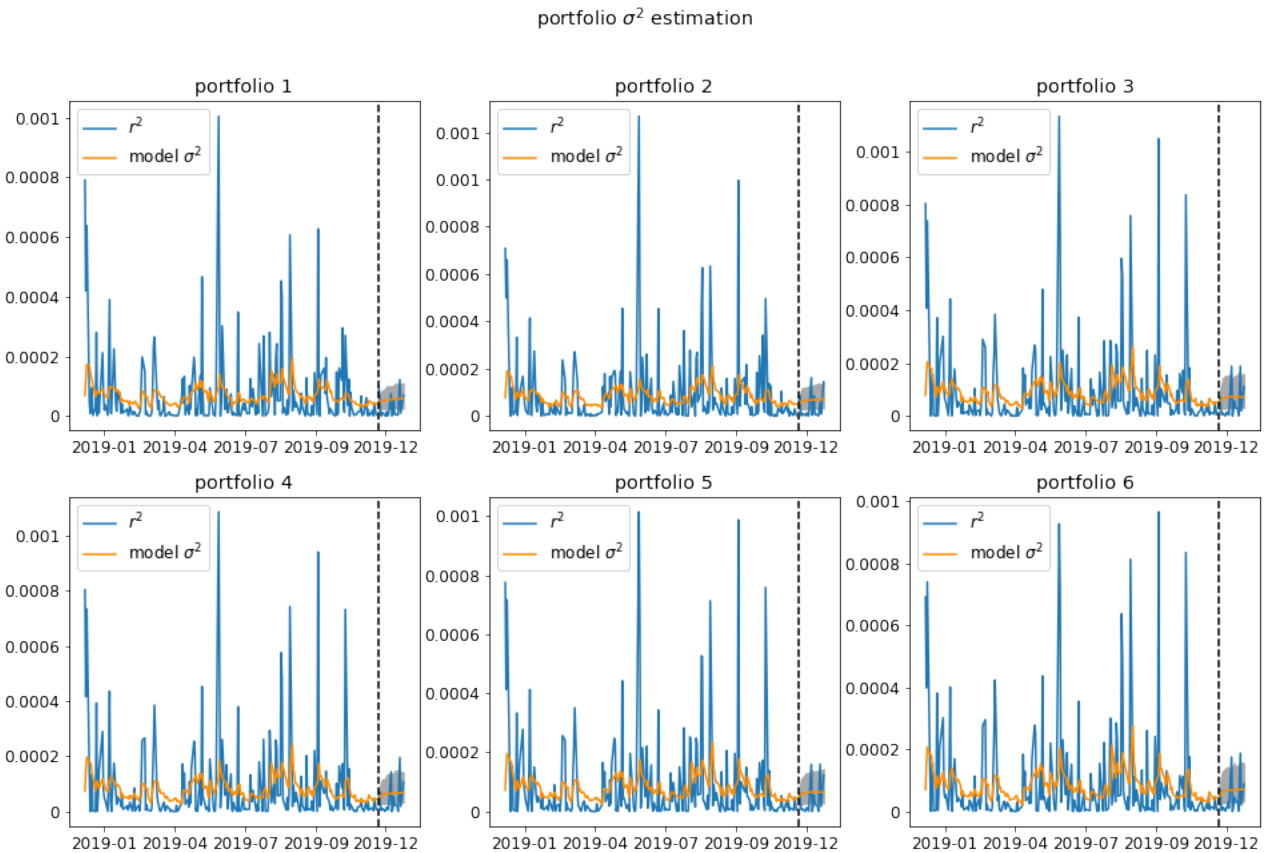


Figure 3: The model behaviour on 6 random portfolios with random weights. The fitted σ_t^2 on the training set is plotted as the orange line and the portfolio return squared r^2 is plotted as the blue line. The gray area is 95% confidence interval for the last month σ_t^2 forecast.

4 Discussion

The realized Garch DCC model we introduced in this project is a simple model that can captures the correlation among the shocks and incorporate high frequency realized information. The fit of the model needs more careful algorithm since our fit for Facebook does not give good result. Besides that, there are some points worth to note.

First, we introduce the log transformation in the measurement equation (2.5). If we simply use the linear measurement equation, the long horizon prediction will have mean reverting behaviour. This is because $E[RV_t] = \xi + \phi\hat{\sigma}_t^2$ is the best prediction for future points and $(\hat{\sigma})^2$ is the prediction of σ_t^2. The Garch equation would become

$$\hat{\sigma}_{t+1}^2 = \omega + \gamma\xi + (\beta + \gamma\phi)\hat{\sigma}_t^2 \tag{4.1}$$

and this is mean reverting given $\beta + \gamma\phi < 1$ and this puts extra constrains on the model. In our case, we use log transformation. This transformation makes the residual of the measurement equation to be more like normal and makes the behaviour of forecasting to be more complicated and the forecast require the use of Monte Carlo method.

Second, we consider a simple mean function in our model. The mean function could in principle have some interactions among the symbols. However, we did not consider this possibility explicitly. In our model, the correlation among the symbols come from shocks z_t. If we introduce transformation that diagonalize the covariance among z_i, we would also dependence of r_t^i on other r_t^j for $j \neq i$. Thus, the model assumes concurrent dependence implicitly. However, the σ_t^i still does not have any interaction among the symbols. The only effects of the correlation will appear on the quantities that relates to the covariance matrix of the returns, such as the portfolio volatility.

Third, in estimating the model, we employed the two step estimation method where we first estimate the Garch part and then the measurement equation parameters. Although, this method gives consistent estimate of the parameters, a more accurate way of estimation is helpful. Also, if the parameters are estimated together, the linear measurement equation can be estimated with the stationary constraints $\beta + \gamma\phi < 1$. In our case, the log transformation of the measurement equation removes the constraint and gives us freedom to do the two step estimation.

5 Conclusion

In the project, we constructed a simple correlated Garch model that can take into account of high frequency information. In particular, we constructed realized volatility from intraday data and incorporate it into each stock. In the model, we adopted the measurement view of realized variance and designed a measurement equation for realized variance to enable long horizon prediction. Besides, to preserve the assumption that the measurement is immersed in Gaussian noise and to introduce interesting forecast behaviour, we introduced the log transformation in the measurement equation. Another feature of the model is that the forecast and estimation of conditional variance for each stock is the same as the uncorrelated realized Garch model. We tested our model on the real data consists of AAPL, AMZN, FB, MSFT, TSLA and SP500. The results for FB and AMZN are not satisfactory but for other four symbols the model do pretty good estimation of conditional volatility. Further study is needed to find the reason for the estimation failure on FB and AMZN. However, overall, the realized Garch-DCC model is a simple enough model that could do satisfactory forecast.

References

[1] R. Engle, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 339–350, 2002.

[2] P. Jorion, "Predicting volatility in the foreign exchange market," *The Journal of Finance*, vol. 50, no. 2, pp. 507–528, 1995.

[3] T. G. Andersen and T. Bollerslev, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International economic review*, pp. 885–905, 1998.

[4] B. J. Blair, S.-H. Poon, and S. J. Taylor, "Forecasting s&p 100 volatility: the incremental information content of implied volatilities and high-frequency index returns," in *Handbook of quantitative finance and risk management*, pp. 1333–1344, Springer, 2010.

[5] M. Martens, "Measuring and forecasting s&p 500 index-futures volatility using high-frequency data," *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, vol. 22, no. 6, pp. 497–518, 2002.

[6] M. Martens and J. Zein, "Predicting financial volatility: High-frequency time-series forecasts vis-à-vis implied volatility," *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, vol. 24, no. 11, pp. 1005–1028, 2004.

[7] R. Engle, "New frontiers for arch models," *Journal of Applied Econometrics*, vol. 17, no. 5, pp. 425–446, 2002.

[8] P. R. Hansen and A. Lunde, “Forecasting volatility using high frequency data,” *The Oxford Handbook of Economic Forecasting*, 2011.

[9] I. Archakov, P. R. Hansen, and A. Lunde, “A multivariate realized garch model,” *arXiv preprint arXiv:2012.02708*, 2020.

[10] P. Christoffersen, *Elements of financial risk management.* Academic press, 2012.

[11] “Free intraday data.” https://firstratedata.com/free-intraday-data. Accessed: 2022-04-25.

Appendix

We first show our calculation of the derivative of the likelihood function for the DCC model. This is needed in the parameter estimation.

First the part of log likelihood that involves the parameters of DCC is

$$f = \log |C| + z^T C^{-1} z$$

and the parameters a and b are hidden in the calculation for C. We use chain rule to find the derivative for a and b

$$\frac{df}{da} = \frac{df}{dC_{ij}} \frac{dC_{ij}}{dQ_{kl}} \frac{dQ_{kl}}{da}.$$

In below calculation we do not employ Einstein summation notation since there are a lot of repeated indices that should not be summed. The first term is

$$\frac{df}{dC_{ij}} = C_{ij}^{-1} - \sum_{n,m} z_n C_{ni}^{-1} C_{jm}^{-1} z_m$$

Then the second second term in the chain rule is

$$\begin{aligned}
\frac{dC_{ij}}{dQ_{kl}} &= \frac{d}{dQ_{kl}} \left(\frac{1}{\sqrt{Q_{ii}}} Q_{ij} \frac{1}{\sqrt{Q_{jj}}}\right) \\
&= -\frac{1}{2} \frac{1}{Q_{ii}^{3/2}} \delta_{ik}\delta_{il} Q_{ij} \frac{1}{\sqrt{Q_{jj}}} + \frac{1}{\sqrt{Q_{ii}}} \frac{1}{\sqrt{Q_{jj}}} \delta_{ik}\delta_{jl} - \frac{1}{2} \frac{1}{\sqrt{Q_{ii}}} Q_{ij} \frac{1}{Q_{jj}^{3/2}} \delta_{jk}\delta_{jl} \\
&= -\frac{1}{2} \frac{C_{ij}}{Q_{ii}} \delta_{ik}\delta_{lk} - \frac{1}{2} \frac{C_{ij}}{Q_{jj}} \delta_{jk}\delta_{kl} + \frac{1}{\sqrt{Q_{ii}}} \frac{1}{\sqrt{Q_{jj}}} \delta_{ik}\delta_{jl}
\end{aligned}$$

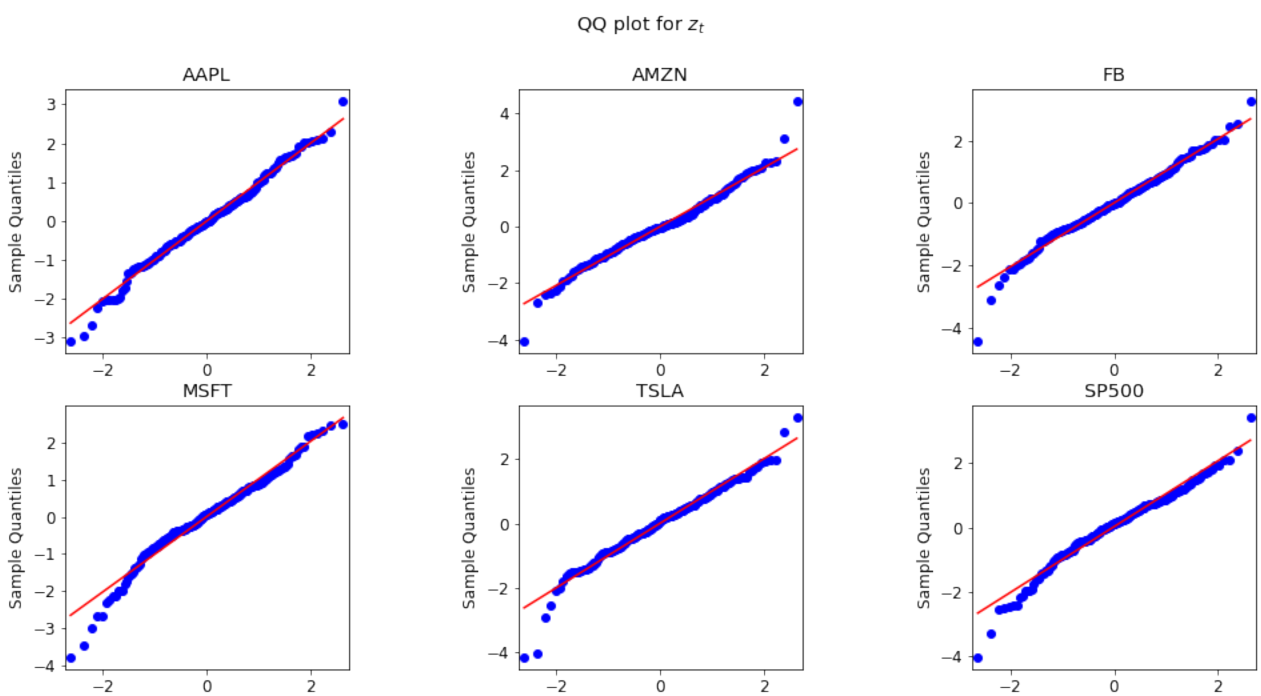
Thus combining the first two terms gives

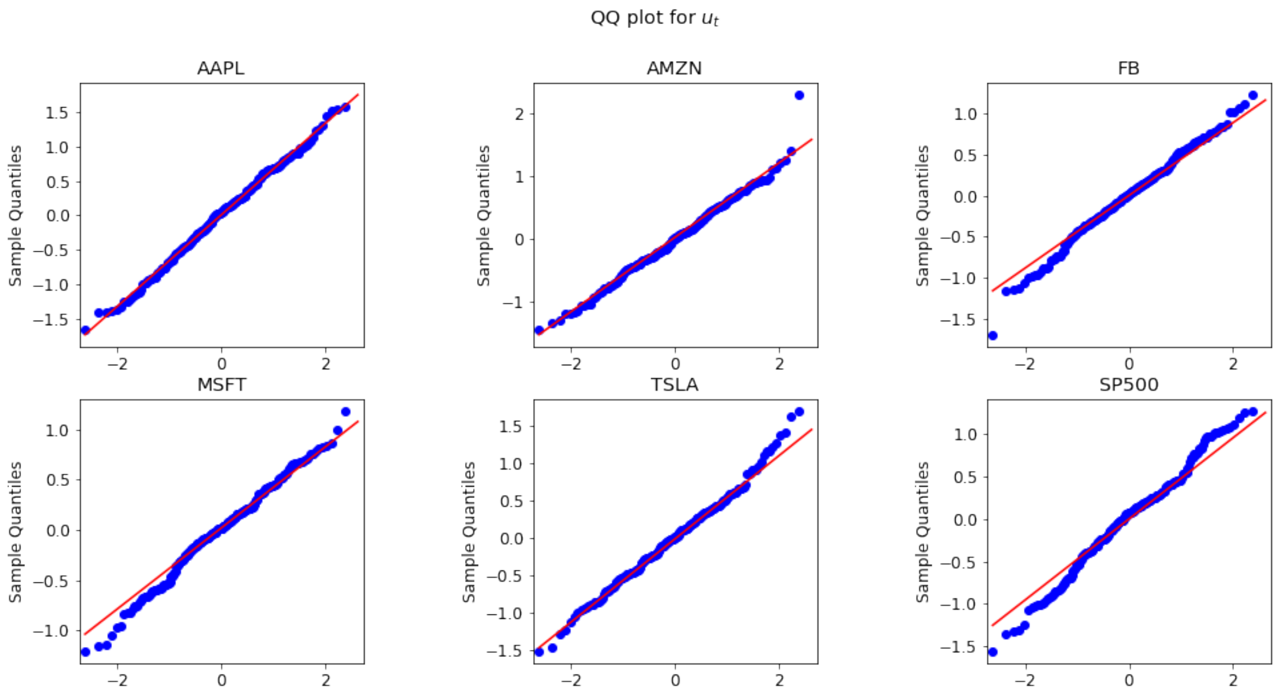
$$\begin{aligned}
&\frac{df}{dQ_{kl}} \\
&= \sum_{i,j} \left(C_{ij}^{-1} - z_n C_{ni}^{-1} C_{jm}^{-1} z_m\right) \left(-\frac{1}{2} \frac{C_{ij}}{Q_{ii}} \delta_{ik}\delta_{lk} - \frac{1}{2} \frac{C_{ij}}{Q_{jj}} \delta_{jk}\delta_{kl} + \frac{1}{\sqrt{Q_{ii}}} \frac{1}{\sqrt{Q_{jj}}} \delta_{ik}\delta_{jl}\right) \\
&= -\frac{1}{2} \frac{1}{Q_{kk}} C_{kj}^{-1} C_{kj} \delta_{lk} - \frac{1}{2} \frac{1}{Q_{kk}} C_{ik}^{-1} C_{ik} \delta_{kl} + \frac{C_{kl}^{-1}}{\sqrt{Q_{kk}}\sqrt{Q_{ll}}} - \frac{1}{\sqrt{Q_{kk}}} \frac{1}{\sqrt{Q_{ll}}} z_n C_{nk}^{-1} C_{lm}^{-1} z_n \\
&\quad + \frac{1}{2} z_n C_{nk}^{-1} C_{jm}^{-1} z_m \frac{C_{kj}}{Q_{kk}} \delta_{lk} + \frac{1}{2} z_n C_{ni}^{-1} C_{km}^{-1} z_m \frac{C_{ik}}{Q_{kk}} \delta_{kl} \\
&= -\frac{1}{Q_{kk}} \delta_{lk} + \frac{C_{kl}^{-1}}{\sqrt{Q_{kk}}\sqrt{Q_{ll}}} - \frac{1}{\sqrt{Q_{kk}}} \frac{1}{\sqrt{Q_{ll}}} z_n C_{nk}^{-1} C_{lm}^{-1} z_n + \frac{1}{2} \frac{1}{Q_{kk}} z_n C_{nk}^{-1} z_k \delta_{lk} + \frac{1}{2} \frac{1}{Q_{kk}} z_k C_{km}^{-1} z_m \delta_{lk} \\
&= \frac{1}{Q_{kk}} \delta_{lk} \left(-1 + \sum_n z_n C_{nk}^{-1} z_k\right) + Q_{kl}^{-1} - \sum_{n,m} \frac{1}{\sqrt{Q_{kk}}} \frac{1}{\sqrt{Q_{ll}}} z_n C_{nk}^{-1} C_{lm}^{-1} z_m
\end{aligned}$$

The last term in the chain rule
$$\frac{dQ_{kl}}{da} = z_k z_l - S_{kl} + b\frac{dQ_{kl,t-1}}{da}$$
which is a recursive equation and could be calculated easily in computer. As a result the derivative to a is a simple product of the above two equations.

Similarly, we could get the derivative with b by only replaceing the last term in the chain rule equation. The derivative of Q_t on b is
$$\frac{dQ_t}{db} = -S + Q_t + b\frac{dQ_{t-1}}{db}$$
and it is a recursive equation with the base case
$$\frac{dQ_1}{db} = -S + Q_0 + 0 = Q_0 - S$$
If we let Q_0 be the sample covariance matrix, the base case is $\frac{dQ_1}{db} = 0$. As a result, the derivative is
$$\begin{aligned}
\frac{df}{db} &= \left\{\frac{1}{Q_{kk}}\delta_{lk}\left(-1 + z_n C^{-1}_{nk} z_k\right) + Q^{-1}_{kl} - \frac{1}{\sqrt{Q_{kk}}}\frac{1}{\sqrt{Q_{ll}}} z_n C^{-1}_{nk} C^{-1}_{lm} z_m\right\}\frac{dQ_{kl}}{db} \\
&= \frac{1}{Q_{kk}}\frac{dQ_{kk}}{db}(-1 + z_n C^{-1}_{nk} z_k) + Q^{-1}_{kl}\frac{dQ_{kl}}{db} - \frac{1}{\sqrt{Q_{kk}}}\frac{1}{\sqrt{Q_{ll}}} z_n C^{-1}_{nk}\frac{dQ_{kl}}{db} C^{-1}_{lm} z_m \\
&= -\frac{1}{Q_{kk}}\frac{dQ_{kk}}{db} + z_n C^{-1}_{nk}\frac{z_k}{Q_{kk}}\frac{dQ_{kk}}{db} + Q^{-1}_{kl}\frac{dQ_{kl}}{db} - \frac{1}{\sqrt{Q_{kk}}}\frac{1}{\sqrt{Q_{ll}}} z_n C^{-1}_{nk}\frac{dQ_{kl}}{db} C^{-1}_{lm} z_m \\
&= -\frac{1}{Q_{kk}}\frac{dQ_{kk}}{db} + z_n\sqrt{Q_{nn}}Q^{-1}_{nk}\sqrt{Q_{kk}}z_k\frac{dQ_{kk}}{db}\frac{1}{Q_{kk}} + Q^{-1}_{kl}\frac{dQ_{kl}}{db} \\
&\quad - z_n\sqrt{Q_{nn}}Q^{-1}_{nk}\frac{dQ_{kl}}{db}Q^{-1}_{lm}\sqrt{Q_{mm}}z_m \\
&= -\frac{1}{Q_{kk}}\frac{dQ_{kk}}{db} + z'_n Q^{-1}_{nk} z'_k\frac{dQ_{kk}}{db}\frac{1}{Q_{kk}} + Q^{-1}_{kl}\frac{dQ_{kl}}{db} - z'_n Q^{-1}_{nk}\frac{dQ_{kl}}{db}Q^{-1}_{lm} z'_m
\end{aligned}$$

The second part of the appendix the diagnositc plot of the fitted residuals z_t and u_t in the Garch and measurement equation respectively. The plot shows that z_t and u_t roughly follows normal distribution. This shows that the estimation preserves the assumption.




QQ plot for u_t
AAPL
AMZN
FB
MSFT
TSLA
SP500
Sample Quantiles