Quantitative Risk Management
Homework for Week 3

John Dodson

September 27, 2021

Solutions to these problems are due at the beginning of the next session, which is 5:30 PM on Monday, October 4. Please turn in your solutions to the TA. Include your U of Minn. student identification number on your submission to facilitate recording marks in the Canvas learning management system. Also include the names of any classmates you consulted with in developing your solutions.

Data

Yahoo! Inc. contracts with Commodity Systems, Inc. to provide historical valuation data on many financial assets, including the Invesco PowerShares QQQ exchange-traded fund, which is managed to replicate the NASDAQ-100 index:

`https://finance.yahoo.com/quote/QQQ/history?p=QQQ`

CSI calculates a daily "adjusted close" timeseries which incorporates dividends and splits into historical prices and is a good source[1] for calculating log-returns.

Problems

For the following, acquire 500 trading days (about two years) of daily total returns of the QQQ through Friday, September 24, 2021[2].

1. Fit GARCH(1,1) using quasi-MLE (normal residuals) and report the parameters and the unconditional variance. **(3 points)**

2. Report which dates corresponded to the smallest and largest conditional variance forecasts during the period. **(2 points)**

3. Use the model to extract a sample of standardized residuals **(2 points)**. Assume that these residuals are i.i.d., and apply the Jarque-Bera normality test **(2 points)**. Can you reject normality with 95% confidence **(1 point)**?

[1] You are welcome to use a different source if you prefer, such as a Bloomberg terminal or a Quandl subscription.

[2] A previous version misspecified this date.

Assumptions

- You may assume that the conditional means of the daily total returns are zero.
- The JB test statistic is asymptotically (in sample size) $\chi^2(2)$ for a normal sample. For a $X \sim \chi^2(2)$, $\mathrm{P}(X > 5.991) \approx 0.05$. You may assume that our sample is large enough to use the asymptotic distribution.

Suggested Approach

You are welcome to use the BHHH algorithm from the case to estimate the parameters. This is efficient and fast. But it requires some development. Here is an alternative using a general function minimization algorithm such as Nelder-Mead:

1. define a function that returns a vector of conditional variances given a vector of innovations and numerical values for the parameters, $\vec{\sigma}^2(\vec{\varepsilon}, \theta)$;
2. implement the (negative) log-likelihood in terms of a vector of residuals and a vector of conditional variances, $H\left(\vec{\varepsilon}, \vec{\sigma}^2\right)$;
3. set up and solve the optimization problem $\hat{\theta} = \arg\min_\theta H\left(\vec{\varepsilon}, \vec{\sigma}^2(\vec{\varepsilon}, \theta)\right)$.

Solution

The timeseries is plotted in Figure 1.

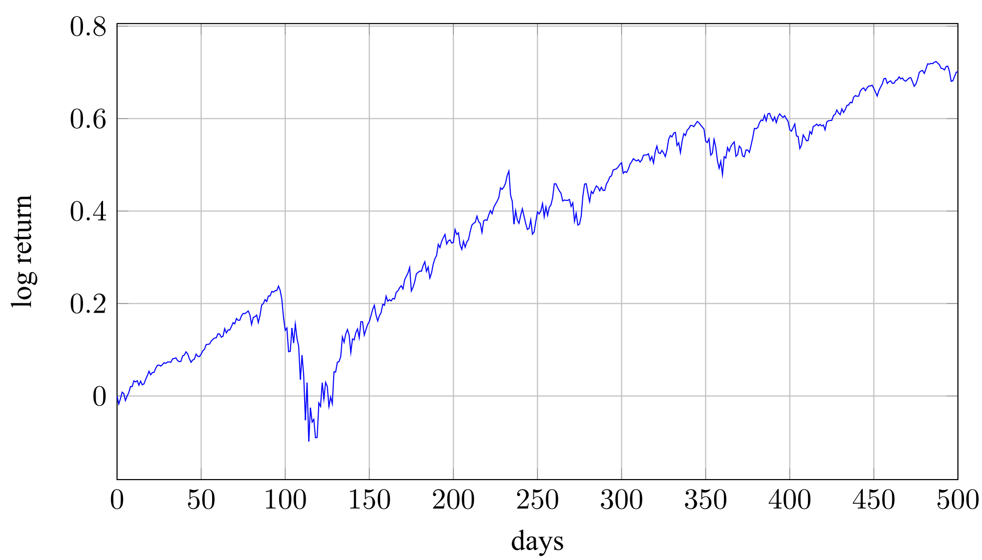

Figure 1: Cumulative log total return from October 10, 2019, through September 24, 2021.

The GARCH(1,1) linear recursion model for the conditional variance is

$$\sigma_1^2 = \frac{\alpha_0}{1-\alpha_1-\beta_1}$$
$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2$$

for $t = 2\ldots n = 500$ days. Recall $\mu_t = \mathrm{E}_{t-1}[X_t]$ is the conditional mean, $\sigma_t^2 = \mathrm{var}_{t-1}[X_t]$ is the conditional variance, and $\varepsilon_t = X_t - \mu_t$ is the residual. The subscripts on the expectation and variance operators represent the sigma algebra and $(X_t)_{t=1,\ldots,n}$ is the timeseries of invariants.

Estimation

The data extends back to the session on October 10, 2019. Since we are assuming $\mu_t = 0$, we can take the log-return innovations as residuals.

```
"data frame"
df = DataFrame(CSV.File("QQQ.csv"))
df.LogReturn=[missing; diff(log.(df."Adj Close"))]
```

The second line assumes that the rows are sorted ascending in date. Note there is no return for the oldest row.

In this next line, I use a boolean array and a symbol to select the log-returns up to and including last Friday. Then I sub-select the last five hundred per the specification.

```
"invariants"
x = df[df.:Date.<=Date(2021,9,24),:LogReturn][end-499:end]

"conditional means"
μ = zeros(500)

"residuals"
ε = x - μ
```

Per the hint, I suggest writing a method for evaluating the entire timeseries of conditional variances for a given set of parameters. This approach avoids any unnecessary recursion you might face were you chose to handle each term in the log-likelihood separately.

```
"GARCH(1,1) conditional variance with parameters θ = [α₀ α₁ β₁]"
function garch(ε,θ)
    α₀, α₁, β₁ = θ
    σ² = fill(NaN,length(ε))
    if α₀ > 0 && α₁ ≥ 0 && β₁ ≥ 0 && α₁+β₁ < 1
        σ²[1] = α₀ / (1-α₁-β₁)
        for i = 2:length(σ²)
            σ²[i] = α₀ + α₁*ε[i-1]^2 + β₁*σ²[i-1]
        end
    end
    return σ²
end
```

Notice that I initialized σ_1^2 with the implied unconditional variance. I am *not* using variance targeting here. For the version with variance targeting, there are only two parameters, not three, and you can set `α₀ = var(ε)*(1-α₁-β₁)`.

The maximum likelihood parameter fit is based on minimizing the entropy of the sample, which is the sample mean of the (negative) logarithm of the probability density. We are assuming $\varepsilon_t \sim \mathcal{N}\left(0, \sigma_t^2\right)$, so

```
"quasi-MLE entropy for GARCH"
function H(ε,σ²)
    return sum(log.(2π*σ²)+ε.^2 ./σ²)/2
end
```

Since we are using a search algorithm to identify the optimal parameters, we need a starting point. I recommend starting at a point with a relatively high β_1 and an implied unconditional variance close to the sample variance.

```
"initial value for parameter search"
θ₀ = [var(x)*0.1 0.2 0.7]
```

I used the standard algorithm in the `Optim` module for the solver,

```
qMLE = optimize(θ->H(ε,garch(ε,θ)),θ₀)
α₀, α₁, β₁ = qMLE.minimizer
```

The results are in Table 1. I got essentially identical results with the BHHH implementation from this week's case.

	without targeting	with targeting
$\hat{\alpha}_0$	6.96×10^{-6}	7.03×10^{-6}
$\hat{\alpha}_1$	0.191	0.201
$\hat{\beta}_1$	0.782	0.776
σ_0^2	$257. \times 10^{-6}$	$311. \times 10^{-6}$

Table 1: quasi-MLE parameters for GARCH(1,1) fit to the NASDAQ-100 daily returns

Recall that the variance targeting technique entails fixing the unconditional variance implied by the model to the sample variance of the residuals, thereby eliminating the need to estimate one of the parameters.

This involves a trade-off: While simplifying the numerical estimation and possibly reducing the standard errors of the parameter estimates, it also introduces a bias by mixing information from different layers of the filtration. In particular, it introduces the problematic assumption that the unknown σ_1^2 is equal to the sample variance of the next n observations. While logically consistent in the limit $n \to \infty$, for finite n it introduces a fallacy.

It seems that the consensus among professional econometricians is that variance targeting is valid and proper, particularly if your goal is to forecast near-term conditional variance. You can see from the results in the table that, while the sample and unconditional variance without targeting are somewhat different, the parameter estimates are similar and the one-period forecasts are within a few percent.

In the succeeding analysis, I will use the results without variance targeting.

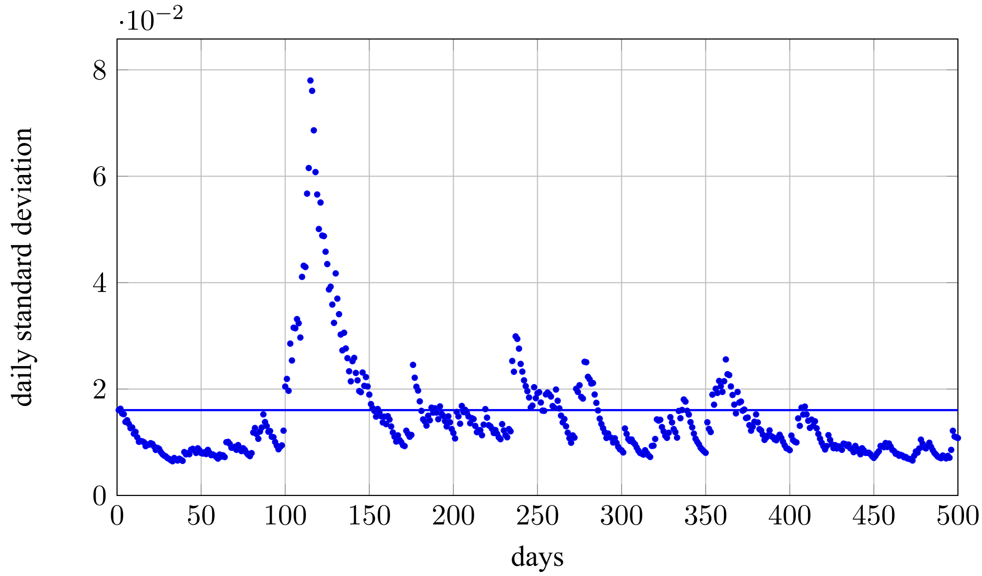

Figure 2: Forecasted conditional and unconditional variance of daily returns from October 10, 2019 to September 24, 2021

Extremes of Conditional Variance

It should come as no surprise that March 2020 contained the period of highest conditional variance is our data, centering on the week of March 16 to March 20. This corresponded to the first coordinated CoViD-19 pandemic response in the U.S. At it's peak, the daily conditional standard deviations was almost five times larger than the unconditional value.

There were two periods of relative calm: November 12 to 25, 2019; and August 13 to 17, 2021. During these periods, the daily standard deviation was about 40% of the unconditional value.

Normality of Residuals

If GARCH(1,1) is a correct description of the data, then the standardized residuals $\{z_i\}_{i=1,\ldots,n}$ where

$$z_i = \frac{\varepsilon_i}{\sqrt{\sigma_i^2}}$$

should be an i.i.d. sample.

Let's consider the hypothesis that the standardized residuals are normal. Looking at the histogram in Figure 3, the left tail in particular appears both heavier than the right tail and possibly heavier than a normal. The Jarque-Bare test is based on the observations that, asymptotically in sample size, the sample skewness and sample kurtosis of a normal sample are uncorrelated. The Central Limit Theorem says that both are also asymptotically normal. So an appropriately shifted and scaled sum of squares of these two statistics is asymptotically $\chi^2(2)$, in the family of Gamma random variables, and a normality test can be constructed.

It is not a particularly efficient test but it gives us a chance to review the descriptive statistics of the residuals, listed in Table 2. We would expect the sample mean to be close to zero (by assumption), and the sample variance to be close to one (by design). Higher sample moments are typically standardized.

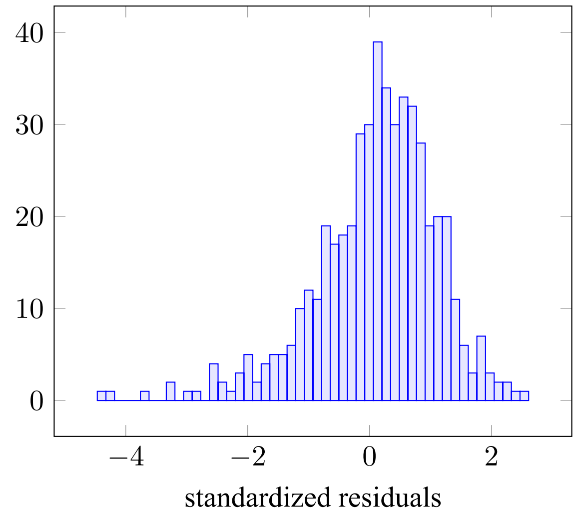

Figure 3: Histogram of standardized residuals

	sample	model
mean	0.109	0
variance	1.016	1
skew	-0.929	0
kurtosis	4.986	3

Table 2: sample moments of standardized residuals ($N = 500$)

The skewness seems to be too low and the the kurtosis too high. In fact, the JB statistic evaluates to about 156, which is clearly high enough to reject the hypothesis of normality (and any reasonable confidence).

Naively (without adjusting for conditional heteroskedasticity), negative skew and high kurtosis are typical of asset returns. But, since sample paths for asset total returns seem to be continuous, we would like to believe that there is a model for conditional variance that would yield normal residuals—at least for a sufficiently fine time grid. GARCH(1,1) seems to be a step in that direction, but it is probably not the final word.

Discussion

Considering the results plotted in Figure 2, it seems that the unconditional variance is a poor forecast of the GARCH conditional variance most of the time. In fact, about 75% of the time it is too high; and, when it is not too high, it can be substantially too low for an extended period.

Using unconditional variance forecasts in risk control can lead to a perception by users that the model is just a drag on business. Most of the time it seems consistently too conservative. Then, during periods when higher volatility does arise, the business can experience multiple outsized losses relative to the supposedly conservative forecast risk level. This dynamic can lead users and management to publicly disavow the model—sometimes spectacularly—in order to deflect blame for poor outcomes.

On the other hand, using a conditional variance forecast model in risk management can lead to radical swings in risk measurements, many of which seem like false alarms in retrospect. This “pro-cyclicality” in a regulatory or market utility setting can increase overall systemic fragility.