

Self directed activities 10

Josh Copeland (SID: 1444772)

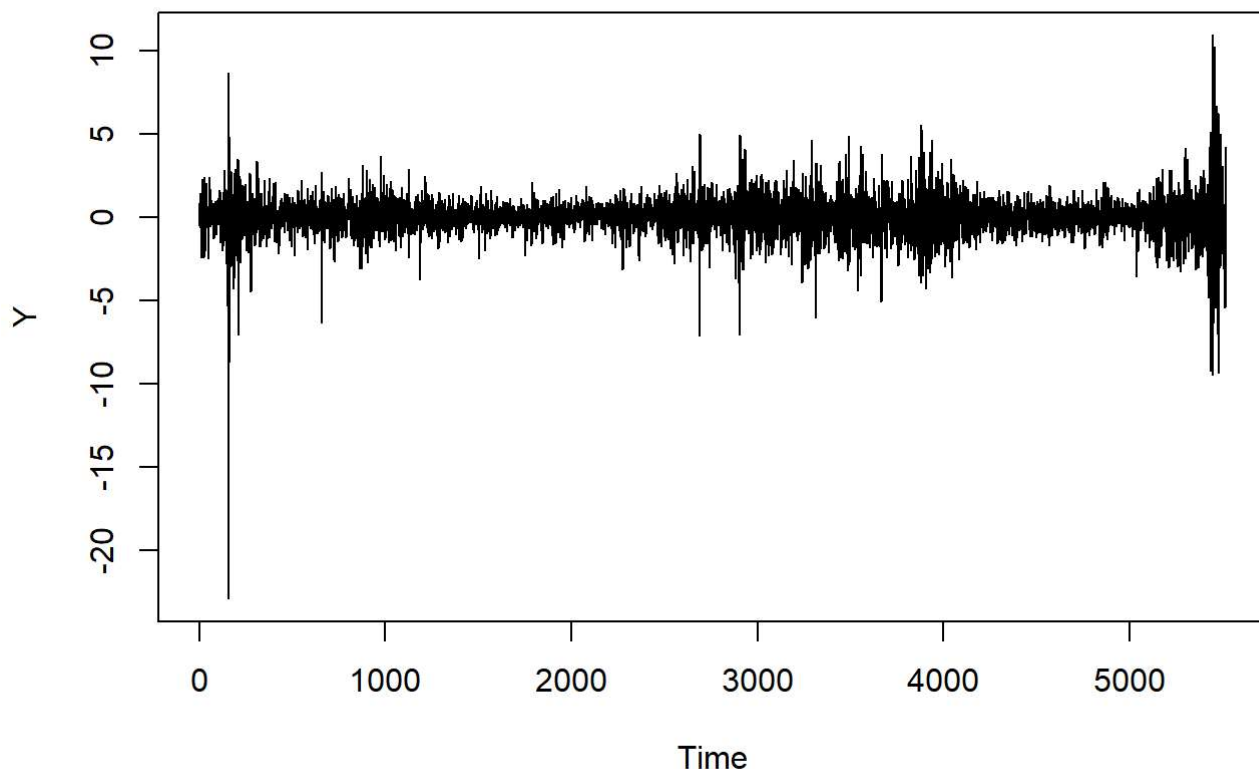
2025-10-08

Q2

Conditional heteroskedasticity is especially common in high frequency (eg. daily or weekly) financial time series. The rugarch package contains data on daily returns on the SP500 share index from 10-Mar-1987 to 30-Jan-2009. Here is how we can set up this data, using the last five days (business week) as the forecast period.

```
data("sp500ret")
Y <- sp500ret$SP500RET*100

hmax <- 5
n <- length(Y)-hmax
Yf <- ts(Y[(n+1):length(Y)], start=n+1)
Y <- ts(Y[1:n])
plot(Y)
```



(a)

Carry out an ARMA model selection search for this time series, using $p_{\max}=10$ and $q_{\max}=2$. Running this may take a minute, since ARMA model estimation with large sample sizes (over 5000) can take a bit longer.

```

# Model selection
pmax <- 10
qmax <- 2

# Storage for AIC and LBp
M <- matrix(nrow=(1+pmax)*(1+qmax), ncol=4)
colnames(M) <- c("p","q","AICc","LBp")
rownames(M) <- 1:nrow(M)
M[,1] <- rep(0:pmax,each=1+qmax)
M[,2] <- rep(0:qmax,1+pmax)

m <- 1
for (p in 0:pmax){
  for (q in 0:qmax){
    # Estimate ARMA(p,q) model
    eq <- Arima(Y, order=c(p,0,q))

    # Store AIC and LB p-value
    M[m,"AICc"] <- eq$aicc
    M[m,"LBp"] <- round(Box.test(eq$residuals, lag=p+q+5,
                                type="Ljung-Box", fitdf=p+q)$p.value,4)

    m <- m+1
  }
}

# Restrict to models that pass LB test
M <- M[which(M[, "LBp"]>0.05),]

# Minimise AICc
m_AICc <- which.min(M[, "AICc"])
p <- M[m_AICc, "p"]
q <- M[m_AICc, "q"]
cat(paste0("p = ", p, ", q = ", q))

```

```
## p = 5, q = 1
```

(b)

Estimate the chosen model and plot the inverse roots. What does this plot suggest about the model specification?

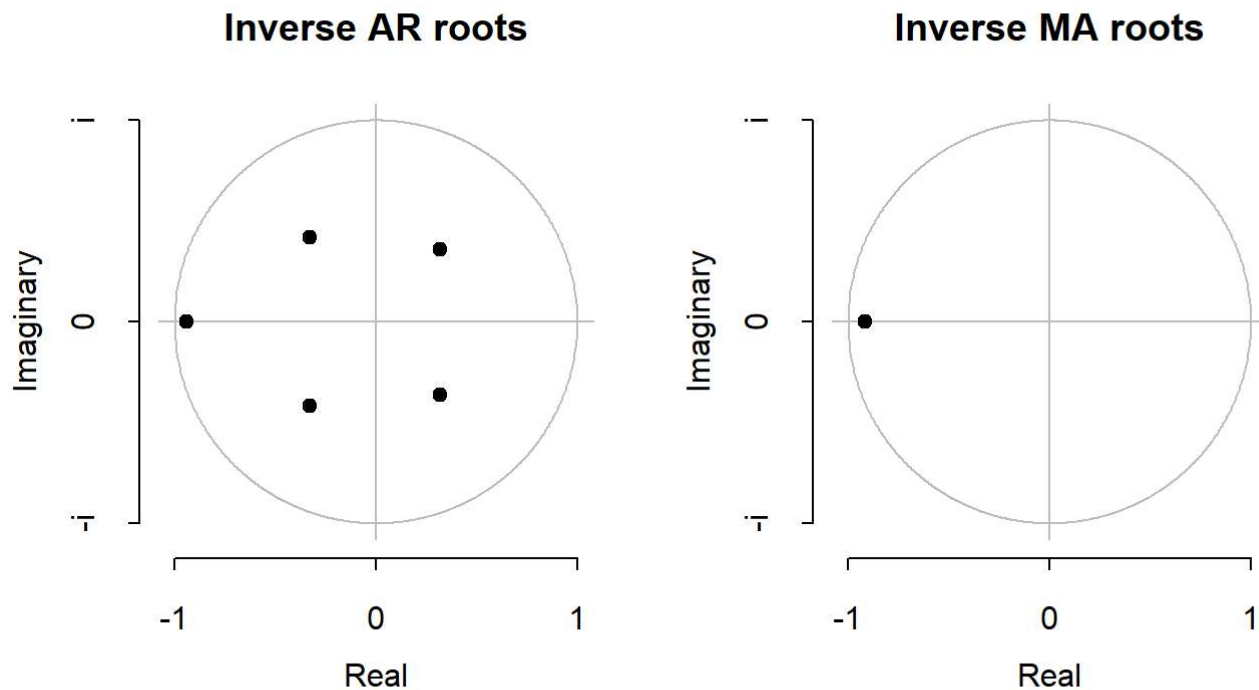
- There is a root in both the AR and MA unit circles which are very close to its border. Each suggest nonstationarity and noninvertibility respectively, which means the model is not well specified and should not be used for forecasting.

```

p <- 5
q <- 1

ARMApq <- Arima(Y, order=c(p,0,q))
plot(ARMApq)

```



(c)

Estimate AR(4)-ARCH(1) and AR(4)-GARCH(1,1) models for Y. Which of these two models is preferred, and why?

- Although both models pass the ARCH LM tests the AR(4)-ARCH(1) model has the lower AIC and is thus preferred as this indicates a better in-sample fit.

```
# estimate and fitr
Y_est <- window(Y, end = c(2024, 6))

AR4ARCH1 <- ugarchfit(
  data = Y_est,
  ugarchspec(
    variance.model = list(garchOrder = c(1, 0)),
    mean.model      = list(armaOrder  = c(4, 0))
  )
)

AR4GARCH11 <- ugarchfit(
  data = Y_est,
  ugarchspec(
    variance.model = list(garchOrder = c(1, 1)),
    mean.model      = list(armaOrder  = c(4, 0))
  )
)
show(AR4GARCH11)
```

```

##
## *-----*
## *          GARCH Model Fit          *
## *-----*
##
## Conditional Variance Dynamics
## -----
## GARCH Model   : sGARCH(1,1)
## Mean Model    : ARFIMA(4,0,0)
## Distribution   : norm
##
## Optimal Parameters
## -----
##      Estimate  Std. Error  t value Pr(>|t|)
## mu      0.047396   0.015227  3.11259 0.001855
## ar1      0.014947   0.025032  0.59713 0.550420
## ar2     -0.003401   0.024709 -0.13764 0.890524
## ar3     -0.036222   0.024480 -1.47963 0.138971
## ar4     -0.075179   0.024242 -3.10121 0.001927
## omega    0.027461   0.007453  3.68478 0.000229
## alpha1   0.117523   0.016032  7.33061 0.000000
## beta1    0.859179   0.020423 42.06827 0.000000
##
## Robust Standard Errors:
##      Estimate  Std. Error  t value Pr(>|t|)
## mu      0.047396   0.016972  2.79262 0.005228
## ar1      0.014947   0.026589  0.56217 0.573999
## ar2     -0.003401   0.022253 -0.15283 0.878531
## ar3     -0.036222   0.023580 -1.53614 0.124504
## ar4     -0.075179   0.026528 -2.83391 0.004598
## omega    0.027461   0.021957  1.25067 0.211055
## alpha1   0.117523   0.076498  1.53630 0.124465
## beta1    0.859179   0.080272 10.70334 0.000000
##
## LogLikelihood : -2525.395
##
## Information Criteria
## -----
##
## Akaike          2.4972
## Bayes           2.5193
## Shibata         2.4972
## Hannan-Quinn    2.5053
##
## Weighted Ljung-Box Test on Standardized Residuals
## -----
##              statistic p-value
## Lag[1]                3.001 0.08322
## Lag[2*(p+q)+(p+q)-1][11] 5.950 0.52010
## Lag[4*(p+q)+(p+q)-1][19] 8.217 0.76331
## d.o.f=4
## H0 : No serial correlation
##
## Weighted Ljung-Box Test on Standardized Squared Residuals
## -----

```

```

##                                statistic p-value
## Lag[1]                        0.8224  0.3645
## Lag[2*(p+q)+(p+q)-1][5]      1.0241  0.8542
## Lag[4*(p+q)+(p+q)-1][9]      1.5322  0.9529
## d.o.f=2
##
## Weighted ARCH LM Tests
## -----
##              Statistic Shape Scale P-Value
## ARCH Lag[3]   0.05308 0.500 2.000  0.8178
## ARCH Lag[5]   0.25094 1.440 1.667  0.9530
## ARCH Lag[7]   0.70711 2.315 1.543  0.9560
##
## Nyblom stability test
## -----
## Joint Statistic:  2.2499
## Individual Statistics:
## mu      0.1526
## ar1     0.1319
## ar2     0.2359
## ar3     0.1816
## ar4     0.1728
## omega   0.8566
## alpha1  0.6751
## beta1   0.8172
##
## Asymptotic Critical Values (10% 5% 1%)
## Joint Statistic:      1.89 2.11 2.59
## Individual Statistic:  0.35 0.47 0.75
##
## Sign Bias Test
## -----
##              t-value      prob sig
## Sign Bias      0.09511 0.924235
## Negative Sign Bias 3.17268 0.001533 ***
## Positive Sign Bias 0.98962 0.322480
## Joint Effect    15.10342 0.001730 ***
##
##
## Adjusted Pearson Goodness-of-Fit Test:
## -----
##   group statistic p-value(g-1)
## 1    20      112.2   3.100e-15
## 2    30      120.3   4.311e-13
## 3    40      138.0   5.445e-13
## 4    50      151.5   2.108e-12
##
##
## Elapsed time : 0.2339649

```

```
# weighted ARCH LM tests
```

```
summaryAR4GARCH11 <- capture.output(show(AR4GARCH11))
r1 <- which(summaryAR4GARCH11 == "Weighted ARCH LM Tests")
r2 <- which(summaryAR4GARCH11 == "Nyblom stability test") - 2
ARCHLMtests_AR4GARCH11 <- summaryAR4GARCH11[r1:r2]
print(ARCHLMtests_AR4GARCH11, quote = "FALSE")
```

```
## [1] Weighted ARCH LM Tests
## [2] -----
## [3]           Statistic Shape Scale P-Value
## [4] ARCH Lag[3]    0.05308 0.500 2.000  0.8178
## [5] ARCH Lag[5]    0.25094 1.440 1.667  0.9530
## [6] ARCH Lag[7]    0.70711 2.315 1.543  0.9560
```

```
# AIC comparison
```

```
AIC_AR4 <- Arima(Y_est, order = c(4,0,0))$aic
AIC_AR4ARCH1 <- infocriteria(AR4ARCH1)[1] * length(Y_est)
AIC_AR4GARCH11 <- infocriteria(AR4GARCH11)[1] * length(Y_est)

cat(paste0(
  "AIC AR(4)           = ", round(AIC_AR4, 2), "\n",
  "AIC AR(4)-ARCH(1)   = ", round(AIC_AR4ARCH1, 2), "\n",
  "AIC AR(4)-GARCH(1,1) = ", round(AIC_AR4GARCH11, 2), "\n"
))
```

```
## AIC AR(4)           = 6003.95
## AIC AR(4)-ARCH(1)   = 5335.41
## AIC AR(4)-GARCH(1,1) = 5066.79
```