單變量黃金時間數列分析

林子立

2025-06-03

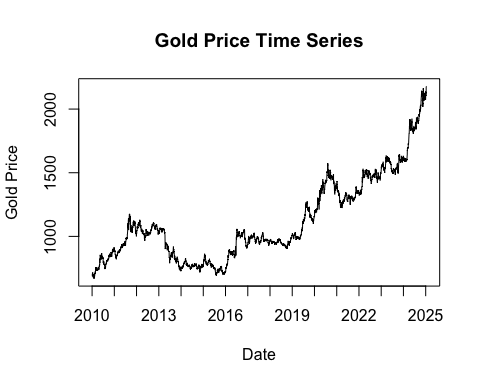
—Section 1: Time Series Analysis for the Gold Price (XUFIX)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ ggplot2 3.5.1 ✔ dplyr 1.1.4  
## ✔ tibble 3.2.1 ✔ stringr 1.5.0  
## ✔ tidyr 1.3.0 ✔ forcats 0.5.2  
## ✔ purrr 1.0.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
##   
## 載入套件：'zoo'  
##   
##   
## 下列物件被遮斷自 'package:base':  
##   
## as.Date, as.Date.numeric  
##   
##   
## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo   
##   
## Rows: 14497 Columns: 2  
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## dbl (1): LBMA Gold Prices - daily - euro - AM (LBMA/gold\_D/gold\_D\_EUR\_AM)  
## date (1): period  
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.  
## Rows: 10268 Columns: 2  
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## dbl (1): DCOILWTICO  
## date (1): observation\_date  
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.  
##   
## 載入套件：'aTSA'  
##   
##   
## 下列物件被遮斷自 'package:tseries':  
##   
## adf.test, kpss.test, pp.test  
##   
##   
## 下列物件被遮斷自 'package:graphics':  
##   
## identify  
##   
##   
##   
## 載入套件：'forecast'  
##   
##   
## 下列物件被遮斷自 'package:aTSA':  
##   
## forecast  
##   
##   
## 載入需要的套件：parallel  
##   
##   
## 載入套件：'rugarch'  
##   
##   
## 下列物件被遮斷自 'package:purrr':  
##   
## reduce  
##   
##   
## 下列物件被遮斷自 'package:stats':  
##   
## sigma

#Step1: Exploratory Data Analysis

##Data Visualization

# 時間向量與價格  
dates <- AU\_OIL$Date  
gold <- AU\_OIL$GoldPrice\_interp  
  
# 繪圖（線圖，不畫 x 軸）  
plot(dates, gold, type = "l",  
 main = "Gold Price Time Series",  
 ylab = "Gold Price",  
 xlab = "Date",  
 xaxt = "n") # 不畫預設 x 軸  
  
# 建立每年 1 月 1 日的日期向量  
years <- seq(from = as.Date("2010-01-01"), to = as.Date("2025-01-01"), by = "year")  
  
# 標上年份作為 x 軸刻度  
axis(side = 1, at = years, labels = format(years, "%Y"))

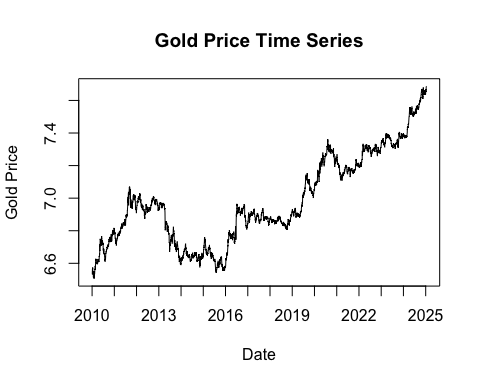


adf.test(gold)

## Augmented Dickey-Fuller Test   
## alternative: stationary   
##   
## Type 1: no drift no trend   
## lag ADF p.value  
## [1,] 0 2.71 0.99  
## [2,] 1 2.52 0.99  
## [3,] 2 2.46 0.99  
## [4,] 3 2.50 0.99  
## [5,] 4 2.54 0.99  
## [6,] 5 2.46 0.99  
## [7,] 6 2.47 0.99  
## [8,] 7 2.46 0.99  
## [9,] 8 2.53 0.99  
## [10,] 9 2.58 0.99  
## Type 2: with drift no trend   
## lag ADF p.value  
## [1,] 0 1.37 0.99  
## [2,] 1 1.18 0.99  
## [3,] 2 1.11 0.99  
## [4,] 3 1.16 0.99  
## [5,] 4 1.24 0.99  
## [6,] 5 1.17 0.99  
## [7,] 6 1.20 0.99  
## [8,] 7 1.16 0.99  
## [9,] 8 1.25 0.99  
## [10,] 9 1.31 0.99  
## Type 3: with drift and trend   
## lag ADF p.value  
## [1,] 0 -0.133 0.990  
## [2,] 1 -0.319 0.990  
## [3,] 2 -0.383 0.987  
## [4,] 3 -0.333 0.989  
## [5,] 4 -0.288 0.990  
## [6,] 5 -0.362 0.988  
## [7,] 6 -0.345 0.989  
## [8,] 7 -0.356 0.988  
## [9,] 8 -0.287 0.990  
## [10,] 9 -0.233 0.990  
## ----   
## Note: in fact, p.value = 0.01 means p.value <= 0.01

此圖呈現的是自2010年以來，LMBA黃金每日拍賣價（AM）的價格。從圖中可以看到，價格以大週期來看可以以2016年為分界找到兩個趨勢。在2016年以前黃金價格經歷先升後降的趨勢；2016年之後則是持續攀升。其中值得注意的現象是，黃金價格兩、三年就會存在一次跳動情形，造成短期價格劇烈變化的情況，如2013年終、2016年底、2019年底、2021年初及2024年初。 除了部分時間存在劇烈跳動的價格變化以外，長期趨勢也存在單調遞增與增加速度加快的趨勢。從此觀點出發，合理懷疑此資料序列存在單根。近一步使用Augmented Dickey-Fuller檢定（以下簡稱ADF檢定），可以發現不管有無趨勢(Trend)或飄移（Drift）的加入，以及滯後期為0-9期之間，強烈的證據都指向單根存在於此序列資料中。為了符合對於穩定態（Stationary）資料的統計假設，以下使用常見的資料處理方法：取對數及差分進行探討。

# 時間向量與價格  
dates <- AU\_OIL$Date  
gold <- log(AU\_OIL$GoldPrice\_interp)  
  
# 繪圖（線圖，不畫 x 軸）  
plot(dates, gold, type = "l",  
 main = "Gold Price Time Series",  
 ylab = "Gold Price",  
 xlab = "Date",  
 xaxt = "n") # 不畫預設 x 軸  
  
# 建立每年 1 月 1 日的日期向量  
years <- seq(from = as.Date("2010-01-01"), to = as.Date("2025-01-01"), by = "year")  
  
# 標上年份作為 x 軸刻度  
axis(side = 1, at = years, labels = format(years, "%Y"))

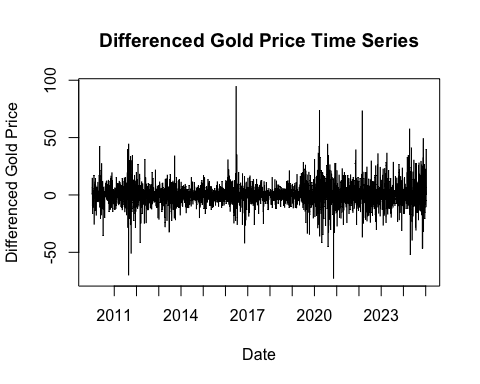


adf.test(gold)

## Augmented Dickey-Fuller Test   
## alternative: stationary   
##   
## Type 1: no drift no trend   
## lag ADF p.value  
## [1,] 0 2.14 0.990  
## [2,] 1 2.01 0.990  
## [3,] 2 1.97 0.988  
## [4,] 3 1.99 0.989  
## [5,] 4 1.98 0.989  
## [6,] 5 1.92 0.986  
## [7,] 6 1.92 0.986  
## [8,] 7 1.92 0.987  
## [9,] 8 1.93 0.987  
## [10,] 9 1.92 0.987  
## Type 2: with drift no trend   
## lag ADF p.value  
## [1,] 0 0.00958 0.956  
## [2,] 1 -0.11378 0.944  
## [3,] 2 -0.15154 0.939  
## [4,] 3 -0.12829 0.942  
## [5,] 4 -0.04057 0.952  
## [6,] 5 -0.09845 0.946  
## [7,] 6 -0.05907 0.951  
## [8,] 7 -0.11796 0.944  
## [9,] 8 -0.07078 0.950  
## [10,] 9 -0.04267 0.952  
## Type 3: with drift and trend   
## lag ADF p.value  
## [1,] 0 -1.15 0.914  
## [2,] 1 -1.30 0.876  
## [3,] 2 -1.34 0.857  
## [4,] 3 -1.31 0.869  
## [5,] 4 -1.27 0.887  
## [6,] 5 -1.34 0.857  
## [7,] 6 -1.32 0.865  
## [8,] 7 -1.35 0.853  
## [9,] 8 -1.32 0.867  
## [10,] 9 -1.31 0.872  
## ----   
## Note: in fact, p.value = 0.01 means p.value <= 0.01

首先嘗試的分析方法為一次差分。從圖中可以發現資料仍舊存在短期的劇烈變動以及穩定的增加趨勢。而且可以發現此增加趨勢似乎沒有緩和的現象，表明此時間數列可能存在單根。是以，近一步透過ADF檢定方法，分析趨勢、飄移及滯後期之下的資料是否存在單根。從檢定結果的報表中可以發現不管有無趨勢(Trend)或飄移（Drift）的加入，以及滯後期為0-9期之間，強烈的證據都指向單根存在於此序列資料中。為了符合對於穩定態（Stationary）資料的統計假設，以下使用常見的資料處理方法：取對數及差分進行探討。

# Step 1: 差分後的價格  
gold\_diff <- diff(AU\_OIL$GoldPrice\_interp)  
  
# Step 2: 調整時間向量（去掉第一天，因為差分少一天）  
dates\_diff <- AU\_OIL$Date[-1] # 或 tail(dates, -1)  
  
# Step 3: 繪圖（關閉預設 x 軸）  
plot(dates\_diff, gold\_diff, type = "l",  
 main = "Differenced Gold Price Time Series",  
 ylab = "Differenced Gold Price",  
 xlab = "Date",  
 xaxt = "n")  
  
# Step 4: 標出每年 1 月 1 日的位置  
years <- seq(from = as.Date("2010-01-01"), to = as.Date("2025-01-01"), by = "year")  
  
# 只選擇 years 有在日期範圍內的  
valid\_years <- years[years %in% dates\_diff]  
  
# Step 5: 畫上 x 軸標籤（以年份顯示）  
axis(side = 1, at = valid\_years, labels = format(valid\_years, "%Y"))

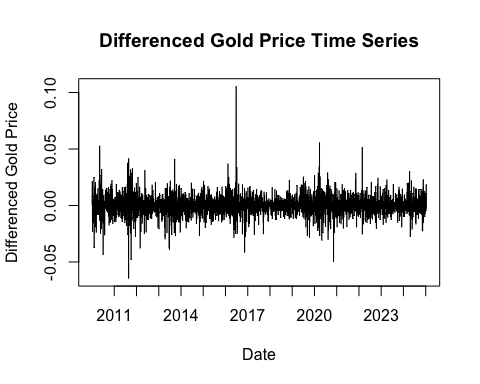


adf.test(diff(AU\_OIL$GoldPrice\_interp))

## Augmented Dickey-Fuller Test   
## alternative: stationary   
##   
## Type 1: no drift no trend   
## lag ADF p.value  
## [1,] 0 -69.8 0.01  
## [2,] 1 -49.7 0.01  
## [3,] 2 -41.9 0.01  
## [4,] 3 -37.0 0.01  
## [5,] 4 -32.2 0.01  
## [6,] 5 -29.6 0.01  
## [7,] 6 -27.3 0.01  
## [8,] 7 -26.2 0.01  
## [9,] 8 -25.1 0.01  
## [10,] 9 -23.4 0.01  
## Type 2: with drift no trend   
## lag ADF p.value  
## [1,] 0 -69.8 0.01  
## [2,] 1 -49.8 0.01  
## [3,] 2 -41.9 0.01  
## [4,] 3 -37.1 0.01  
## [5,] 4 -32.2 0.01  
## [6,] 5 -29.7 0.01  
## [7,] 6 -27.4 0.01  
## [8,] 7 -26.3 0.01  
## [9,] 8 -25.2 0.01  
## [10,] 9 -23.5 0.01  
## Type 3: with drift and trend   
## lag ADF p.value  
## [1,] 0 -69.9 0.01  
## [2,] 1 -49.8 0.01  
## [3,] 2 -42.0 0.01  
## [4,] 3 -37.1 0.01  
## [5,] 4 -32.3 0.01  
## [6,] 5 -29.8 0.01  
## [7,] 6 -27.5 0.01  
## [8,] 7 -26.4 0.01  
## [9,] 8 -25.3 0.01  
## [10,] 9 -23.6 0.01  
## ----   
## Note: in fact, p.value = 0.01 means p.value <= 0.01

其次，嘗試的資料屬於進行差分處理的資料，從資料來看，雖然仍舊有部分時間的資料有劇烈波動，但是相較前兩個圖形，明顯少了穩定上升的現象。進一步檢測資料的單根，可以發現與前兩個方法完全不同的結果。不論趨勢是否存在以及落後期數，在以0.05的信心水準之下，所以組合的檢定結果都表明這個方法處理的資料不具有單根，合理推測此筆資料適合進行分析。總結來看，透過差分處理的資料不再有單根，僅剩劇烈的波動問題存在於此筆資料。

# Step 1: 差分後的價格  
gold\_diff <- diff(log(AU\_OIL$GoldPrice\_interp))  
  
# Step 2: 調整時間向量（去掉第一天，因為差分少一天）  
dates\_diff <- AU\_OIL$Date[-1] # 或 tail(dates, -1)  
  
# Step 3: 繪圖（關閉預設 x 軸）  
plot(dates\_diff, gold\_diff, type = "l",  
 main = "Differenced Gold Price Time Series",  
 ylab = "Differenced Gold Price",  
 xlab = "Date",  
 xaxt = "n")  
  
# Step 4: 標出每年 1 月 1 日的位置  
years <- seq(from = as.Date("2010-01-01"), to = as.Date("2025-01-01"), by = "year")  
  
# 只選擇 years 有在日期範圍內的  
valid\_years <- years[years %in% dates\_diff]  
  
# Step 5: 畫上 x 軸標籤（以年份顯示）  
axis(side = 1, at = valid\_years, labels = format(valid\_years, "%Y"))



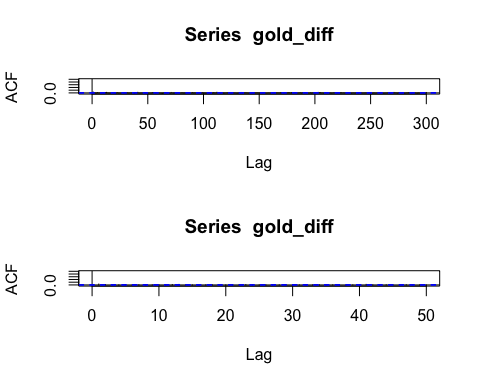
adf.test(diff(log(AU\_OIL$GoldPrice\_interp)))

## Augmented Dickey-Fuller Test   
## alternative: stationary   
##   
## Type 1: no drift no trend   
## lag ADF p.value  
## [1,] 0 -69.5 0.01  
## [2,] 1 -49.7 0.01  
## [3,] 2 -41.8 0.01  
## [4,] 3 -36.8 0.01  
## [5,] 4 -31.9 0.01  
## [6,] 5 -29.5 0.01  
## [7,] 6 -27.1 0.01  
## [8,] 7 -25.7 0.01  
## [9,] 8 -24.4 0.01  
## [10,] 9 -22.8 0.01  
## Type 2: with drift no trend   
## lag ADF p.value  
## [1,] 0 -69.6 0.01  
## [2,] 1 -49.8 0.01  
## [3,] 2 -41.8 0.01  
## [4,] 3 -36.9 0.01  
## [5,] 4 -32.0 0.01  
## [6,] 5 -29.6 0.01  
## [7,] 6 -27.2 0.01  
## [8,] 7 -25.8 0.01  
## [9,] 8 -24.5 0.01  
## [10,] 9 -22.9 0.01  
## Type 3: with drift and trend   
## lag ADF p.value  
## [1,] 0 -69.6 0.01  
## [2,] 1 -49.8 0.01  
## [3,] 2 -41.8 0.01  
## [4,] 3 -36.9 0.01  
## [5,] 4 -32.0 0.01  
## [6,] 5 -29.6 0.01  
## [7,] 6 -27.2 0.01  
## [8,] 7 -25.8 0.01  
## [9,] 8 -24.5 0.01  
## [10,] 9 -22.9 0.01  
## ----   
## Note: in fact, p.value = 0.01 means p.value <= 0.01

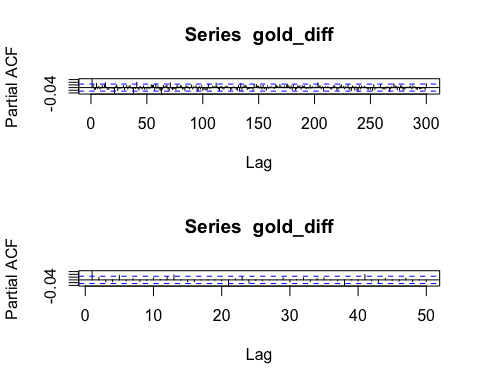
再來，嘗試結合差分與對數處理的方法。從圖中可以看到劇烈波動趨於穩定，且同樣少了長期穩定上升的趨勢，合理推測此種方法不只消除單根、也緩和劇烈波動的現象。進一步檢測資料的單根，可以發現與前兩個方法完全不同的結果。不論趨勢是否存在以及落後期數，在以0.05的信心水準之下，所以組合的檢定結果都表明這個方法處理的資料不具有單根，合理推測此筆資料適合進行分析。 總結個段資料處理方法可以發現，原始資料存在的劇烈波動及穩定上升趨勢（單根），透過結合差分及對數的處理為相對較佳的方法。故而，後續將以此方法進行深入分析與建立解釋模型。

##Data mining

par(mfrow=c(2,1))  
acf(gold\_diff, lag.max=300)  
acf(gold\_diff, lag.max=50)



pacf(gold\_diff, lag.max=300)  
pacf(gold\_diff, lag.max=50)



par(mfrow=c(1,1))

AR(1) ,MA(1)

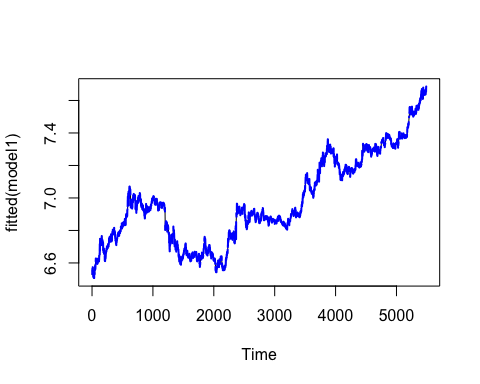
model1 <- arima(log(AU\_OIL$GoldPrice\_interp), order = c(1,1,1))  
model1

##   
## Call:  
## arima(x = log(AU\_OIL$GoldPrice\_interp), order = c(1, 1, 1))  
##   
## Coefficients:  
## ar1 ma1  
## 0.2256 -0.1624  
## s.e. 0.1665 0.1678  
##   
## sigma^2 estimated as 5.253e-05: log likelihood = 19256.16, aic = -38506.31

plot(fitted(model1))  
nrow(AU\_OIL)

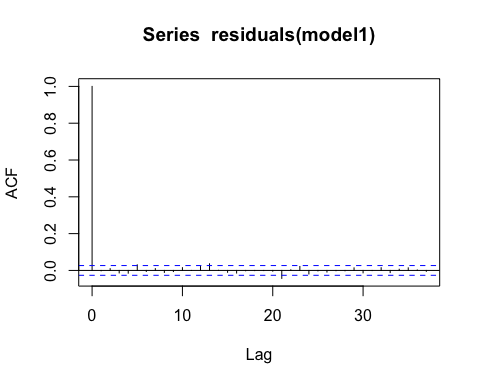
## [1] 5490

lines(1:nrow(AU\_OIL), y=log(AU\_OIL$GoldPrice), type="l", lwd=2, col="blue")

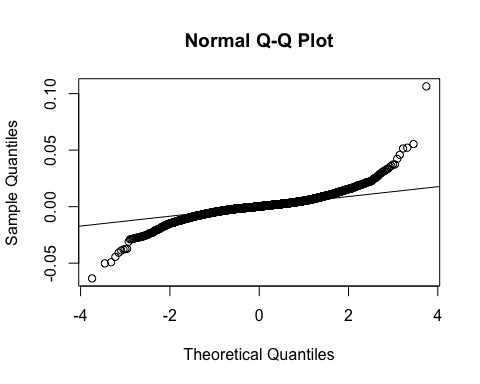


AR1, MA1的模型看起來不錯，

acf(residuals(model1))



qqnorm(residuals(model1)); qqline(residuals(model1))



ks.test(residuals(model1), "pnorm")

##   
## Asymptotic one-sample Kolmogorov-Smirnov test  
##   
## data: residuals(model1)  
## D = 0.48664, p-value < 2.2e-16  
## alternative hypothesis: two-sided

library(nortest)  
ad.test(residuals(model1))

##   
## Anderson-Darling normality test  
##   
## data: residuals(model1)  
## A = 119.85, p-value < 2.2e-16

Box.test(residuals(model1), lag = 20, type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: residuals(model1)  
## X-squared = 21.921, df = 20, p-value = 0.3448

根據 Gauss-Markov 定理，線性模型若要具備最佳線性無偏估計量（BLUE）的性質，必須滿足數項假設，包括：誤差項彼此獨立、具同質變異（homoscedasticity），且通常進一步假設服從常態分布（normality），以利後續推論。在進行模型殘差分析時，本文依序透過殘差分布圖與常態分布的分位數圖（Q-Q plot），以及 Anderson-Darling 檢定（A-D 檢定） 來檢驗殘差的常態性；使用 Permutation Test 以非參數方式比較前後期殘差分布是否一致；最後再以 Box-Ljung Test 檢驗殘差的自我相關性。

從殘差分析圖中可觀察到明顯的厚尾現象（heavy tail），顯示殘差偏離常態分布；同時，A-D 檢定的 p 值小於 0.01，也顯示有統計顯著的證據拒絕常態假設。此外，Box-Ljung Test 的 p 值為 0.3448，遠高於 0.05 的信心水準，表示殘差並無明顯的自相關性。

基於上述殘差不符合常態分布的結果，傳統的 F 檢定將可能導致誤導性的推論，因其對常態性具高度敏感性。因此，為避免違反 F 檢定前提所帶來的偏誤，本文改以置換檢定（Permutation Test）來比較不同時間段的殘差變異，作為檢驗異質變異的替代方案。置換檢定不依賴資料的分布型態，能在較少前提下提供可靠的推論結果，特別適合應用於殘差存在偏態或厚尾的情境下。

set.seed(42) # 固定隨機種子  
  
# 假設你用的是模型殘差  
res <- residuals(model1)  
  
# 分組  
n <- length(res)  
half <- floor(n / 2)  
group1 <- res[1:half]  
group2 <- res[(half + 1):n]  
  
# 真實標準差差異  
obs\_diff <- abs(sd(group1) - sd(group2))  
  
# 置換檢定  
n\_perm <- 10000  
perm\_diffs <- replicate(n\_perm, {  
 perm <- sample(res) # 隨機重組  
 g1 <- perm[1:half]  
 g2 <- perm[(half + 1):n]  
 abs(sd(g1) - sd(g2))  
})  
  
# p 值  
p\_value <- mean(perm\_diffs >= obs\_diff)  
cat("Permutation test p-value:", p\_value, "\n")

## Permutation test p-value: 0

基於模型殘差是否服從常態分布的疑慮，為進一步檢驗殘差變異是否隨時間改變，亦即是否存在異質變異（Heteroskedasticity）的情形，本研究採用置換檢定（Permutation Test）進行分析。置換檢定是一種非參數統計方法，適用於在不依賴特定分布假設（如常態性或等變異性）下，檢驗兩組資料在某統計量上的差異是否顯著。其核心概念在於：當虛無假設成立時，觀察值間具有可交換性（exchangeability），因此可透過隨機重排資料順序以建立統計量的虛無分布（null distribution），並比較實際觀察值的位置以計算 p 值，進而檢驗前後期殘差是否來自相同變異結構。本次分析中，最終所得的 permutation test p-value 為小於0.01，顯示殘差變異在時間上存在顯著差異，支持異質變異的存在。

總結本段可以得知，ARIMA模型對於黃金日價格估計上儘管符合多項OLS假設，然而異質變異問題存在估計問題中，不僅導致最小平方法估計（OLS）不為不偏估計量(Unbiased Estimator)，且無法根據Gauss-Markov定理計算出最佳線性不偏估計式(Best Linear Unbiased Estimater)。為解決此問題，以下嘗試使用AutoRegressive Conditional Heteroskedasticity Model(以下簡稱ARCH 模型)進行分析。

# library(forecast)  
#   
# # Fit ARIMA(1,1,1) model  
# model1 <- arima(log(AU\_OIL$GoldPrice\_interp), order = c(1,1,1))  
#   
# # Extract fitted differences (Δlog\_price)  
# z\_fit <- fitted(model1) # length = N - 1  
#   
# # Get full time series  
# log\_gold <- log(AU\_OIL$GoldPrice\_interp)  
# dates <- AU\_OIL$Date  
#   
# # Initialize fitted log-price vector  
# log\_price\_hat <- rep(NA, length(log\_gold))  
# log\_price\_hat[2:length(log\_gold)] <- z\_fit  
#   
# # Residual standard deviation  
# resid\_sd <- sd(residuals(model1), na.rm = TRUE)  
#   
# # CI quantiles  
# z95 <- qnorm(0.975)  
# z90 <- qnorm(0.95)  
#   
# # Confidence intervals on log scale  
# upper95 <- log\_price\_hat + z95 \* resid\_sd  
# lower95 <- log\_price\_hat - z95 \* resid\_sd  
# upper90 <- log\_price\_hat + z90 \* resid\_sd  
# lower90 <- log\_price\_hat - z90 \* resid\_sd  
#   
# # Plot  
# plot(dates, log\_gold, type = "l", col = "black", lwd = 1.5,  
# main = "ARIMA(1,1,1) Fitted Values with 95% and 90% CI",  
# xlab = "Date", ylab = "log(Gold Price)")  
#   
# lines(dates, log\_price\_hat, col = "blue", lwd = 2)  
# lines(dates, upper95, col = "red", lty = 2)  
# lines(dates, lower95, col = "red", lty = 2)  
# lines(dates, upper90, col = "orange", lty = 3)  
# lines(dates, lower90, col = "orange", lty = 3)  
#   
# legend("topleft", legend = c("Actual log price", "Fitted", "95% CI", "90% CI"),  
# col = c("black", "blue", "red", "orange"),  
# lty = c(1, 1, 2, 3), lwd = c(1.5, 2, 1, 1))

# p <-predict(model1, n.ahead = 12)  
# exp(p$pred)

————–STEP2 ARCH model

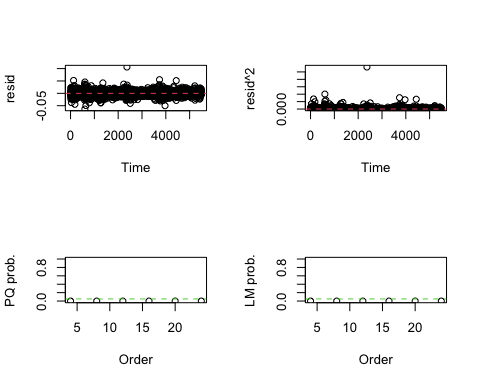
Box.test(res,type="Ljung-Box", fitdf=1)

##   
## Box-Ljung test  
##   
## data: res  
## X-squared = 0.023729, df = 0, p-value < 2.2e-16

Recalled the result of the Ljung-Box test for the residuals which has sufficient evidence showing the series is autocorrelated.

arch.test(model1, output = TRUE)

## ARCH heteroscedasticity test for residuals   
## alternative: heteroscedastic   
##   
## Portmanteau-Q test:   
## order PQ p.value  
## [1,] 4 58.2 7.09e-12  
## [2,] 8 178.1 0.00e+00  
## [3,] 12 202.4 0.00e+00  
## [4,] 16 250.1 0.00e+00  
## [5,] 20 264.2 0.00e+00  
## [6,] 24 279.1 0.00e+00  
## Lagrange-Multiplier test:   
## order LM p.value  
## [1,] 4 18784 0  
## [2,] 8 7547 0  
## [3,] 12 4942 0  
## [4,] 16 3482 0  
## [5,] 20 2762 0  
## [6,] 24 2282 0

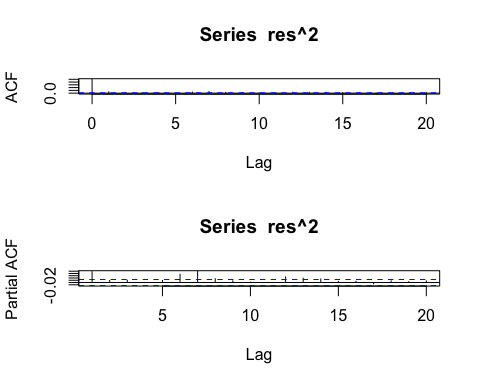
 為檢驗殘差中是否存在條件異質變異，本研究進一步對 ARIMA 模型殘差進行 ARCH LM 檢定（Lagrange Multiplier Test for ARCH effects）。從結果圖中可觀察到，無論是 Portmanteau Q test（PQ）或 Lagrange Multiplier test（LM）於各滯後階數下的 p 值皆遠低於 0.05，顯示在顯著水準下拒絕「不存在 ARCH 效應」的虛無假設，亦即殘差中存在顯著的變異聚集現象（volatility clustering）。因此，採用 ARCH 或 GARCH 類模型進行條件變異數建模是有其統計根據的。

library(FinTS)

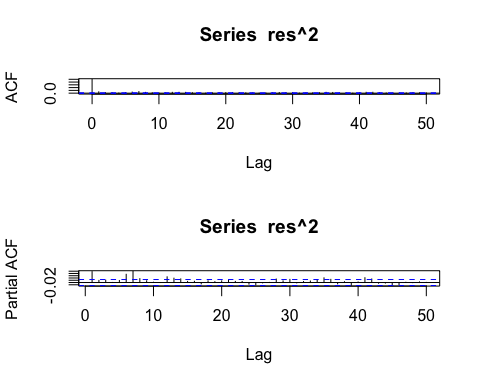
##   
## 載入套件：'FinTS'

## 下列物件被遮斷自 'package:forecast':  
##   
## Acf

par(mfrow=c(2,1))  
acf(res^2, lag.max=20)  
pacf(res^2, lag.max=20)



acf(res^2, lag.max=50)  
pacf(res^2, lag.max=50)



par(mfrow=c(1,1))  
  
ArchTest(res, lags = 5)

##   
## ARCH LM-test; Null hypothesis: no ARCH effects  
##   
## data: res  
## Chi-squared = 56.471, df = 5, p-value = 6.5e-11

ACF與PACF的圖形是用來決定ARCH模型滯後期數的判斷依據之一，良好的ARCH模型應該具有拖尾（Tails off）的ACF圖形以及截斷的（Cut-Off）的PACF圖形，PACF截斷的期數即為ARCH模型的滯後期數。從圖中可以看到ACF在第0期之後的期數，幾乎都與0沒有顯著差異；然而PACF圖形則有兩種可能：截斷在第一期以及截斷在第7期。歸納上述的圖形分析結論，可以推測殘差可能有兩個潛在模型：ARCH(1)或是ARCH(7)。以下將擬合兩個模型後，透過評估係數及相關模型指標後進行選擇。

# library(zoo)  
# roll\_sd <- rollapply(log(AU\_OIL$GoldPrice\_interp), width = 365, FUN = sd, align = "right", fill = NA)  
#   
# plot(AU\_OIL$Date, roll\_sd, type = "l",  
# main = "Rolling Standard Deviation (1 Year)",  
# ylab = "SD", xlab = "Date")  
#所有的觀測值都落在fitted value的95%信賴區間內

# 建立 ARCH(1) with ARIMA(1,1,1) 均值結構  
spec1 <- ugarchspec(  
 variance.model = list(model = "sGARCH", garchOrder = c(1, 0)), # ARCH(1)  
 mean.model = list(armaOrder = c(1, 1), include.mean = TRUE),  
 distribution.model = "norm"  
)  
  
# 建立 ARCH(2) with ARIMA(1,1,1) 均值結構  
spec2 <- ugarchspec(  
 variance.model = list(model = "sGARCH", garchOrder = c(7, 0)), # ARCH(7)  
 mean.model = list(armaOrder = c(1, 1), include.mean = TRUE),  
 distribution.model = "norm"  
)  
  
# 使用對數轉換後的價格資料  
log\_gold <- log(AU\_OIL$GoldPrice\_interp)  
  
# 擬合模型  
model1.1 <- ugarchfit(spec = spec1, data = log\_gold)  
model1.2 <- ugarchfit(spec = spec2, data = log\_gold)  
model1.1

##   
## \*---------------------------------\*  
## \* GARCH Model Fit \*  
## \*---------------------------------\*  
##   
## Conditional Variance Dynamics   
## -----------------------------------  
## GARCH Model : sGARCH(1,0)  
## Mean Model : ARFIMA(1,0,1)  
## Distribution : norm   
##   
## Optimal Parameters  
## ------------------------------------  
## Estimate Std. Error t value Pr(>|t|)  
## mu 6.860429 0.010720 639.9485 0  
## ar1 0.999961 0.000555 1802.3836 0  
## ma1 0.097892 0.000097 1005.7119 0  
## omega 0.000045 0.000001 42.6109 0  
## alpha1 0.057560 0.008386 6.8641 0  
##   
## Robust Standard Errors:  
## Estimate Std. Error t value Pr(>|t|)  
## mu 6.860429 1.3233e+03 0.005184 0.99586  
## ar1 0.999961 7.1316e+01 0.014022 0.98881  
## ma1 0.097892 1.2930e+01 0.007571 0.99396  
## omega 0.000045 7.3290e-03 0.006179 0.99507  
## alpha1 0.057560 4.7773e+01 0.001205 0.99904  
##   
## LogLikelihood : 19318.7   
##   
## Information Criteria  
## ------------------------------------  
##   
## Akaike -7.0360  
## Bayes -7.0299  
## Shibata -7.0360  
## Hannan-Quinn -7.0339  
##   
## Weighted Ljung-Box Test on Standardized Residuals  
## ------------------------------------  
## statistic p-value  
## Lag[1] 1.728 1.887e-01  
## Lag[2\*(p+q)+(p+q)-1][5] 6.576 6.940e-06  
## Lag[4\*(p+q)+(p+q)-1][9] 10.054 8.741e-03  
## d.o.f=2  
## H0 : No serial correlation  
##   
## Weighted Ljung-Box Test on Standardized Squared Residuals  
## ------------------------------------  
## statistic p-value  
## Lag[1] 0.003553 0.9525  
## Lag[2\*(p+q)+(p+q)-1][2] 0.003837 0.9956  
## Lag[4\*(p+q)+(p+q)-1][5] 0.077507 0.9989  
## d.o.f=1  
##   
## Weighted ARCH LM Tests  
## ------------------------------------  
## Statistic Shape Scale P-Value  
## ARCH Lag[2] 0.000568 0.500 2.000 0.9810  
## ARCH Lag[4] 0.092311 1.397 1.611 0.9856  
## ARCH Lag[6] 0.149427 2.222 1.500 0.9978  
##   
## Nyblom stability test  
## ------------------------------------  
## Joint Statistic: 9.4696  
## Individual Statistics:   
## mu 1.0000  
## ar1 0.9999  
## ma1 0.2469  
## omega 1.8392  
## alpha1 6.6346  
##   
## Asymptotic Critical Values (10% 5% 1%)  
## Joint Statistic: 1.28 1.47 1.88  
## Individual Statistic: 0.35 0.47 0.75  
##   
## Sign Bias Test  
## ------------------------------------  
## t-value prob sig  
## Sign Bias 0.962 3.361e-01   
## Negative Sign Bias 3.784 1.560e-04 \*\*\*  
## Positive Sign Bias 2.336 1.952e-02 \*\*  
## Joint Effect 21.695 7.548e-05 \*\*\*  
##   
##   
## Adjusted Pearson Goodness-of-Fit Test:  
## ------------------------------------  
## group statistic p-value(g-1)  
## 1 20 843.1 1.489e-166  
## 2 30 905.9 6.298e-172  
## 3 40 935.5 6.784e-171  
## 4 50 982.8 5.705e-174  
##   
##   
## Elapsed time : 0.8839941

model1.2

##   
## \*---------------------------------\*  
## \* GARCH Model Fit \*  
## \*---------------------------------\*  
##   
## Conditional Variance Dynamics   
## -----------------------------------  
## GARCH Model : sGARCH(7,0)  
## Mean Model : ARFIMA(1,0,1)  
## Distribution : norm   
##   
## Optimal Parameters  
## ------------------------------------  
## Estimate Std. Error t value Pr(>|t|)  
## mu 6.525533 0.007248 9.0030e+02 0.000000  
## ar1 1.000000 0.000201 4.9699e+03 0.000000  
## ma1 0.165986 0.014909 1.1134e+01 0.000000  
## omega 0.000018 0.000001 1.9144e+01 0.000000  
## alpha1 0.243791 0.035500 6.8673e+00 0.000000  
## alpha2 0.012882 0.006386 2.0171e+00 0.043682  
## alpha3 0.000000 0.010276 2.8000e-05 0.999978  
## alpha4 0.000000 0.035824 5.0000e-06 0.999996  
## alpha5 0.016922 0.006962 2.4305e+00 0.015079  
## alpha6 0.179089 0.029815 6.0066e+00 0.000000  
## alpha7 0.355204 0.031522 1.1268e+01 0.000000  
##   
## Robust Standard Errors:  
## Estimate Std. Error t value Pr(>|t|)  
## mu 6.525533 0.005294 1.2326e+03 0.000000  
## ar1 1.000000 0.000417 2.3998e+03 0.000000  
## ma1 0.165986 0.071870 2.3095e+00 0.020914  
## omega 0.000018 0.000004 4.6864e+00 0.000003  
## alpha1 0.243791 0.286725 8.5026e-01 0.395182  
## alpha2 0.012882 0.011126 1.1578e+00 0.246950  
## alpha3 0.000000 0.108039 3.0000e-06 0.999998  
## alpha4 0.000000 0.401670 0.0000e+00 1.000000  
## alpha5 0.016922 0.023252 7.2775e-01 0.466764  
## alpha6 0.179089 0.232793 7.6931e-01 0.441711  
## alpha7 0.355204 0.149947 2.3689e+00 0.017843  
##   
## LogLikelihood : 19849.64   
##   
## Information Criteria  
## ------------------------------------  
##   
## Akaike -7.2272  
## Bayes -7.2139  
## Shibata -7.2272  
## Hannan-Quinn -7.2226  
##   
## Weighted Ljung-Box Test on Standardized Residuals  
## ------------------------------------  
## statistic p-value  
## Lag[1] 3.621 5.704e-02  
## Lag[2\*(p+q)+(p+q)-1][5] 13.318 0.000e+00  
## Lag[4\*(p+q)+(p+q)-1][9] 15.629 1.794e-05  
## d.o.f=2  
## H0 : No serial correlation  
##   
## Weighted Ljung-Box Test on Standardized Squared Residuals  
## ------------------------------------  
## statistic p-value  
## Lag[1] 1.755 1.852e-01  
## Lag[2\*(p+q)+(p+q)-1][20] 35.556 7.138e-06  
## Lag[4\*(p+q)+(p+q)-1][34] 68.998 5.429e-11  
## d.o.f=7  
##   
## Weighted ARCH LM Tests  
## ------------------------------------  
## Statistic Shape Scale P-Value  
## ARCH Lag[8] 1.671 0.500 2.000 0.19611  
## ARCH Lag[10] 6.470 1.488 1.815 0.06665  
## ARCH Lag[12] 10.468 2.451 1.700 0.02883  
##   
## Nyblom stability test  
## ------------------------------------  
## Joint Statistic: 8.2622  
## Individual Statistics:   
## mu 0.0006515  
## ar1 0.4418652  
## ma1 0.1218307  
## omega 0.2216975  
## alpha1 0.7544903  
## alpha2 0.1142603  
## alpha3 0.5043903  
## alpha4 4.3426062  
## alpha5 0.1049096  
## alpha6 0.2166597  
## alpha7 0.1517165  
##   
## Asymptotic Critical Values (10% 5% 1%)  
## Joint Statistic: 2.49 2.75 3.27  
## Individual Statistic: 0.35 0.47 0.75  
##   
## Sign Bias Test  
## ------------------------------------  
## t-value prob sig  
## Sign Bias 0.4837 0.6286   
## Negative Sign Bias 0.3956 0.6924   
## Positive Sign Bias 1.0306 0.3028   
## Joint Effect 1.2643 0.7376   
##   
##   
## Adjusted Pearson Goodness-of-Fit Test:  
## ------------------------------------  
## group statistic p-value(g-1)  
## 1 20 645.7 1.112e-124  
## 2 30 678.4 3.208e-124  
## 3 40 730.4 2.360e-128  
## 4 50 737.6 1.184e-123  
##   
##   
## Elapsed time : 1.61722

為評估不同階數之 ARCH 模型在解釋黃金價格變異性方面的表現，本文分別建構以 ARIMA(1,1,1) 為均值結構的 ARCH(1) 與 ARCH(7) 模型進行比較。從樣本內擬合指標觀察，ARCH(7) 模型的對數似然值（LogLikelihood = 19849.64）高於 ARCH(1) 模型（19318.7），且其 AIC 值（-7.2272）亦優於 ARCH(1) 的 -7.0360，顯示在樣本內擬合層面 ARCH(7) 模型較佳。

然而，深入觀察模型診斷結果可發現，ARCH(7) 存在多項統計上的潛在問題。其殘差平方的 Ljung-Box 檢定在高階（Lag 20 與 Lag 34）下出現顯著結果（p < 0.01），顯示仍存在殘差結構未被捕捉；此外，ARCH LM 檢定在 Lag 12 時亦呈現邊界顯著性（p = 0.02883），說明變異數模型可能未完全描述波動性。穩定性分析方面，ARCH(7) 模型的 alpha4 參數對應 Nyblom 統計量為 4.34，遠超過 1% 臨界值（0.75），顯示參數存在嚴重不穩定問題，且多數 alpha 參數在 robust 標準誤下變得不顯著，增加模型解釋上的不確定性。

相比之下，ARCH(1) 模型雖然 AIC 略高，但模型結構簡潔，所有參數顯著，殘差與殘差平方均符合白噪音假設，且未觀察到殘留的 ARCH 效應。其 Nyblom 檢定中，唯一顯著不穩定的是 alpha1，但幅度相對 ARCH(7) 為低，整體結構較為穩定。

綜上所述，ARCH(7) 雖在擬合表現上佔優，但模型不穩定且解釋力不一致，存在過度擬合風險；而 ARCH(1) 模型則在統計顯著性、殘差結構與穩定性方面表現更為穩健。因此，本文建議採用 ARCH(1) 作為黃金價格變異數建模的主要架構，並可於後續進一步考慮 ARCH(2) 或 ARCH(3) 作為潛在折衷方案。

LL\_full <- likelihood(model1.2) # ARCH(7)  
LL\_restricted <- likelihood(model1.1) # ARCH(1)  
  
LR\_stat <- 2 \* (LL\_full - LL\_restricted)  
df <- 6  
p\_value <- pchisq(LR\_stat, df = df, lower.tail = FALSE)  
  
cat("Likelihood Ratio Statistic:", LR\_stat, "\n")

## Likelihood Ratio Statistic: 1061.871

cat("p-value:", p\_value, "\n")

## p-value: 3.701049e-226

為比較 ARCH(1) 與 ARCH(7) 模型在條件變異數建模上的解釋力，本文進行似然比檢定。結果顯示，ARCH(7) 模型相較於 ARCH(1) 的對數似然值顯著提高（LR 統計量 = 1061.871，df = 6，p 值 <2.2e-16），表示多加入的延遲變異數項對模型整體擬合具有統計上的顯著貢獻。然而，雖然 ARCH(7) 在樣本內表現優越，其模型參數顯著性、穩定性與殘差結構檢定結果顯示存在過度擬合與不穩定性問題，多數 項於 robust 標準誤下不顯著，Nyblom個別檢定亦出現超過臨界值情形。因此，在統計顯著與模型穩健性間取得平衡下，本文仍建議採用較為簡潔且結構穩定的 ARCH(1) 模型，並視情況進行中階（例如 ARCH(2)–ARCH(3)）模型之敏感性分析以尋求最佳折衷。

library(e1071) # for skewness and kurtosis  
  
# 取標準化殘差  
resid\_std <- residuals(model1.1, standardize = TRUE)  
  
# 計算指標  
sk <- skewness(resid\_std)  
ku <- kurtosis(resid\_std)  
  
cat("Skewness:", sk, "\n")

## Skewness: -6.554653

cat("Kurtosis:", ku, "\n")

## Kurtosis: 237.8278

# 取標準化殘差  
resid\_std <- residuals(model1.2, standardize = TRUE)  
  
# 計算指標  
sk <- skewness(resid\_std)  
ku <- kurtosis(resid\_std)  
  
cat("Skewness:", sk, "\n")

## Skewness: -0.2203127

cat("Kurtosis:", ku, "\n")

## Kurtosis: 6.646643

模型的 標準化殘差 不僅不是常態分布，還呈現極度偏態與厚尾，這會嚴重低估風險與尾端事件的機率。

# ARCH(1) with skewed-t  
spec\_sstd <- ugarchspec(  
 variance.model = list(model = "sGARCH", garchOrder = c(1, 0)),  
 mean.model = list(armaOrder = c(1, 1), include.mean = TRUE),  
 distribution.model = "sstd"  
)  
  
  
# ARCH(7) with t-distribution  
spec\_std <- ugarchspec(  
 variance.model = list(model = "sGARCH", garchOrder = c(7, 0)),  
 mean.model = list(armaOrder = c(1, 1), include.mean = TRUE),  
 distribution.model = "std"  
)  
  
  
  
  
model1.1.sstd <- ugarchfit(spec = spec\_sstd, data = log\_gold)  
model1.2.std <- ugarchfit(spec = spec\_std, data = log\_gold)  
model1.1.sstd

##   
## \*---------------------------------\*  
## \* GARCH Model Fit \*  
## \*---------------------------------\*  
##   
## Conditional Variance Dynamics   
## -----------------------------------  
## GARCH Model : sGARCH(1,0)  
## Mean Model : ARFIMA(1,0,1)  
## Distribution : sstd   
##   
## Optimal Parameters  
## ------------------------------------  
## Estimate Std. Error t value Pr(>|t|)  
## mu 6.531953 0.014191 460.2958 0  
## ar1 1.000000 0.000171 5846.9024 0  
## ma1 0.141898 0.012380 11.4615 0  
## omega 0.000045 0.000005 8.9861 0  
## alpha1 0.998990 0.141163 7.0768 0  
## skew 0.976734 0.012969 75.3158 0  
## shape 2.556399 0.088840 28.7752 0  
##   
## Robust Standard Errors:  
## Estimate Std. Error t value Pr(>|t|)  
## mu 6.531953 0.002104 3104.7875 0  
## ar1 1.000000 0.000217 4616.8002 0  
## ma1 0.141898 0.010760 13.1873 0  
## omega 0.000045 0.000005 9.6728 0  
## alpha1 0.998990 0.126004 7.9282 0  
## skew 0.976734 0.015164 64.4115 0  
## shape 2.556399 0.074564 34.2848 0  
##   
## LogLikelihood : 20301.64   
##   
## Information Criteria  
## ------------------------------------  
##   
## Akaike -7.3933  
## Bayes -7.3849  
## Shibata -7.3933  
## Hannan-Quinn -7.3904  
##   
## Weighted Ljung-Box Test on Standardized Residuals  
## ------------------------------------  
## statistic p-value  
## Lag[1] 2.751 9.720e-02  
## Lag[2\*(p+q)+(p+q)-1][5] 10.853 2.998e-15  
## Lag[4\*(p+q)+(p+q)-1][9] 13.082 3.513e-04  
## d.o.f=2  
## H0 : No serial correlation  
##   
## Weighted Ljung-Box Test on Standardized Squared Residuals  
## ------------------------------------  
## statistic p-value  
## Lag[1] 6.390 0.01148  
## Lag[2\*(p+q)+(p+q)-1][2] 6.610 0.01513  
## Lag[4\*(p+q)+(p+q)-1][5] 8.861 0.01795  
## d.o.f=1  
##   
## Weighted ARCH LM Tests  
## ------------------------------------  
## Statistic Shape Scale P-Value  
## ARCH Lag[2] 0.4403 0.500 2.000 0.506976  
## ARCH Lag[4] 2.8832 1.397 1.611 0.279162  
## ARCH Lag[6] 13.4937 2.222 1.500 0.001855  
##   
## Nyblom stability test  
## ------------------------------------  
## Joint Statistic: 5.2957  
## Individual Statistics:   
## mu 0.0001138  
## ar1 0.3388701  
## ma1 0.0959737  
## omega 1.6359461  
## alpha1 2.1127347  
## skew 0.6173650  
## shape 1.8134559  
##   
## Asymptotic Critical Values (10% 5% 1%)  
## Joint Statistic: 1.69 1.9 2.35  
## Individual Statistic: 0.35 0.47 0.75  
##   
## Sign Bias Test  
## ------------------------------------  
## t-value prob sig  
## Sign Bias 0.364 0.715851   
## Negative Sign Bias 1.827 0.067778 \*  
## Positive Sign Bias 2.726 0.006439 \*\*\*  
## Joint Effect 11.653 0.008673 \*\*\*  
##   
##   
## Adjusted Pearson Goodness-of-Fit Test:  
## ------------------------------------  
## group statistic p-value(g-1)  
## 1 20 61.64 2.121e-06  
## 2 30 61.64 3.845e-04  
## 3 40 80.03 1.189e-04  
## 4 50 89.67 3.504e-04  
##   
##   
## Elapsed time : 1.711666

model1.2.std

##   
## \*---------------------------------\*  
## \* GARCH Model Fit \*  
## \*---------------------------------\*  
##   
## Conditional Variance Dynamics   
## -----------------------------------  
## GARCH Model : sGARCH(7,0)  
## Mean Model : ARFIMA(1,0,1)  
## Distribution : std   
##   
## Optimal Parameters  
## ------------------------------------  
## Estimate Std. Error t value Pr(>|t|)  
## mu 6.532090 0.015083 4.3307e+02 0.00000  
## ar1 1.000000 0.000165 6.0564e+03 0.00000  
## ma1 0.170524 0.011190 1.5239e+01 0.00000  
## omega 0.000014 0.000001 1.0957e+01 0.00000  
## alpha1 0.441645 0.068745 6.4244e+00 0.00000  
## alpha2 0.012865 0.011992 1.0728e+00 0.28336  
## alpha3 0.000000 0.009962 2.3000e-05 0.99998  
## alpha4 0.000000 0.008553 3.3000e-05 0.99997  
## alpha5 0.005319 0.007911 6.7234e-01 0.50137  
## alpha6 0.160335 0.028116 5.7026e+00 0.00000  
## alpha7 0.378832 0.043225 8.7643e+00 0.00000  
## shape 3.297236 0.102222 3.2256e+01 0.00000  
##   
## Robust Standard Errors:  
## Estimate Std. Error t value Pr(>|t|)  
## mu 6.532090 0.002278 2.8679e+03 0.000000  
## ar1 1.000000 0.000276 3.6198e+03 0.000000  
## ma1 0.170524 0.010264 1.6613e+01 0.000000  
## omega 0.000014 0.000004 3.0391e+00 0.002373  
## alpha1 0.441645 0.197314 2.2383e+00 0.025203  
## alpha2 0.012865 0.033958 3.7884e-01 0.704807  
## alpha3 0.000000 0.026681 8.0000e-06 0.999993  
## alpha4 0.000000 0.024748 1.2000e-05 0.999991  
## alpha5 0.005319 0.020993 2.5336e-01 0.799987  
## alpha6 0.160335 0.059319 2.7029e+00 0.006873  
## alpha7 0.378832 0.088616 4.2750e+00 0.000019  
## shape 3.297236 0.094649 3.4837e+01 0.000000  
##   
## LogLikelihood : 20496.19   
##   
## Information Criteria  
## ------------------------------------  
##   
## Akaike -7.4624  
## Bayes -7.4479  
## Shibata -7.4624  
## Hannan-Quinn -7.4573  
##   
## Weighted Ljung-Box Test on Standardized Residuals  
## ------------------------------------  
## statistic p-value  
## Lag[1] 1.406 2.358e-01  
## Lag[2\*(p+q)+(p+q)-1][5] 12.029 0.000e+00  
## Lag[4\*(p+q)+(p+q)-1][9] 14.549 6.488e-05  
## d.o.f=2  
## H0 : No serial correlation  
##   
## Weighted Ljung-Box Test on Standardized Squared Residuals  
## ------------------------------------  
## statistic p-value  
## Lag[1] 5.867 1.543e-02  
## Lag[2\*(p+q)+(p+q)-1][20] 33.221 2.557e-05  
## Lag[4\*(p+q)+(p+q)-1][34] 57.785 2.751e-08  
## d.o.f=7  
##   
## Weighted ARCH LM Tests  
## ------------------------------------  
## Statistic Shape Scale P-Value  
## ARCH Lag[8] 0.420 0.500 2.000 0.5169  
## ARCH Lag[10] 3.287 1.488 1.815 0.3015  
## ARCH Lag[12] 6.337 2.451 1.700 0.1807  
##   
## Nyblom stability test  
## ------------------------------------  
## Joint Statistic: 68.6197  
## Individual Statistics:   
## mu 0.00063  
## ar1 1.24743  
## ma1 0.13534  
## omega 16.18311  
## alpha1 4.57779  
## alpha2 0.38396  
## alpha3 9.47955  
## alpha4 4.30646  
## alpha5 0.20720  
## alpha6 1.67443  
## alpha7 3.36859  
## shape 1.02076  
##   
## Asymptotic Critical Values (10% 5% 1%)  
## Joint Statistic: 2.69 2.96 3.51  
## Individual Statistic: 0.35 0.47 0.75  
##   
## Sign Bias Test  
## ------------------------------------  
## t-value prob sig  
## Sign Bias 0.4506 0.652318   
## Negative Sign Bias 1.6655 0.095870 \*  
## Positive Sign Bias 2.7831 0.005402 \*\*\*  
## Joint Effect 10.5197 0.014628 \*\*  
##   
##   
## Adjusted Pearson Goodness-of-Fit Test:  
## ------------------------------------  
## group statistic p-value(g-1)  
## 1 20 42.94 1.321e-03  
## 2 30 67.79 6.041e-05  
## 3 40 79.50 1.375e-04  
## 4 50 84.97 1.093e-03  
##   
##   
## Elapsed time : 2.530394

ARCH(7)裡面的參數在更換分布後仍舊有不同的部分

兩個模型解釋力有顯著不同

library(e1071) # for skewness and kurtosis  
  
# 取標準化殘差  
resid\_std <- residuals(model1.1.sstd, standardize = TRUE)  
  
# 計算指標  
sk <- skewness(resid\_std)  
ku <- kurtosis(resid\_std)  
  
cat("Skewness:", sk, "\n")

## Skewness: 0.1415584

cat("Kurtosis:", ku, "\n")

## Kurtosis: 8.583846

# 取標準化殘差  
resid\_std <- residuals(model1.2.std, standardize = TRUE)  
  
# 計算指標  
sk <- skewness(resid\_std)  
ku <- kurtosis(resid\_std)  
  
cat("Skewness:", sk, "\n")

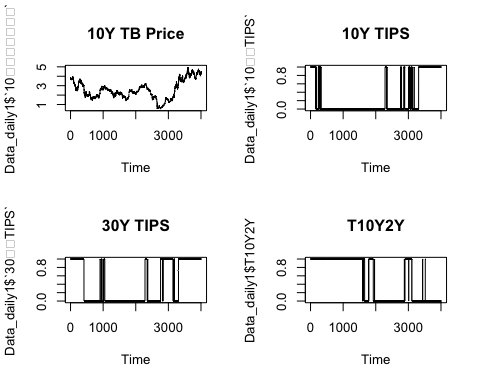
## Skewness: -0.3464603

cat("Kurtosis:", ku, "\n")

## Kurtosis: 8.205279

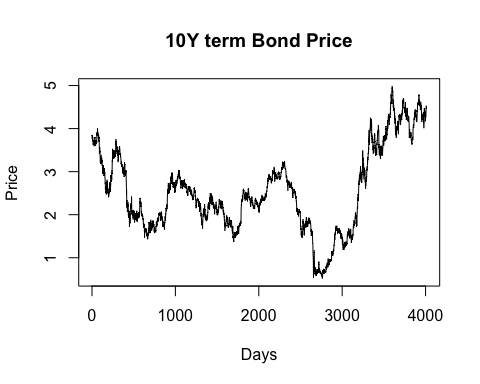
在本次分析中，我們針對四種不同的模型，評估它們的標準化殘差與理論分布的契合程度。首先，採用常態分布誤差的 Model 1.1（ARIMA(1,1,1) 搭配 ARCH(1)）所產生的標準化殘差表現極差，呈現出強烈的負偏態（skewness 為 -6.55）與極高的峰度（kurtosis 達 237.83），顯示誤差分布高度偏離對稱且具有非常厚的尾部，遠遠超出常態分布的假設，這使得模型在預測或風險評估上可能非常不可靠。 相較之下，Model 1.2（ARIMA(1,1,1) 搭配 ARCH(7)）同樣假設常態分布，但在殘差表現上略有改善，偏態降至 -0.22，峰度亦降至 6.65，雖仍偏離理論常態分布，卻已不再如前者般極端。這顯示增加 ARCH 項次可在一定程度上捕捉變異結構，但仍無法解決常態分布誤差對金融數列厚尾的低適應性問題。 進一步地，當我們將誤差分布改為 skewed-t 分布時，情況有明顯改善。Model 1.1.sstt（ARIMA(1,1,1) + ARCH(1) + skewed-t error）所產生的殘差具有良好的對稱性（skewness 約 0.14）與可接受的厚尾（kurtosis 為 8.58），明顯顯示 skewed-t 分布能有效捕捉金融資料中的非對稱與重尾特性。該模型的誤差分布與資料行為較為吻合，使模型在預測與風險估計上更具可靠性。 最後，在 Model 1.2.std（ARIMA(1,1,1) + ARCH(7) + t 分布誤差）中，誤差雖仍略帶負偏態（skewness -0.35），但峰度為 8.21，整體上仍可視為合理且穩定的模型。t 分布具備對極端值與波動的容納能力，因此在高波動金融資產建模中，具有實用價值。 綜合而言，純常態分布對誤差的假設無法充分描述金融時間數列的實際行為，尤其在厚尾與偏態明顯的情況下。相較之下，skewed-t 與 t 分布能更好地捕捉資料特性，其中又以 Model 1.1.sstt 的誤差行為最為貼近理論分布，是目前四者中最具代表性的模型選擇。

—Section2 Time Series Analysis for the Yield rate of the Treasury Bond

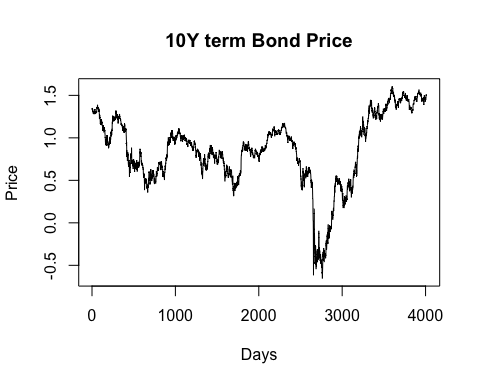


繪圖結果類似

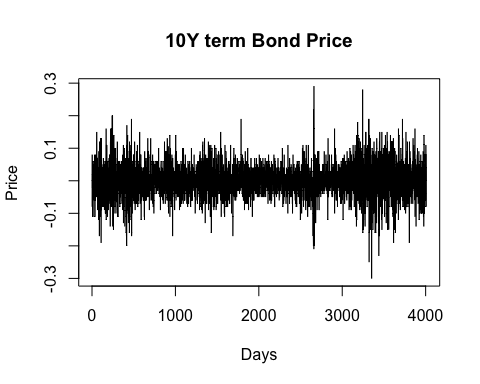
做線性插補



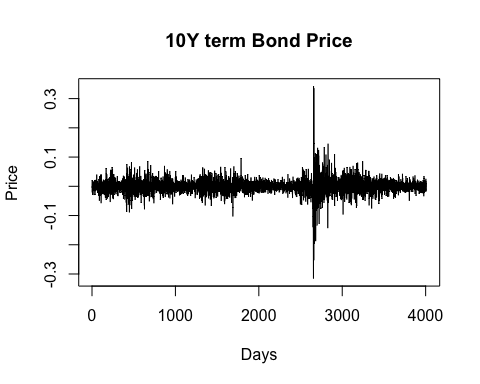
## Augmented Dickey-Fuller Test   
## alternative: stationary   
##   
## Type 1: no drift no trend   
## lag ADF p.value  
## [1,] 0 -0.330 0.549  
## [2,] 1 -0.301 0.557  
## [3,] 2 -0.299 0.558  
## [4,] 3 -0.288 0.561  
## [5,] 4 -0.287 0.561  
## [6,] 5 -0.299 0.558  
## [7,] 6 -0.250 0.572  
## [8,] 7 -0.254 0.571  
## [9,] 8 -0.233 0.577  
## [10,] 9 -0.214 0.582  
## Type 2: with drift no trend   
## lag ADF p.value  
## [1,] 0 -1.41 0.557  
## [2,] 1 -1.39 0.565  
## [3,] 2 -1.34 0.581  
## [4,] 3 -1.32 0.589  
## [5,] 4 -1.33 0.586  
## [6,] 5 -1.34 0.581  
## [7,] 6 -1.29 0.598  
## [8,] 7 -1.27 0.608  
## [9,] 8 -1.25 0.615  
## [10,] 9 -1.24 0.618  
## Type 3: with drift and trend   
## lag ADF p.value  
## [1,] 0 -1.93 0.607  
## [2,] 1 -1.89 0.622  
## [3,] 2 -1.88 0.630  
## [4,] 3 -1.86 0.637  
## [5,] 4 -1.86 0.636  
## [6,] 5 -1.88 0.629  
## [7,] 6 -1.81 0.656  
## [8,] 7 -1.81 0.658  
## [9,] 8 -1.78 0.670  
## [10,] 9 -1.76 0.678  
## ----   
## Note: in fact, p.value = 0.01 means p.value <= 0.01



## Augmented Dickey-Fuller Test   
## alternative: stationary   
##   
## Type 1: no drift no trend   
## lag ADF p.value  
## [1,] 0 -0.752 0.410  
## [2,] 1 -0.738 0.415  
## [3,] 2 -0.698 0.429  
## [4,] 3 -0.643 0.449  
## [5,] 4 -0.598 0.465  
## [6,] 5 -0.674 0.438  
## [7,] 6 -0.646 0.448  
## [8,] 7 -0.580 0.471  
## [9,] 8 -0.527 0.490  
## [10,] 9 -0.547 0.483  
## Type 2: with drift no trend   
## lag ADF p.value  
## [1,] 0 -1.85 0.390  
## [2,] 1 -1.84 0.392  
## [3,] 2 -1.74 0.434  
## [4,] 3 -1.63 0.477  
## [5,] 4 -1.54 0.510  
## [6,] 5 -1.69 0.452  
## [7,] 6 -1.66 0.464  
## [8,] 7 -1.51 0.521  
## [9,] 8 -1.42 0.554  
## [10,] 9 -1.47 0.534  
## Type 3: with drift and trend   
## lag ADF p.value  
## [1,] 0 -2.00 0.576  
## [2,] 1 -1.99 0.580  
## [3,] 2 -1.90 0.620  
## [4,] 3 -1.80 0.663  
## [5,] 4 -1.72 0.697  
## [6,] 5 -1.86 0.637  
## [7,] 6 -1.82 0.651  
## [8,] 7 -1.69 0.709  
## [9,] 8 -1.60 0.746  
## [10,] 9 -1.65 0.725  
## ----   
## Note: in fact, p.value = 0.01 means p.value <= 0.01



## Augmented Dickey-Fuller Test   
## alternative: stationary   
##   
## Type 1: no drift no trend   
## lag ADF p.value  
## [1,] 0 -63.0 0.01  
## [2,] 1 -46.4 0.01  
## [3,] 2 -37.8 0.01  
## [4,] 3 -32.2 0.01  
## [5,] 4 -28.6 0.01  
## [6,] 5 -26.1 0.01  
## [7,] 6 -24.8 0.01  
## [8,] 7 -23.1 0.01  
## [9,] 8 -21.6 0.01  
## [10,] 9 -19.8 0.01  
## Type 2: with drift no trend   
## lag ADF p.value  
## [1,] 0 -63.0 0.01  
## [2,] 1 -46.4 0.01  
## [3,] 2 -37.8 0.01  
## [4,] 3 -32.2 0.01  
## [5,] 4 -28.6 0.01  
## [6,] 5 -26.1 0.01  
## [7,] 6 -24.8 0.01  
## [8,] 7 -23.1 0.01  
## [9,] 8 -21.6 0.01  
## [10,] 9 -19.8 0.01  
## Type 3: with drift and trend   
## lag ADF p.value  
## [1,] 0 -63.0 0.01  
## [2,] 1 -46.5 0.01  
## [3,] 2 -37.9 0.01  
## [4,] 3 -32.3 0.01  
## [5,] 4 -28.7 0.01  
## [6,] 5 -26.2 0.01  
## [7,] 6 -24.9 0.01  
## [8,] 7 -23.2 0.01  
## [9,] 8 -21.6 0.01  
## [10,] 9 -19.9 0.01  
## ----   
## Note: in fact, p.value = 0.01 means p.value <= 0.01

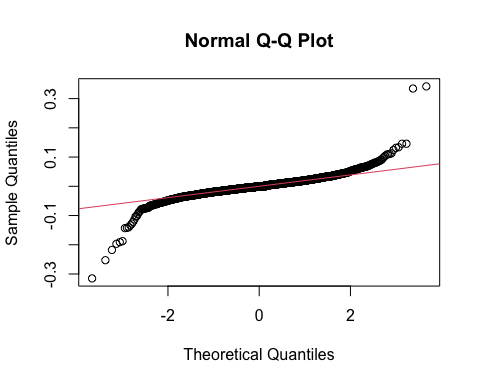


## Augmented Dickey-Fuller Test   
## alternative: stationary   
##   
## Type 1: no drift no trend   
## lag ADF p.value  
## [1,] 0 -63.1 0.01  
## [2,] 1 -47.3 0.01  
## [3,] 2 -40.1 0.01  
## [4,] 3 -35.4 0.01  
## [5,] 4 -28.6 0.01  
## [6,] 5 -26.2 0.01  
## [7,] 6 -26.4 0.01  
## [8,] 7 -25.6 0.01  
## [9,] 8 -22.8 0.01  
## [10,] 9 -20.6 0.01  
## Type 2: with drift no trend   
## lag ADF p.value  
## [1,] 0 -63.1 0.01  
## [2,] 1 -47.3 0.01  
## [3,] 2 -40.1 0.01  
## [4,] 3 -35.4 0.01  
## [5,] 4 -28.6 0.01  
## [6,] 5 -26.2 0.01  
## [7,] 6 -26.4 0.01  
## [8,] 7 -25.6 0.01  
## [9,] 8 -22.8 0.01  
## [10,] 9 -20.6 0.01  
## Type 3: with drift and trend   
## lag ADF p.value  
## [1,] 0 -63.1 0.01  
## [2,] 1 -47.4 0.01  
## [3,] 2 -40.1 0.01  
## [4,] 3 -35.5 0.01  
## [5,] 4 -28.6 0.01  
## [6,] 5 -26.2 0.01  
## [7,] 6 -26.4 0.01  
## [8,] 7 -25.6 0.01  
## [9,] 8 -22.9 0.01  
## [10,] 9 -20.7 0.01  
## ----   
## Note: in fact, p.value = 0.01 means p.value <= 0.01

## Warning in ks.test.default(diff((Bond\_interp)), "pnorm"): Kolmogorov - Smirnov  
## 檢驗裡不應該有連結

##   
## Asymptotic one-sample Kolmogorov-Smirnov test  
##   
## data: diff((Bond\_interp))  
## D = 0.43924, p-value < 2.2e-16  
## alternative hypothesis: two-sided

## [1] 1.800422



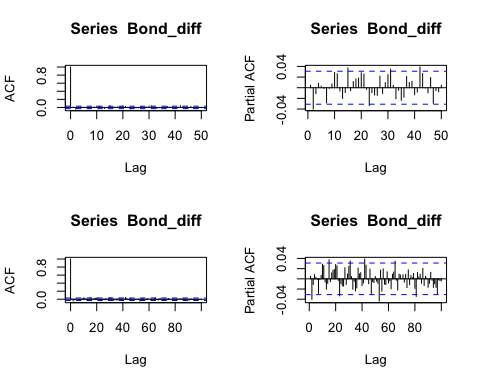
## Warning in ks.test.default(diff(log(Bond\_interp)), "pnorm"): Kolmogorov -  
## Smirnov 檢驗裡不應該有連結

##   
## Asymptotic one-sample Kolmogorov-Smirnov test  
##   
## data: diff(log(Bond\_interp))  
## D = 0.46525, p-value < 2.2e-16  
## alternative hypothesis: two-sided

## [1] 25.01949

結論是用差分但不要log的會比較好

Bond\_diff<- diff(Bond\_interp)  
##EFA  
par(mfrow=c(2,2))  
acf(Bond\_diff, lag.max=50)  
pacf(Bond\_diff, lag.max=50)  
acf(Bond\_diff, lag.max=100)  
pacf(Bond\_diff, lag.max=100)



par(mfrow=c(1,1))

ARMA(1,1) model is good for the differenced Bond Price, or ARIMA(1,1,1) model for the non-differenced Bond Price Data.

model0<- auto.arima(Bond\_diff)  
model0

## Series: Bond\_diff   
## ARIMA(0,0,0) with zero mean   
##   
## sigma^2 = 0.002727: log likelihood = 6145.44  
## AIC=-12288.88 AICc=-12288.88 BIC=-12282.58

model1 <- arima(Bond\_diff, order=c(1,0,1))  
model1

##   
## Call:  
## arima(x = Bond\_diff, order = c(1, 0, 1))  
##   
## Coefficients:  
## ar1 ma1 intercept  
## -0.4194 0.4351 2e-04  
## s.e. 0.3995 0.4005 8e-04  
##   
## sigma^2 estimated as 0.002726: log likelihood = 6146.06, aic = -12284.13

summary(model1)

##   
## Call:  
## arima(x = Bond\_diff, order = c(1, 0, 1))  
##   
## Coefficients:  
## ar1 ma1 intercept  
## -0.4194 0.4351 2e-04  
## s.e. 0.3995 0.4005 8e-04  
##   
## sigma^2 estimated as 0.002726: log likelihood = 6146.06, aic = -12284.13  
##   
## Training set error measures:  
## ME RMSE MAE MPE MAPE MASE ACF1  
## Training set -8.141533e-06 0.0522149 0.03945228 NaN Inf 0.7235024 -0.01043022

—Section 3 the Vector Autoregressive model