金融危機預測- HW1

魏上傑

2023-02-24

目錄

1 加權指數的年增率

1.1 讀取並清理資料

```
library(readr)

#read csv file
stock_raw <- read_csv("stock_raw.csv",
                      col_names = c("year","original","growth"))
```

```
## Rows: 288 Columns: 3
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (2): year, growth
```

```
## dbl (1): original
##
## 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.
```

```
#remove missing value
stock_raw <- stock_raw[stock_raw$growth!="...",]

#transform the data type
stock_raw$original <- as.numeric(stock_raw$original)
stock_raw$growth <- as.numeric(stock_raw$growth)

#transform to date
 # transform to string first
stock_raw$year <- as.character(stock_raw$year)

stock_raw$year <- as.Date(paste0(substr(stock_raw$year, 1, 4),
                                 "-",
                                 substr(stock_raw$year, 6, 7), "-01"),
                          format = "%Y-%m-%d") #default day=01
```

```
str(stock_raw)
```

```
## tibble [281 x 3] (S3: tbl_df/tbl/data.frame)
##  $ year    : Date[1:281], format: "1999-08-01" "1999-09-01" ...
##  $ original: num [1:281] 8158 7599 7855 7721 8449 ...
##  $ growth  : num [1:281] 24.54 11.19 9.61 7.57 31.63 ...
```

```
head(stock_raw)
```

```
## # A tibble: 6 x 3
##   year       original growth
##   <date>        <dbl>  <dbl>
## 1 1999-08-01    8158.  24.5
## 2 1999-09-01    7599.  11.2
## 3 1999-10-01    7855.   9.61
## 4 1999-11-01    7721.   7.57
## 5 1999-12-01    8449.  31.6
## 6 2000-01-01    9745.  62.5
```

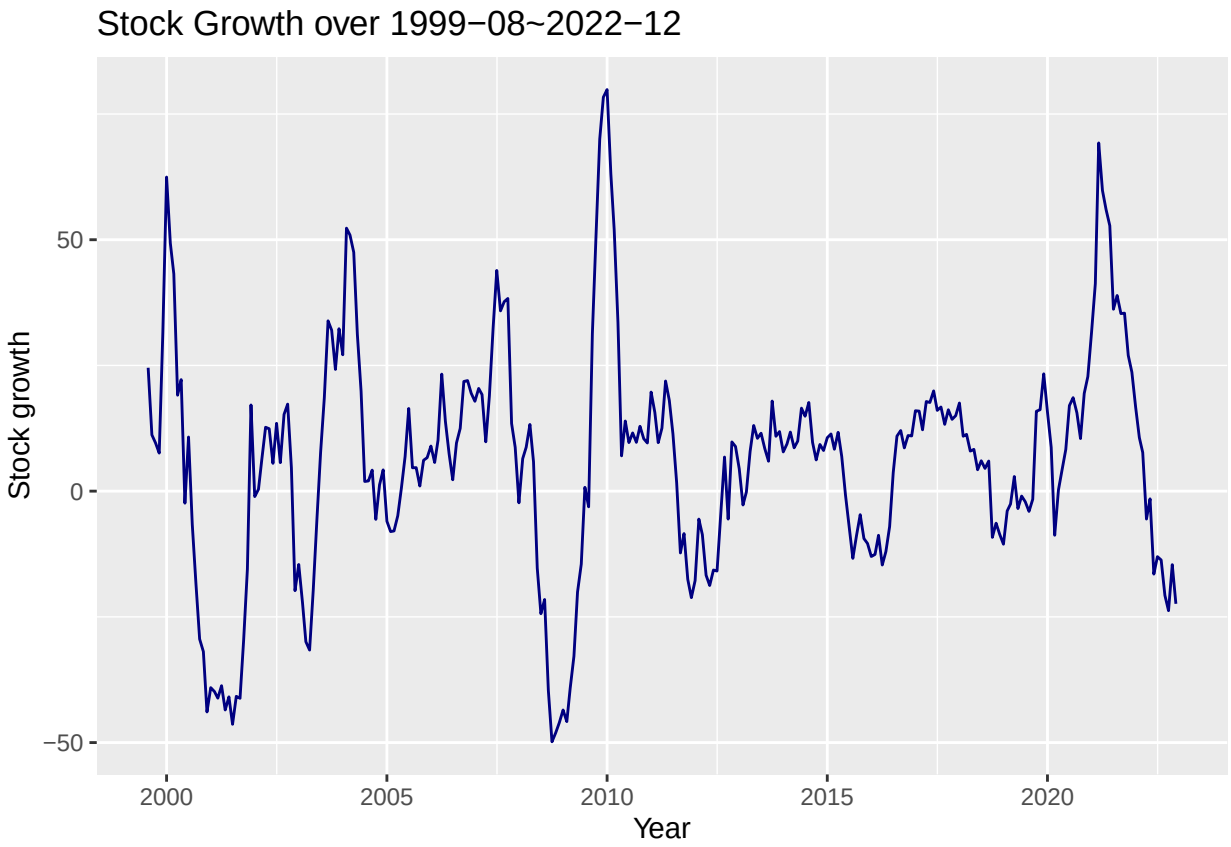
```
tail(stock_raw)
```

```
## # A tibble: 6 x 3
##   year       original growth
##   <date>        <dbl>  <dbl>
## 1 2022-07-01   15000.  -13.0
## 2 2022-08-01   15095.  -13.7
## 3 2022-09-01   13425.  -20.7
## 4 2022-10-01   12950.  -23.8
## 5 2022-11-01   14880.  -14.6
## 6 2022-12-01   14138.  -22.4
```

1.2 作圖

```
library(ggplot2)

graph1 <- ggplot(data=stock_raw, mapping=aes(x=year,y=growth))+
  geom_line(color="navyblue")+
  labs(x="Year",y="Stock growth",
       title="Stock Growth over 1999-08~2022-12")
graph1
```



1.3 建立虛擬變數

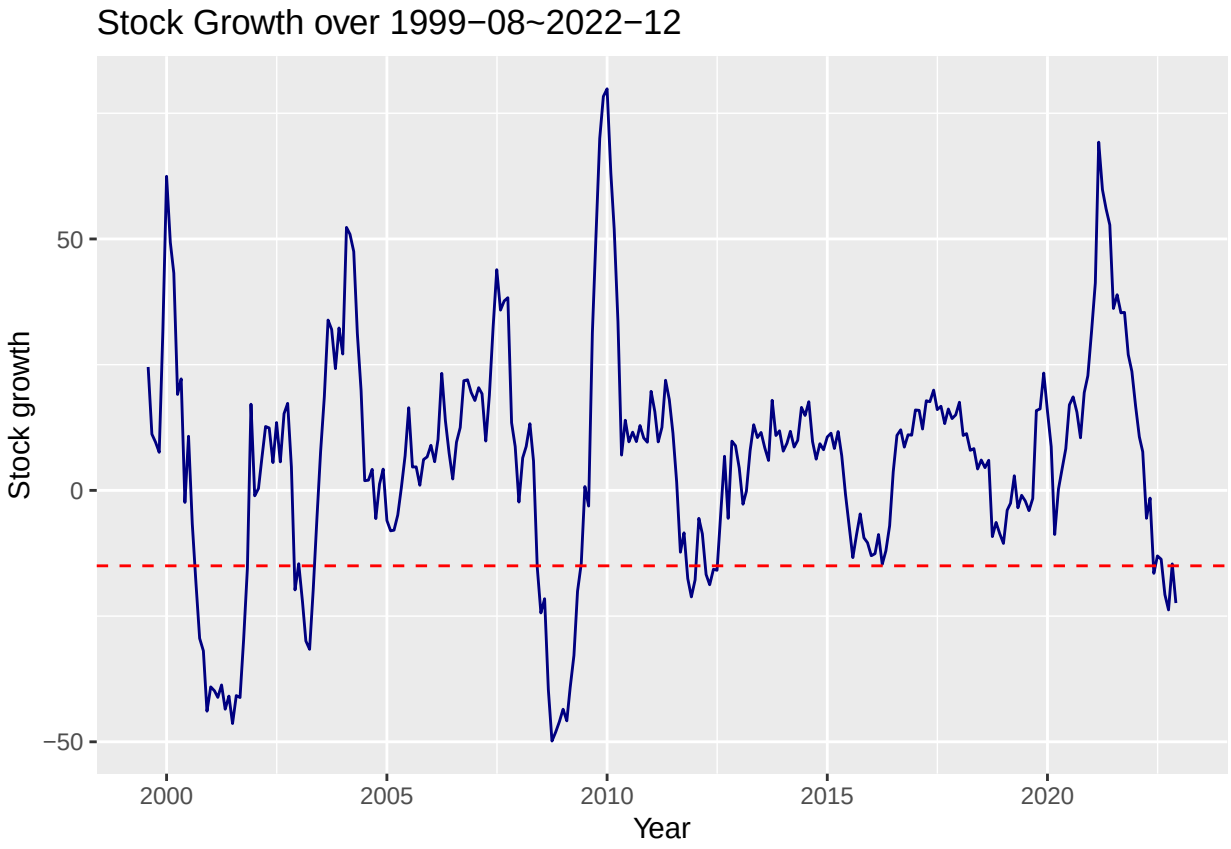
```
# add dummy variable
stock_raw$fall <- stock_raw$growth<(-15)
head(stock_raw)
```

```
## # A tibble: 6 x 4
##   year       original growth fall
##   <date>        <dbl>  <dbl> <lgl>
## 1 1999-08-01    8158.  24.5  FALSE
## 2 1999-09-01    7599.  11.2  FALSE
## 3 1999-10-01    7855.   9.61 FALSE
## 4 1999-11-01    7721.   7.57 FALSE
## 5 1999-12-01    8449.  31.6  FALSE
## 6 2000-01-01    9745.  62.5  FALSE
```

```
tail(stock_raw)
```

```
## # A tibble: 6 x 4
##   year       original growth fall
##   <date>        <dbl>  <dbl> <lgl>
## 1 2022-07-01   15000.  -13.0 FALSE
## 2 2022-08-01   15095.  -13.7 FALSE
## 3 2022-09-01   13425.  -20.7 TRUE
## 4 2022-10-01   12950.  -23.8 TRUE
## 5 2022-11-01   14880.  -14.6 FALSE
## 6 2022-12-01   14138.  -22.4 TRUE
```

```
graph1+geom_hline(yintercept = -15, linetype="dashed", color="red")
```



2 工業生產指數

```
library(readr)

IPI_raw <- read_csv("IPI_raw.csv",
                    col_names = c("year","original","growth"))
```

```
## Rows: 288 Columns: 3
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (1): year
## dbl (2): original, growth
##
## 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.
```

```
#transform the data type
IPI_raw$original <- as.numeric(IPI_raw$original)
IPI_raw$growth <- as.numeric(IPI_raw$growth)
```

```
#transform to date
IPI_raw$year <- as.character(IPI_raw$year) # transform to string first

IPI_raw$year <- as.Date(paste0(substr(IPI_raw$year, 1, 4),
                               "-",
                               substr(IPI_raw$year, 6, 7), "-01"),
                        format = "%Y-%m-%d") #default day=01
```

```
head(IPI_raw)
```

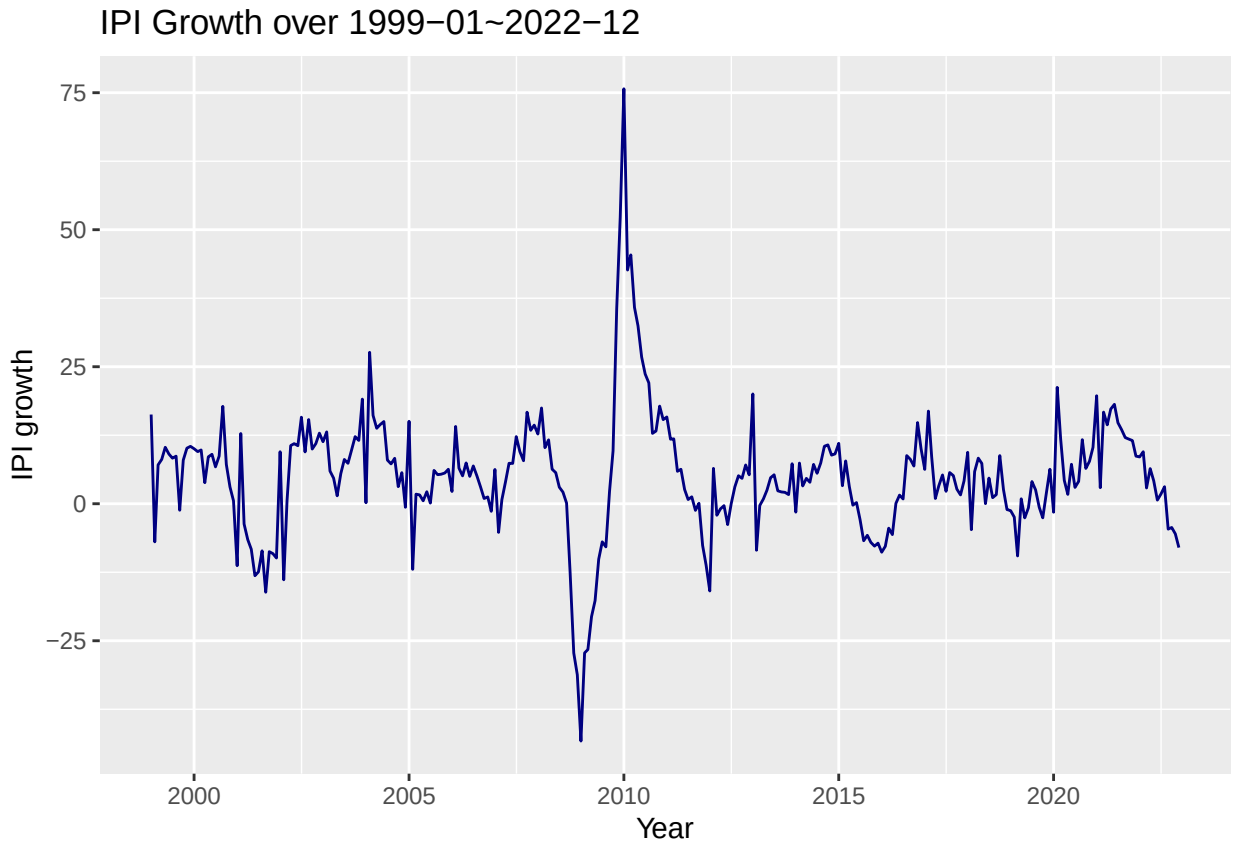
```
## # A tibble: 6 x 3
##   year       original growth
##   <date>        <dbl>  <dbl>
## 1 1999-01-01     47.9  16.3
## 2 1999-02-01     39.6  -6.95
## 3 1999-03-01     51.4   7.09
## 4 1999-04-01     51.0   8.1
## 5 1999-05-01     51.7  10.3
## 6 1999-06-01     51.4   9.07
```

```
tail(IPI_raw)
```

```
## # A tibble: 6 x 3
##   year       original growth
##   <date>        <dbl>  <dbl>
## 1 2022-07-01     137.   1.75
## 2 2022-08-01     141.   3.11
## 3 2022-09-01     132.  -4.63
## 4 2022-10-01     130.  -4.33
## 5 2022-11-01     130.  -5.5
## 6 2022-12-01     130.  -8
```

```
library(ggplot2)

graph2 <- ggplot(data=IPI_raw, aes(x=year, y= growth)) +
  geom_line(color="navyblue")+
  labs(x="Year", y="IPI growth", title="IPI Growth over 1999-01~2022-12")
graph2
```

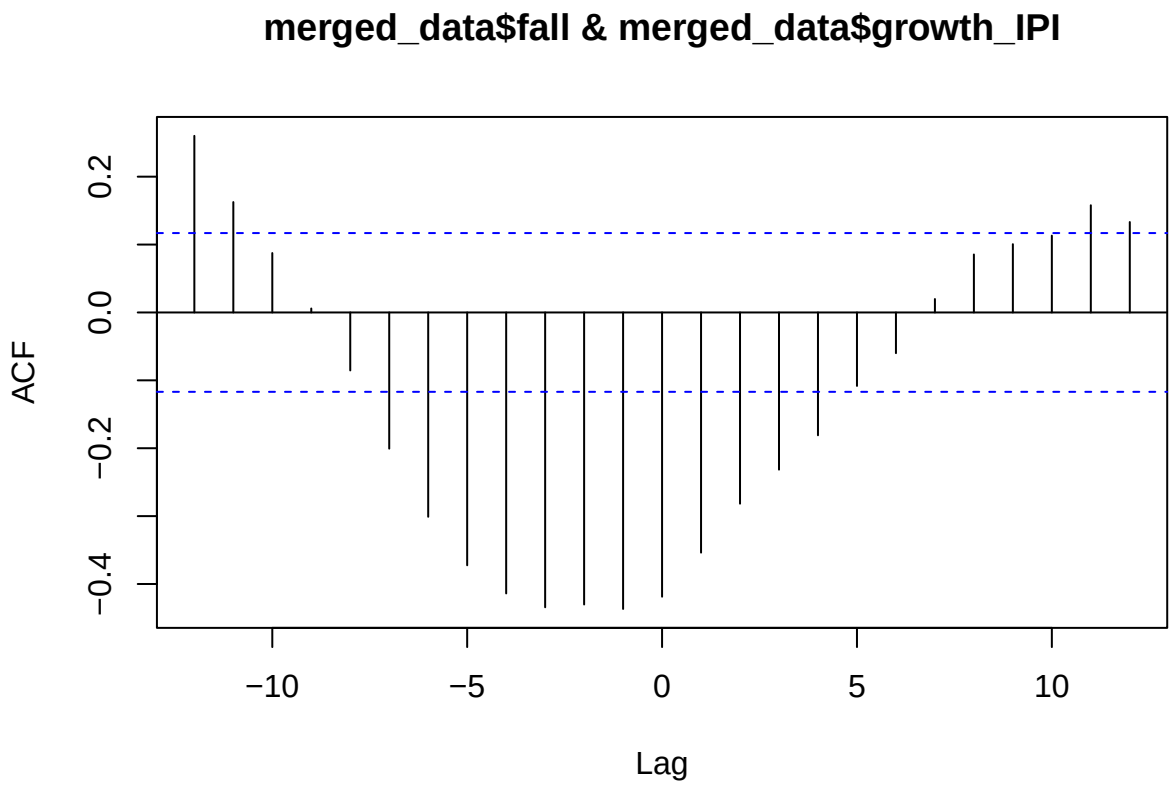


3 Cross Correlation Function(CCF)

```
merged_data <- merge(IPI_raw, stock_raw, by="year", all.y=TRUE)
merged_data$year <- as.Date(merged_data$year, format = "%Y-%m-%d")

colnames(merged_data) <- c("year","original_IPI","growth_IPI",
                           "original_stock","growth_stock","fall")
```

```
ccf_result <- ccf(merged_data$fall,merged_data$growth_IPI, lag.max=12)
```




```
ccf_values <- ccf_result$acf

lags <- ccf_result$lag

ccf_values_0_to_12 <- subset(ccf_values,lags>=0, lags<=12)

ccf_values_0_to_12
```

```
##  [1] -0.41875623 -0.35379838 -0.28169609 -0.23155740 -0.18096555 -0.10824282
##  [7] -0.05998879  0.01979543  0.08543237  0.10057198  0.11317794  0.15791655
## [13]  0.13319973
```

4 解釋

根據上面的計算，可以看出股價大跌與工業生產指數有一定程度的負相關，但相關性不代表因果關係，若要嚴謹判斷股價指數是否會影響與工業生產指數，還需要應用其他時間序列的模型。