Problem 1

A VAR(p) model is specified as:

$$X_t = a + B_1 X_{t-1} + \cdots + B_p X_{t-p} + \varepsilon_t$$

where $a \in \mathbb{R}^n$ is a constant vector, and $B_1, \cdots, B_p \in \mathbb{R}^{n \times n}$ are constant matrices. ε_t are i.i.d random vectors with $\varepsilon_t \sim N(0, \Omega)$.
Using lag operator, we can re-write the formula as:

$$\Psi(L) X_t = (1 - B_1 L - \cdots - B_p L^p) X_t = a + \varepsilon_t \tag{1}$$

Let

$$B = B_1 + B_2 + \cdots + B_p$$
$$\Gamma_j = -(B_{j+1} + \cdots + B_p)$$

Then we can re-write:

$$\begin{aligned}
B_1 L &= BL - (B_2 + \cdots + B_p)L = BL + \Gamma_1 L \\
B_2 L^2 &= -(B_3 + \cdots + B_p)L^2 + (B_2 + \cdots + B_p)L^2 = (\Gamma_2 - \Gamma_1)L^2 \\
B_3 L^3 &= (\Gamma_3 - \Gamma_2)L^3 \\
&\vdots \\
B_{p-1} L^{p-1} &= (\Gamma_{p-1} - \Gamma_{p-2})L^{p-1} \\
B_p L^p &= -\Gamma_{p-1} L^p
\end{aligned}$$

To sum of the two sides:

$$\begin{aligned}
B_1 L + B_2 L^2 + \cdots + B_{p-1} L^{p-1} + B_p L^p &= BL + \Gamma_1 L + (\Gamma_2 - \Gamma_1)L^2 + \cdots + (\Gamma_{p-1} - \Gamma_{p-2})L^{p-1} - \Gamma_{p-1} L^p \\
&= BL + (\Gamma_1 L + \Gamma_2 L^2 + \cdots + \Gamma_{p-1} L^{p-1})(1 - L)
\end{aligned}$$

Thereofre, equation (1) can be re-written as:

$$\Psi(L) X_t = \left[1 - BL - (\Gamma_1 L + \Gamma_2 L^2 + \cdots + \Gamma_{p-1} L^{p-1})(1 - L)\right] X_t = a + \varepsilon_t \tag{2}$$

which proves equation (35) in lecture 3.
Since

$$X_t - X_{t-1} = (1 - L) X_t = \Delta X_t$$

Equation (2) can be written as:

$$X_t - BLX_t - (\Gamma_1 L + \Gamma_2 L^2 + \cdots + \Gamma_{p-1} L^{p-1}) \Delta X_t = a + \varepsilon_t$$

$$\Leftrightarrow X_t - BX_{t-1} - \Gamma_1 \Delta X_{t-1} - \Gamma_2 \Delta X_{t-1} - \cdots - \Gamma_{p-1} \Delta X_{t-p+1} = a + \varepsilon_t$$

Finally, we get the same formula as equation (36) from lecture 3:

$$X_t = a + BX_{t-1} + \Gamma_1 \Delta X_{t-1} + \Gamma_2 \Delta X_{t-1} + \cdots + \Gamma_{p-1} \Delta X_{t-p+1} + \varepsilon_t$$

MTH9893 Time Series Analysis HW3

- Group 01
- Author: Pan, Hongchao & Sun, Yu
- Kernel version: Python 3.5
- Packages: pandas, matplotlib,statsmodels, time
- Data:
  - Data were stored in the folder **datasets**
  - Q2: PX_LAST of SPX and VIX over last 5 years
- Notes:
  - The running time of this notebook is around 3s
  - The original **SPX index is not stationary, but the 1st order differential of SPX index is stationary**. Thereby test the Granger Causality between VIX and 1st order differential of SPX
  - The VIX index **does have Granger Causality** on SPX index, but SPX index does not have Granger Causality on VIX index with 95% confidence interval.
  - Guideline of Granger Causality Test:
    - Step 1: Staionary test of two time series data. If they both are stationary, go to Step 3. If not, go to Step 2.
    - Step 2: Data manipulation. Use methods (1st order differential works here) to make the time series data be stationary.
    - Step 3: Find the lag of VAR(p) model by using VAR.select_order() function
    - Step 4: Fit the model by using the desired lag and criterion (aic or bic)found in Step 3
    - Step 5: Granger Causality test by using test_causality function
    - All the test functions are in Python statsmodels package

Question 2

In [1]:
```
# import packages
import pandas as pd
import matplotlib.pyplot as plt
import statsmodels.tsa.vector_ar.var_model as var_model  # For VAR(p)
import statsmodels.tsa.stattools as statools
import statsmodels.tsa.stattools as tsa  # adfuller test
import time
```

In [2]:
```
# Record the running time of the notebook
startTALL=time.time()
```

In [3]:
```
# Define a function to get the data
def get_data():
    # Retrive the historical daily end of the day (PX_LAST)
    # of SPX and VIX over the last 5 years from folder 'datasets'

    # Get the data of SPX
    df_spx=pd.read_excel(io='datasets/SPX_Daily.xlsx',sheetname=0,pars
e_cols='A:B',skiprows=1)
    # Get the data of VIX
    df_vix=pd.read_excel(io='datasets/VIX_Daily.xlsx',sheetname=0,pars
e_cols='A:B',skiprows=1)

    # Both data has been sorted by the date from oldest to latest and
has same length

    # Rename the columns of both dataframe
    df_spx.columns=['Date','SPX PX_LAST']
    df_vix.columns=['Date','VIX PX_LAST']
    # Combine the two dataframe
    df=df_spx.merge(df_vix,on='Date',how='inner')

    return df
```

In [4]:
```
df_Q2=get_data()
```

In [5]:
```
# Plot the data in 1 figure
plt.figure(1,figsize=(12,6)) # Change figure size
# Plot SPX index
plt.subplot(211)
plt.plot(df_Q2['Date'],df_Q2['SPX PX_LAST'],'r-')
plt.ylabel('Close index')
plt.legend(['SPX'],loc='upper left')

# Plot VIX index
plt.subplot(212)
plt.plot(df_Q2['Date'],df_Q2['VIX PX_LAST'],'b-')
plt.xlabel('Date')
plt.ylabel('Close index')
plt.legend(['VIX'],loc='upper left')
plt.show()
```

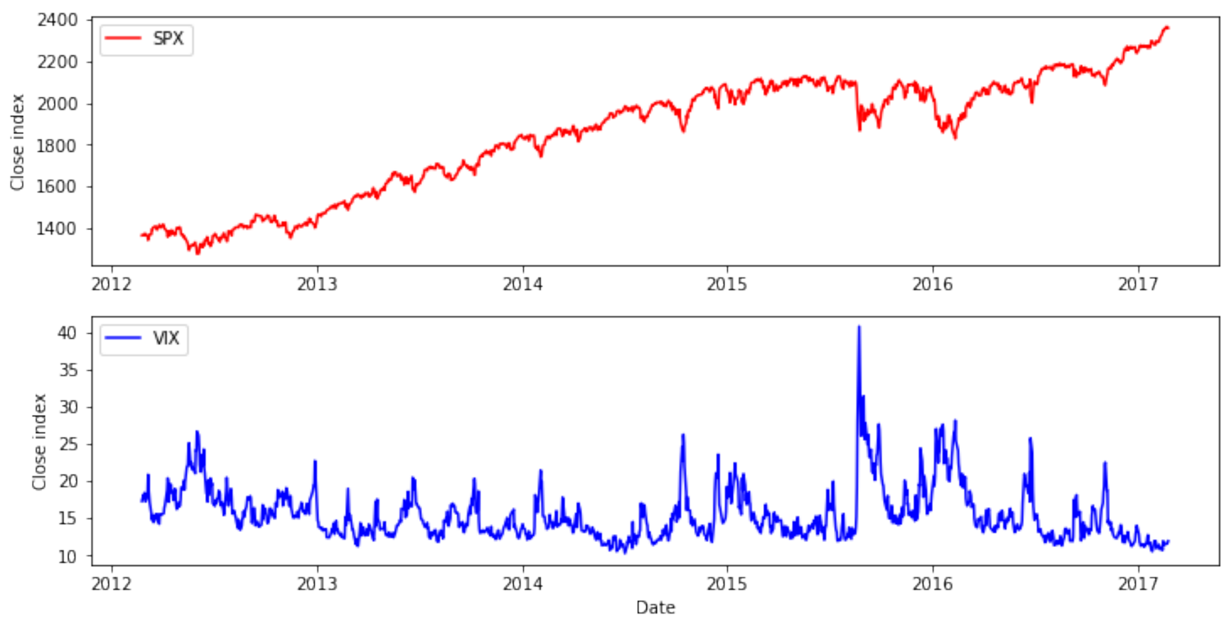

Figure 1 The overview of SPX and VIX close index over past 5 years

- **Observations**:
  - Figure 1 shows that VIX index jupms up when SPX index drops over the past 5 years
  - SPX index has a increasing trend indicates that it is **not stationary** over past 5 years, but VIX index looks more stable indicates that it is **stationary** over past 5 years. Thereby, further stationary test needs to be applied.

Stationary test

Define a test funtion to test whether a time series is stationary with **confidence interval 95%**

In [6]:
```
def stationary_test(TsData,CI):
    # TsData: to be tested time series data
    # CI: the critical value for test statistic, string of 1%, 5%, or
10%
    res=tsa.adfuller(TsData)

    # reference: https://en.wikipedia.org/wiki/Augmented_Dickey%E2%80%
93Fuller_test
    if(res[0]>res[4][CI]):
        return (res[0],res[4][CI],'non-stationary')
    else:
        return (res[0],res[4][CI],'stationary')
```

In [7]:
```
# Test the stationary of SPX index and VIX index
sta_SPX=stationary_test(df_Q2['SPX PX_LAST'],'5%')
sta_VIX=stationary_test(df_Q2['VIX PX_LAST'],'5%')
sta_df=pd.DataFrame({'SPX':list(sta_SPX),'VIX':list(sta_VIX)},
                   index=['adf(test statistic)','CI value','stationary
/non-stationary'])
sta_df
```

Out[7]:

	SPX	VIX
adf(test statistic)	-0.860425	-6.87366
CI value	-2.86384	-2.86384
stationary/non-stationary	non-stationary	stationary

Table 1 Stationary test of SPX index and VIX index over past 5 years

Change SPX index to stationary

Since stationary of two time series is required for Granger Causality test, we need make the SPX index to stationary.

Try differential method

In [8]:
```
# Take the 1st-order difference of SPX index data
df_Q2['SPX_diff1']=df_Q2['SPX PX_LAST'].diff()  # The 1st row of 'SPX_
diff1' will be NaN
plt.figure(1,figsize=(12,3)) # Change figure size
plt.plot(df_Q2.iloc[1:]['SPX_diff1'])  # Avoid NaN in 1st row
plt.xlabel('Date')
plt.ylabel('Close index')
plt.legend(['SPX_diff1'],loc='upper left')
plt.show()
```

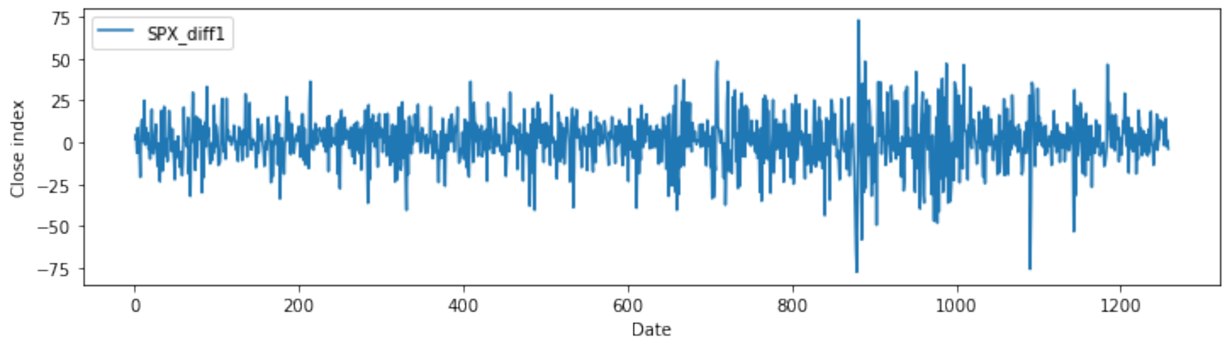

Figure 2 SPX close index with 1st order difference over past 5 years. Clearly, the 1st order differential data should be stationary.

- Stationary test of 1st order differential data of SPX index

In [9]:
```
# Stationary test of 1st order differential data of SPX index
tsa_SPX_diff=stationary_test(df_Q2.iloc[1:]['SPX_diff1'],'5%') # Avoid
1st row (NaN)
sta_df['SPX_diff1']=list(tsa_SPX_diff)  # Append result to the new col
umn of stationary test result
sta_df
```

Out[9]:

	SPX	VIX	SPX_diff1
adf(test statistic)	-0.860425	-6.87366	-19.4979
CI value	-2.86384	-2.86384	-2.86385
stationary/non-stationary	non-stationary	stationary	stationary

Table 2 Stationary test of SPX index, VIX index, and 1st order differential of SPX index over past 5 years

- Notes: Since the differential of SPX index has the information of SPX index and is stationary, we should use SPX_diff1 to test Granger Causality with VIX index

Test the Granger Causality between VIX and SPX (SPX_diff1)

- Calculate the lag of time series X=matrix(SPX_diff1,VIX) with VAR model

In [10]:
```
# Change index of df_Q2 to the date object for VAR model
df_Q2.set_index('Date',inplace=True)
# Erase the row has NaN data and the column of 'SPX PX_LAST'
df_gc=df_Q2.iloc[1:][['VIX PX_LAST','SPX_diff1']]
df_gc.head()
```

Out[10]:

	VIX PX_LAST	SPX_diff1
Date		
2012-02-27	18.19	1.85
2012-02-28	17.96	4.59
2012-02-29	18.43	-6.50
2012-03-01	17.26	8.41
2012-03-02	17.29	-4.46

Table 3 Overview of SPX and VIX dataframe

In [11]:
```
# VAR model
# Reference: http://statsmodels.sourceforge.net/devel/vector_ar.html
VAR_model2=var_model.VAR(df_gc)
# Select the lag
# Since the running time of the select_order function of VAR model, se
t the maxlags=30
res=VAR_model2.select_order(maxlags=30)
```

```
                VAR Order Selection
=====================================================
           aic          bic          fpe         hqic
-----------------------------------------------------
0        7.871        7.880        2621.        7.874
1        4.736        4.761        114.0        4.745
2        4.708       4.750*        110.9       4.724*
3       4.707*        4.766       110.8*        4.729
4        4.708        4.783        110.8        4.736
5        4.709        4.801        111.0        4.744
6        4.714        4.822        111.4        4.754
7        4.716        4.841        111.7        4.763
8        4.718        4.859        111.9        4.771
9        4.719        4.877        112.0        4.778
10       4.719        4.894        112.0        4.785
11       4.723        4.915        112.6        4.796
12       4.728        4.936        113.0        4.806
13       4.730        4.955        113.3        4.814
14       4.736        4.977        113.9        4.827
15       4.735        4.993        113.9        4.832
16       4.740        5.015        114.4        4.843
17       4.744        5.036        114.9        4.854
18       4.746        5.054        115.1        4.862
19       4.748        5.073        115.4        4.871
20       4.751        5.093        115.7        4.880
21       4.749        5.107        115.5        4.884
22       4.752        5.127        115.8        4.893
23       4.754        5.146        116.1        4.902
24       4.755        5.163        116.2        4.909
25       4.759        5.183        116.6        4.918
26       4.753        5.195        116.0        4.920
27       4.759        5.217        116.6        4.931
28       4.762        5.237        117.0        4.941
29       4.764        5.256        117.3        4.949
30       4.761        5.269        116.9        4.952
=====================================================
* Minimum
```

Table 4 Result of VAR(p) order selection: best model is **VAR(3)** based on AIC and **VAR(2)** based on BIC

In [12]:

```
# Fit the model based on the results above from select_order function
# Fit the model with desired lag, i.e., set the maxlags=3, ic='aic'
model_res=VAR_model2.fit(maxlags=3,method='ols',ic='aic',verbose=False
)
model_res.summary()
```

Out[12]: Summary of Regression Results

```
==================================
Model:                         VAR
Method:                        OLS
Date:           Tue, 28, Feb, 2017
Time:                     11:31:37
--------------------------------------------------------------------
No. of Equations:         2.00000    BIC:                    4.75074
Nobs:                     1255.00    HQIC:                   4.71499
Log likelihood:          -6492.68    FPE:                    109.230
AIC:                      4.69346    Det(Omega_mle):         108.022
--------------------------------------------------------------------
Results for equation VIX PX_LAST
====================================================================
=============
                    coefficient       std. error           t-stat
prob
--------------------------------------------------------------------
-------------
const                  1.082261         0.183868            5.886
0.000
L1.VIX PX_LAST         0.995805         0.052465           18.980
0.000
L1.SPX_diff1           0.005975         0.004664            1.281
0.200
L2.VIX PX_LAST        -0.086796         0.067920           -1.278
0.202
L2.SPX_diff1          -0.001931         0.004629           -0.417
0.677
L3.VIX PX_LAST         0.020567         0.051726            0.398
0.691
L3.SPX_diff1           0.001959         0.002563            0.764
0.445
====================================================================
=============

Results for equation SPX_diff1
====================================================================
=============
                    coefficient       std. error           t-stat
prob
--------------------------------------------------------------------
-------------
const                 -4.568336         2.067693           -2.209
0.027
L1.VIX PX_LAST        -2.406130         0.590000           -4.078
0.000
L1.SPX_diff1          -0.197727         0.052450           -3.770
0.000
L2.VIX PX_LAST         2.073284         0.763801            2.714
0.007
```

```
L2.SPX_diff1          -0.038136         0.052051            -0.733
0.464
L3.VIX PX_LAST          0.689379         0.581685             1.185
0.236
L3.SPX_diff1          0.001838         0.028827             0.064
0.949
====================================================================
=============

Correlation matrix of residuals
              VIX PX_LAST  SPX_diff1
VIX PX_LAST      1.000000  -0.843096
SPX_diff1       -0.843096   1.000000
```

Table 5 Summary results of VAR model fit with maxlags=3

Granger Causality test of VIX on SPX (SPX_diff1)

In [13]:

```
# Granger Causality test
# Reference: http://statsmodels.sourceforge.net/devel/generated/statsm
odels.tsa.vector_ar.var_model.VARResults.test_causality.html
# kind 'f': F-test
# confidence interval: 95%
# Test whether 2nd variable has Granger Causality on 1st variable, i.e
., test VIX on SPX_diff1
# Notes: the sequence/order of variables are matter!
gctest=model_res.test_causality('SPX_diff1','VIX PX_LAST',kind='f',sig
nif=0.05)
```

```
Granger causality f-test
==============================================================
   Test statistic   Critical Value          p-value         df
--------------------------------------------------------------
        10.165602         2.608468            0.000  (3, 2496)
==============================================================
H_0: ['VIX PX_LAST'] do not Granger-cause SPX_diff1
Conclusion: reject H_0 at 5.00% significance level
```

Table 6 The Granger Causality test of VIX index on SPX_diff1 index based on VAR(3) model

The results in table 6 shows that VIX index **does have Granger Causality** on SPX_diff1, which indicates that the VIX index **does have Granger Causality** on SPX index

Granger Causality test of SPX index (SPX_diff1) on VIX index

In [14]:
```
# Granger Causality test
# Reference: http://statsmodels.sourceforge.net/devel/generated/statsm
odels.tsa.vector_ar.var_model.VARResults.test_causality.html
# kind 'f': F-test
# confidence interval: 95%
# Test whether 2nd variable has Granger Causality on 1st variable, i.e
. test SPX_diff1 on VIX
# Notes: the sequence/order of variables are matter!
gctest2=model_res.test_causality('VIX PX_LAST','SPX_diff1',kind='f',si
gnif=0.05)
```

```
Granger causality f-test
=============================================================
   Test statistic   Critical Value          p-value         df
-------------------------------------------------------------
         0.853802         2.608468            0.464  (3, 2496)
=============================================================
H_0: ['SPX_diff1'] do not Granger-cause VIX PX_LAST
Conclusion: fail to reject H_0 at 5.00% significance level
```

Table 7 The Granger Causality test of VIX index on SPX index based on VAR(3) model

The results in table 7 shows that SPX_diff1 **does not have Granger Causality** on VIX index, which indicates that the SPX index **does not have Granger Causality** on VIX index

Answers/Observations of Q2

- The original **SPX index is not stationary, but the 1st order differential of SPX index is stationary**. And stationary of two time series is required of Granger Causality test, thereby we tested Granger Causality between 1st order differential data of SPX and VIX.

- The best fitted VAR(p) model of SPX_diff1 and VIX is p=3 based on AIC and fpe and p=2 based on BIC and hqic.

- From the Granger Causality test results, SPX and the SPX does not have Granger-causality on VIX, but **the VIX does have Granger-causality on SPX.**