timeseries_q2

December 16, 2018

1 Time Series Homework Question 2

In [1]:
```
import numpy as np
import pandas as pd
from statsmodels.regression.linear_model import OLS
from itertools import combinations
```

In [2]:
```
YEAR_NUM = 2000
df = pd.read_csv(f'../portfolio-analysis/{YEAR_NUM}_data.csv', index_col=0)
```

In [3]:
```
def is_cointegrated(x, y):
    nonstat_threshold = 0.5
    # Check if the autocorrelations have decayed well enough
    lower = 200
    upper = 210
    corr_threshold = 0.05

    w1 = OLS(x[1:].values, x[:-1].values).fit().params.item()
    w2 = OLS(y[1:].values, y[:-1].values).fit().params.item()
    if np.abs(w1) < nonstat_threshold or np.abs(w2) < nonstat_threshold:
        return False

    resid = OLS(y, x).fit().resid
    corr = np.array([resid.autocorr(lag=i) for i in range(lower, upper)])
    if (np.abs(corr) > 0.5).any():
        return False

    return True
```

In [4]:
```
# It looks like nothing is cointegrated with each other!
for data1, data2 in combinations(df.columns, 2):
    if is_cointegrated(df.loc[:, data1], df.loc[:, data2]):
        print(f'Cointegrated: {data1} and {data2}')
```

In [5]:
```
# We resort to a simulation...
# x = cumsum of Gaussian noise and y = x + noise.
# This makes them cointegrated.
```

```
x = pd.Series(np.random.randn(252).cumsum())
y = x + np.random.randn(252)
msg = 'Able to detect cointegration!' if is_cointegrated(x, y) else 'Unable to detect c
print(msg)
```

```
Able to detect cointegration!
```