It was shown how to identify two co-integrated stocks in the pair trade. In the example below, I will show how to form a mean reverting portfolio with three or more stocks, e.g. stocks with co-integration, and also how to find the linear combination that is stationary for these stocks.

First of all, we downloaded series of three stock prices from finance.yahoo.com.

### GET DATA FROM YAHOO FINANCE

quantmod::getSymbols("FITB", from = "2010-01-01")

FITB <- get("FITB")[, 6]

quantmod::getSymbols("MTB", from = "2010-01-01")

MTB <- get("MTB")[, 6]

quantmod::getSymbols("BAC", from = "2010-01-01")

BAC <- get("BAC")[, 6]

For the residual-based co-integration test, we can utilize the Pu statistic in the Phillips-Ouliaris test to identify the co-integration among three stocks. As shown below, the null hypothesis of no co-integration is rejected, indicating that these three stocks are co-integrated and therefore form a mean reverting portfolio. Also, the test regression to derive the residual for the statistical test is also given.

k <- trunc(4 + (length(FITB) / 100) ^ 0.25)

po.test <- urca::ca.po(cbind(FITB, MTB, BAC), demean = "constant", lag = "short", type = "Pu")

#Value of test-statistic is: 62.7037

#Critical values of Pu are:

# 10pct 5pct 1pct

#critical values 33.6955 40.5252 53.8731

po.test@testreg

# Estimate Std. Error t value Pr(|t|)

#(Intercept) -1.097465 0.068588 -16.00 <2e-16 \*\*\*

#z[, -1]MTB.Adjusted 0.152637 0.001487 102.64 <2e-16 \*\*\*

#z[, -1]BAC.Adjusted 0.140457 0.007930 17.71 <2e-16 \*\*\*

Based on the test regression output, a linear combination can be derived by [FITB + 1.097465 – 0.152637 \* MTB – 0.140457 \* BAC]. The ADF test result confirms that the linear combination of these three stocks are indeed stationary.

ts1 <- FITB + 1.097465 - 0.152637 \* MTB - 0.140457 \* BAC

tseries::adf.test(ts1, k = k)

#Dickey-Fuller = -4.1695, Lag order = 6, p-value = 0.01

Alternatively, we can also utilize the Johansen test that is based upon the likelihood ratio to identify the co-integration. While the null hypothesis of no co-integration (r = 0) is rejected, the null hypothesis of r <= 1 suggests that there exists a co-integration equation at the 5% significance level.

js.test <- urca::ca.jo(cbind(FITB, MTB, BAC), type = "trace", K = k, spec = "longrun", ecdet = "const")

# test 10pct 5pct 1pct

#r <= 2 | 3.26 7.52 9.24 12.97

#r <= 1 | 19.72 17.85 19.96 24.60

#r = 0 | 45.88 32.00 34.91 41.07

js.test@V

# FITB.Adjusted.l6 MTB.Adjusted.l6 BAC.Adjusted.l6 constant

#FITB.Adjusted.l6 1.0000000 1.000000 1.000000 1.0000000

#MTB.Adjusted.l6 -0.1398349 -0.542546 -0.522351 -0.1380191

#BAC.Adjusted.l6 -0.1916826 1.548169 3.174651 -0.9654671

#constant 0.6216917 17.844653 -20.329085 6.8713179

Similarly, based on the above Eigenvectors, a linear combination can be derived by [FITB + 0.6216917 – 0.1398349 \* MTB – 0.1916826 \* BAC]. The ADF test result also confirms that the linear combination of these three stocks are stationary.

ts2 <- FITB + 0.6216917 - 0.1398349 \* MTB - 0.1916826 \* BAC

tseries::adf.test(ts2, k = k)

#Dickey-Fuller = -4.0555, Lag order = 6, p-value = 0.01