Python Companion Course

Kevin Sheppard
University of Oxford
www.kevinsheppard.com

September 2019

©2019 Kevin Sheppard

Contents

Topic 1

Data and Simulation

1.1 Data Set Construction

Functions

`pd.read_csv`, `pd.read_excel`, `np.diff` or `DataFrame.diff`, `DataFrame.resample`

Exercise 1

1. Download all available daily data for the S&P 500 and the Hang Seng Index from Yahoo! Finance.
2. Import both data sets into Python. The final dataset should have a `DateTimeIndex`, and the date column should not be part of the `DataFrame`.
3. Construct weekly price series from each, using Tuesday prices (less likely to be a holiday).
4. Construct monthly price series from each using last day in the month.
5. Save the data to the HDF file "equity-indices.h5".

Exercise 2

Write a function that will correctly aggregate to weekly or monthly respecting the aggregation rules

- High: `max`
- Low: `min`
- Volume: `sum`

The signature should be:

```
def yahoo_agg(data, freq):
    <code here>
    return resampled_data
```

Exercise 3

1. Import the Fama-French benchmark portfolios as well as the 25 sorted portfolios at both the monthly and daily horizon from Ken French"s Data Library. **Note** It is much easier to clean to data file before importing than to find the precise command that will load the unmodified data.
2. Import daily FX rate data for USD against AUD, Euro, JPY and GBP from the Federal Reserve Economic Database (FRED). Use Excel rather than csv files.
3. Save the data to the HDF files "fama-french.h5" and "fx.h5"

Exercise 3 (Alternative method)

1. Install and use `pandas-datareader` to repeat the previous exercise.

Preliminary Step You must first install the module using

```
pip install pandas-datareader
```

from the command line. Then you can run this code. **Note**: Running this code requires access to the internet.

Exercise 4

Download data on 1 year and 10 year US government bond rates from FRED, and construct the term premium as the different in yields on 10 year and 1 year bonds. Combine the two yield series and the term premium into a `DataFrame` and save it as HDF.

1.2 Simulation

Functions

`np.random.standard_normal`, `np.random.standard_t`, `np.random.RandomState`

Exercise 5

Simulate 100 standard Normal random variables

Exercise 6

Simulate 100 random variables from a $N\left(.08,.2^2\right)$

Exercise 7

Simulate 100 random variables from a Students t with 8 degrees of freedom

Exercise 8

Simulate 100 random variables from a Students t with 8 degrees of freedom with a mean of 8% and a volatility of 20%. Note: $V[X]=\frac{\nu}{\nu-2}$ when $X\sim t_\nu$.

Exercise 9

Simulate two identical sets of 100 standard normal random variables by resetting the random number generator.

Exercise 10

Repeat exercise 7 using only `standard_normal`.

1.3 Expectations

Functions

`np.random.RandomState`, `RandomState.standard_normal`, `RandomState.standard_t`, `RandomState.chi2`, `np.exp`, `np.mean`, `np.std`, `scipy.integrate.quadrature`, `scipy.integrate.quad`

Exercise 11

Compute $E[X]$, $E\left[X^2\right]$, $V[X]$ and the kurtosis of X using Monte Carlo integration when X is distributed:

1. Standard Normal
2. $N\left(0.08, 0.2^2\right)$
3. Students t_8
4. χ^2_5

Standard Normal

t_8

χ^2_5

$N(8\%, 20\%^2)$

Function are useful for avoiding many blocks of repetitive code.

Exercise 12

1. Compute $E[\exp(X)]$ when $X \sim N\left(0.08, 0.2^2\right)$.
2. Compare this to the analytical result for a Log-Normal random variable.

Exercise 13

Explore the role of uncertainty in Monte Carlo integration by increasing the number of simulations 300% relative to the base case.

Exercise 14

Compute the $N(8\%, 20\%^2)$ expectation in exercise 11 using quadrature.

Note: This requires writing a function which will return $\exp(x) \times \phi(x)$ where $\phi(x)$ is the pdf evaluated at x.

Exercise 15

Optional (Much more challenging)

Suppose log stock market returns are distributed according to a Students t with 8 degrees of freedom, mean 8% and volatility 20%. Utility maximizers hold a portfolio consisting of a risk-free asset paying 1% and the stock market. Assume that they are myopic and only care about next period wealth, so that

$$U(W_{t+1}) = U(\exp(r_p) W_t)$$

and that $U(W) = \frac{W^{1-\gamma}}{1-\gamma}$ is CRRA with risk aversion γ. The portfolio return is $r_p = wr_s + (1-w) r_f$ where s is for stock market and f is for risk-free. A 4th order expansion of this utility around the expected wealth next period is

$$E_t[U(W_{t+1})] \approx \phi_0 + \phi_1 \mu_1' + \phi_2 \mu_2' + \phi_3 \mu_3' + \phi_4 \mu_4'$$

where

$$\phi_j = (j!)^{-1} U^{(j)}(E_t[W_{t+1}]),$$

$$U^{(j)} = \frac{\partial^j U}{\partial W^j},$$

$$\mu_k' = E_t\left[(r-\mu)_p^k\right],$$

and $\mu = E_t[r_p]$. Use Monte Carlo integration to examine how the weight in the stock market varies as the risk aversion varies from 1.5 to 10. Note that when $\gamma = 1$, $U(W) = \ln(W)$. Use $W_t = 1$ without loss of generality since the portfolio problem is homogeneous of degree 0 in wealth.

Topic 2

Estimation and Inference

2.1 Method of Moment Estimation

Functions

`DataFrame.mean`, `DatFrame.sum`, `plt.subplots`, `plt.plot`, `stats.kurtosis`, `stats.skewness`

Exercise 16

Estimate the mean, variance, skewness and kurtosis of the S&P 500 and Hang Seng using the method of moments using monthly data.

Exercise 17

Estimate the asymptotic covariance of the mean and variance of the two series (separately, but not the skewness and kurtosis).

Exercise 18

Estimate the Sharpe ratio of the two series and compute the asymptotic variance of the Sharpe ratio. See Chapter 2 of the notes for more on this problem.

The asymptotic variance is computed as

$$D\Sigma D'$$

where

$$D = [\sigma^{-1}, -1/2\mu\sigma^{-3}]$$

and Σ is the asymptotic covariance of the mean and variance. Finally, we divide by n the sample size when computing the statistic variance.

Exercise 19

Plot rolling estimates of the four moments using 120 months of consecutive data using a 4 by 1 subplot against the dates.

The simple approach to this problem uses a loop across 120, 121, . . . , n and computes the statistics using 120 observations. The figure is then created with a call to `plt.subplots` and the series can be directly plotted by calling `ax.plot`.

The pandas-centric approach uses teh `rolling` method to compute rolling statistics and then uses `.plot.line` with `subplots=True` to produce the plot.

2.2 Maximum Likelihood Estimation

Functions

`np.log`, `scipy.special.gamma`, `scipy.special.gammaln`, `scipy.stats.norm.cdf`, `scipy.optimize.minimize`, `scipy.stats.t`, `np.var`, `np.std`, `scipy.stats.norm.pdf`

Exercise 20

Simulate a set of i.i.d. Student's t random variables with degree of freedom parameter $\nu = 10$. Standardize the residuals so that they have unit variance using the fact that $V[x] = \frac{\nu}{\nu-2}$. Use these to estimate the degree of freedom using maximum likelihood. Note that the likelihood of a standardized Student's t is

$$f(x;\nu,\mu,\sigma^2) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi(\nu-2)}} \frac{1}{\sigma} \frac{1}{\left(1+\frac{(x-\mu)^2}{\sigma^2(\nu-2)}\right)^{\frac{\nu+1}{2}}}$$

where $\Gamma()$ is known as the gamma function.

Exercise 21

Repeat the previous exercise using daily, weekly and monthly S&P 500 and Hang Seng data. Note that it is necessary to remove the mean and standardize by the standard deviation error before estimating the degree of freedom. What happens over longer horizons?

Exercise 22

Repeat the previous problem by estimating the mean and variance simultaneously with the degree of freedom parameter.

Exercise 23

Simulate a set of Bernoulli random variables y_i where

$$p_i = \Phi(x_i)$$

where $X_i \sim N(0,1)$. (Note: p_i is the probability of success and $\Phi()$ is the standard Normal CDF). Use this simulated data to estimate the Probit model where $p_i = \Phi(\alpha_0 + \alpha_1 x_i)$ using maximum likelihood.

Exercise 24

Estimate the asymptotic covariance of the estimated parameters in the previous.

The derivative of the log-likelihood for a single observation is

$$\frac{\partial\left\{y_i \ln(\Phi(\alpha_0+\alpha_1 x_i)) + (1-y_i)\ln(1-\Phi(\alpha_0+\alpha_1 x_i))\right\}}{\partial \alpha_j}$$

which is

$$y_i \frac{\phi\left(\alpha_0+\alpha_1 x_i\right)}{\Phi\left(\alpha_0+\alpha_1 x_i\right)}-\left(1-y_i\right) \frac{\phi\left(\alpha_0+\alpha_1 x_i\right)}{1-\Phi\left(\alpha_0+\alpha_1 x_i\right)}$$

for α_0 and

$$y_i x_i \frac{\phi\left(\alpha_0+\alpha_1 x_i\right)}{\Phi\left(\alpha_0+\alpha_1 x_i\right)}-\left(1-y_i\right) x_i \frac{\phi\left(\alpha_0+\alpha_1 x_i\right)}{1-\Phi\left(\alpha_0+\alpha_1 x_i\right)}$$

for α_1 where $\Phi(\cdot)$ is the cdf of a standard Normal random variable and $\phi(\cdot)$ is the pdf of a standard Normal random variable.

2.3 Estimation: Bias and Verification of Standard Errors

Methods/Functions

`mean`, `var`, `RandomState`, `RandomState.chisquare`, `array`, `DataFrame.plot.kde`, `stats.norm.ppf`

Exercise 25

Simulate a set of i.i.d. χ^2_5 random variables and use the method of moments to estimate the mean and variance.

Exercise 26

Compute the asymptotic variance of the method of moment estimator.

Exercise 27

Repeat Exercises 24 and 25 a total of 1000 times. Examine the finite sample bias of these estimators relative to the true values.

Exercise 28

Repeat Exercises 24 and 25 a total of 1000 times. Compare the covariance of the estimated means and variance (1000 of each) to the asymptotic covariance of the parameters (use the average of the 1000 estimated variance-covariances). Are these close? How does the sample size affect this?

Exercise 29

In the previous problem, for each parameter, form a standardized parameter estimate as

$$z_i = \frac{\sqrt{n}\left(\hat{\theta}_i - \theta_{i,0}\right)}{\sqrt{\hat{\Sigma}_{ii}}}$$

where

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} N\left(0, \Sigma\right)$$

so that $\hat{\Sigma}$ is the estimated asymptotic covariance. What percent of these z_i are larger in absolute value than 10%, 5% and 1% 2-sided critical values from a normal?

Exercise 30

Produce a density plot of the z_i standardized parameters and compare to a standard normal.

Exercise 31

Repeat the same exercise for the Bernoulli problem from the previous question.

Topic 3

Linear Regression

3.1 Linear Regression: Model Estimation

Functions

`sm.OLS`

Exercise 32

Use the OLS function to estimate the coefficients of the Fama-French portfolios (monthly data) on the market, size and value factors. Include a constant in the regressions. Use only the four extremum portfolios – that is the 1-1, 1-5, 5-1 and 5-5 portfolios. Estimate the model with homoskedastic errors and with White's covariance estimator.

Exercise 33

Are the parameter standard errors similar using the two covariance estimators? If not, what does this mean?

Exercise 34

How much of the variation is explained by these three regressors?

3.2 Linear Regression: Rolling and Recursive

Functions

`sm.OLS`, `plt.title`, `plt.legend`, `plt.subplots`, `plt.plot`

Exercise 35

For the same portfolios in the previous exercise, compute rolling βs using 60 consecutive observations.

Exercise 36

For each of the four βs, produce a plot containing four series:

- A line corresponding to the constant β (full sample)
- The β estimated on the rolling sample
- The constant β plus $1.96\times$ the variance of a 60-observation estimate of β. The 60-month covariance can be estimated using a full sample VCV and rescaling it by T/60 where T is the length of the full sample used to estimate the VCV. Alternatively, the VCV could be estimated by first estimating the 60-month VCV for each sub-sample and then averaging these.
- The constant β minus $1.96\times$ the variance of a 60-observation estimate of β.

Exercise 37

Do the factor exposures appear constant?

Exercise 38

What happens if only the market is used as a factor (repeat the exercise excluding SMB and HML).

3.3 Linear Regression: Model Selection and Cross-Validation

Functions

`RandomState.permute`, `sm.OLS`, `set`, `scipy.random.norm.ppf`, `np.linspace`, `np.mean`

Exercise 39

Four portfolios we have been looking at, and considering all 8 sets of regressors which range from no factor to all 3 factors, which model is preferred by AIC, BIC, GtS and StG?

Exercise 40

Cross-validation is a method of analyzing the in-sample forecasting ability of a cross-sectional model by using $\alpha\%$ of the data to estimate the model and then measuring the fit using the remaining $100-\alpha\%$. The most common forms are 5- and 10-fold cross-validation which use $\alpha = 20\%$ and 10%, respectively. k-fold cross validation is implemented by randomly grouping the data into k-equally-sized groups, using k-1 of the groups to estimate parameters, and then evaluating using the bin that was held out. This is then repeated so that each bin is held out.

1. Implement cross-validation using the 5- and 10-fold methods for all 8 models.
2. For each model, compute the full-sample sum of squared errors as well as the sum-of-squared errors using the held-out sample. Note that all data points will appear exactly once in both of these sum or squared errors. What happens to the cross-validated R^2 when computed on the held out sample when compared to the full-sample R^2? (k-fold cross validated SSE by the TSS).

3.4 Linear Regression: Best Subset and Stepwise Regression

Functions

`np.linalg.lstsq`, `sklearn.model_selection.cross_val_score`, `sklearn.linear_model.LinearRegression`, `sklearn.model_selection.KFold`

Exercise 41

Download data from Ken French's website on the VWM and the 12 industry portfolios. Reformat both data sets so that they has a `DatetimeIndex`.

Exercise 42

Use Best Subset Regression with cross-validation to select the weights of a tracking portfolio where the industry portfolios are used to track the value-weighted-market. Use data until the end of 2014. A tracking portfolio is a portfolio that replicates a portfolio using other assets. The weights can be estimated using a regression model excludes a constant.

$$R_{i,p} = \beta_1 R_{i,1} + \beta_2 R_{i,2} + \ldots + \beta_k R_{i,k} + \varepsilon_i$$

where $R_{i,j}$ are returns on the assets used to construct the tracking portfolio. The constant is excluded since the portfolio should track both the shock and match the mean return. OLS minimizes the variance of the tracking error (in-sample).

Exercise 43

Select the best tracking portfolio using Forward Stepwise Regression.

Exercise 44

Use Backward Stepwise Regression to select the tracking portfolio.

Exercise 45

Using scikit-learn to cross-validate the Backward Stepwise selected models.

Exercise 46

Repeat the cross-validation using scikit-learn using randomly selected values in each block.

Exercise 47

Evaluate the models selected using the sample that was *not* used in fitting to assess the out-of-sample performance based on SSE.

Exercise 47

Compute the in-sample and out-of-sample R^2 for the selected models.
The out-of-sample R^2 defined as

$$1 - \frac{SSE_{OOS}}{TSS_{OOS}}$$

where the SSE_{OOS} is the SSE using the predicted values and the TSS_{OOS} is the TSS computed using the in-sample value (without demeaning since the models we are fitting do not include a constant). **Note**: If a model does not contain a constant, the TSS_{OOS} is *not* demeaned.

Exercise 48

Use scikit-learn to produce out-of-sample fits and compute the out-of-sample SSE. Verify this value is the same as you found previously.

3.5 Linear Regression: Ridge Regression and LASSO

Functions

`sklearn.linear_model.RidgeCV`, `sklearn.linear_model.LassoCV`, `sklearn.preprocessing.StandardScaler`

Exercise 49

Standardize the value-weighted-market return data and the 12 industry portfolios by their standard deviation. You should *not* remove the mean since we want to match the mean in the tracking portfolio.

Exercise 50

Select the optimal tuning parameter in a LASSO and estimate model parameters for the tracking error minimizing portfolio using the standardized data.

Exercise 51

Transform the estimated LASSO coefficients back to the scale of the original, non-standardized data.

Exercise 52

Select the optimal tuning parameter in a Ridge regression and estimate model parameters for the tracking error minimizing portfolio using the standardized data.

Exercise 53

Transform the estimated Ridge regression coefficients back to the scale of the original, non-standardized data.

Exercise 54

Compare the parameter estimates from the LASSO and Ridge regression to those from OLS in a plot. Use the original, non-standardized data.

Exercise 55

Use scikit-learn to scale the standardize the data by changing the scale but not the mean.

Exercise 56

Use the scikit-learn scaler to compute the predicted in-sample values using the Ridge ridge regression.

Exercise 57

Use the scalar from scikit-learn to produce out-of-sample forecasts of the two shrinkage estimators and OLS and evaluate the out-of-sample SSE.

Exercise 58

Directly produce out-of-sample forecasts of the two shrinkage estimators and OLS and evaluate the out-of-sample SSE without using scikit-learn.

3.6 Regression: Tree-based Methods

Functions

`sklearn.ensemble.RandomForestRegressor`, `sklearn.model_selection.GridSearchCV`, `sklearn.ensemble.GradientBoostingRegressor`

Exercise 59

Load the portfolio tracking data and compute the in- and out-of-sample SSE for OLS.

Exercise 60

Fit a default Random Forest in a reproducible manner to the portfolio tracking data and compute the in- and out-of-sample SSE.

Warning: This exercise is simply an example of how to use these methods. In general tree-based models are terrible choices for tracking portfolio construction since the final model is not a weighted combination of the returns, but instead depends on non-linear transformation of the returns. This makes implementation of a tree-based estimator virtually impossible.

Exercise 61

Optimize the key tuning parameters of the Random Forest using cross-validation and compute the out-of-sample SSE of the preferred model.

Exercise 62

Boosting is often a better alternative to Random Forests sine it limits tree depth, and in turn, variable interactions. Fit a default boosted regression tree to the portfolio tracking data, and compute the out-of-sample SSE.

Exercise 63

Optimize the key parameters of the boosted regression tree using cross-validation.

Exercise 64

Compute the out-of-sample SSE for the selected boosted regression tree.

Topic 4

Time Series Modeling

4.1 ARMA Modeling: Estimation

Functions

`tsa.SARIMAX`

Exercise 65

Estimate an AR(1) on the term premium, and compute standard errors for the parameters.

Exercise 66

Estimate an MA(5) on the term premium, and compute standard errors for the parameters.

Exercise 67

Estimate an ARMA(1,1) on the term premium, and compute standard errors for the parameters.

4.2 ARMA Modeling: Model Selection

Functions

`sm.tsa.SARIMAX`

Exercise 68

Perform a model selection exercise on the term premium using

1. General-to-Specific
2. Specific-to-General
3. Minimizing an Information Criteria

4.3 ARMA Modeling: Residual Diagnostics

Functions

`tsa.SARIMAX`, `sm.stats.diagnostic.acorr_ljungbox`, `SARIMAXResults.test_serial_correlation`, `statsmodels.sandbox.stats.diagnostic.acorr_lm`

Exercise 69

Compute the residuals from your preferred model from the previous exercise, as well as a random-walk model.

1. Plot the residuals
2. Is there evidence of autocorrelation in the residuals?
3. Compute the Q statistic from both sets of residuals. Is there evidence of serial correlation?
4. Compute the LM test for serial correlation. Is there evidence of serial correlation?

4.4 ARMA Modeling: Forecasting

Functions

`tsa.SARIMAX.forecast`

Exercise 70

Produce 1-step forecasts from your preferred model in the previous exercise, as well as a random-walk model.

1. Are the forecasts objectively accurate?
2. Compare these forecasts to the random walk models using MSE and MAE.

Note: Use 50% of the sample to estimate the model and 50% to evaluate it.

Exercise 71

Produce 3-step forecasts from the models selected in the previous exercises as well as a random walk model.

1. Compare these forecasts to the random walk models using MSE and MAE.

4.5 ARMA Modeling: Unit Root Testing

Functions

`sm.tsa.stattools.adfuller`, `arch.unitroot.ADF`

Exercise 72

Download data on the AAA and BAA yields (Moodys) from FRED and construct the default premium as the difference between these two.

1. Test the default premium for a unit root.
2. If you find a unit root, test the change.

Exercise 73

Download data on consumer prices in the UK from the ONS.

1. Test the log of CPI for a unit root.
2. If you find a unit root, test inflation for one.

Topic 5

Volatility Modeling

5.1 ARCH Model Estimation

Functions

`arch.arch_model`

Exercise 74

Estimate a GARCH(1,1) and a GJR-GARCH(1,1,1) to the returns of both series.

Note: You need to install arch using

```
pip install arch
```

which contains ARCH and related models.

Documentation and examples for the arch package are available online.

Exercise 75

Comment on the asymmetry.

- Compare robust and non-robust standard errors.
- Plot the fit variance and fit volatility.
- Plot the standardized residuals.

5.2 ARCH Model Selection

Exercise 76

Which model is selected for the S&P 500 among the classes: a. TARCH b. GJR-GARCH c. EGARCH

5.3 ARCH Model Forecasting

Functions

`sm.OLS`, `sm.WLS`

Exercise 77

Use 50% of the sample to estimate your preferred GARCH model for returns to the S&P 500 and the EUR/USD rate, and construct forecasts for the remaining period.

Exercise 78

Evaluate the accuracy of the forecasts.

We can also use a weighted LS estimator where the variance is the inverse of the squared forecast. The creates the MZ-GLS test.

Exercise 79

Evaluate the accuracy of forecasts from a 2-year backward moving average variance.

Exercise 80

Compare the ARCH-model forecasts to a naive 2-year backward looking moving average using QLIKE.

Topic 6

Value-at-Risk

6.1 Value-at-Risk: Using Historical Simulation

Functions

`Series.quantile`, `Series.rolling`

Exercise 81

Compute the 1-, 5- and 10-day historical simulation VaR for the S&P 500 and the EUR/USD rate.

Note: Start the historical simulation at 25% of the data, and then build the additional forecasts using a recursive scheme.

6.2 Value-at-Risk: Using Filtered Historical Simulation

Functions

`arch_model`, `ARCHModelResult.std_resid`, `np.percentile`

Exercise 82

Use a GARCH(1,1) model to construct filtered historical VaR for the S&P 500 and the EUR/USD exchange rate, using 10 years of data.

Note: For simplicity, estimate the model on the full sample, but start the historical simulation at 25% of the data, and then build the additional forecasts using a recursive scheme.

6.3 Value-at-Risk: Forecast Evaluation

Functions

`sm.OLS`, `stats.bernoulli`

Exercise 83

Compare this VaR to the HS VaR in the previous example.
import pandas as pd

Exercise 84

Evaluate the FHS and HS VaR forecasts constructed in the previous exercises using:

- HIT tests
- The Bernoulli test for unconditionally correct VaR
- Christoffersen's test for conditionally correct VaR

Topic 7

Vector Autoregressions

Vector Autoregressions ## Vector Autoregresion (VAR) Estimation

Functions

`tsa.VAR`

Exercise 85

Download data on 10-year interest rates, 1-year interest rates and the GDP deflator from FRED.

Exercise 86

Transform the GDP deflator to be percent returns (e.g. $\Delta \ln(GDP_t)$).

Exercise 87

Estimate a first-order VAR on the spread between the 10-year and 1-year (spread), the one-year, and the growth rate of the GDP deflator.

Exercise 88

What are the "own" effects?

Exercise 89

What are the cross effects between these?

Exercise 90

How could you get a sense of the persistence of this system?

7.1 VAR Model Order Selection

Functions

`tsa.VAR`

Exercise 91

Using the same data as in the previous exercise, determine the optimal VAR order using:

1. AIC
2. HQIC
3. BIC
4. Likelihood-ratio testing using General-to-Specific

7.2 Granger Causality Testing

Functions

`VARResults.test_causality`

Exercise 92

Using the data and the models selected in the previous exercise, is there evidence of Granger Causality between the series?

Exercise 93

What if the 10 year and the 1 year are both used, but the spread is omitted?

7.3 Impulse Responses

Functions

`VARResults.irf`

Exercise 94

Plot the impulse responses from both a first order model and the model selected in the order selection exercise.

Exercise 95

Which covariance factor makes sense for the impulse responses?

Exercise 96

What happens when you re-order the series and use the Cholesky factor in the impulse response? 1. Which series should be first? 2. Which should be last?

7.4 Cointegration Testing using Engle-Granger

Functions

`tsa.adfuller`, `tsa.coint`

Exercise 97

Download data on *CAY* from Martin Lettau's site.
We start by verifying that all three series have unit roots using `adfuller`.

Exercise 98

Is there evidence that these three series are cointegrated in the entire sample?

Exercise 99

What about in the post-Volker era (start in 1981)?

Exercise 100

Download monthly WTI and Brent crude prices from the EIA and import them. Examine whether the log prices are cointegrated.