FRE-GY-6971, Homework #2, due 4/18/2022, 11am

1. Definitions:
   1. Sample1: 3/1/2012 to 7/30/2015
   2. WFLY: 3Y\*w1 – 5Y + 7Y\*w2, weights = (w1,-1,w2)
2. Build a Jupyter Notebook to do the following:
   1. Download a panel of CMT rates into pandas dataframe & remove ‘1M column from the dataset
   2. Perform PCA on the dataset using Sample1
   3. Use this PCA model to analyze the CMT curve move on the 2016 Election Day: 11/8/2016 to 11/9/2016
      1. Plot CMT curve move vs the move explained by the first PCA factor, first 2 PCA factors, first 3 PCA factors
      2. Explain your calculations and results
   4. Compute weights of the WFLY to make sure that WFLY does not have PCA1,2 risk exposure in Sample1. Let’s call this combination WFLY1
   5. Choose weights of the WFLY from cointegration analysis (weights correspond to the best cointegrated vector). Let’s call this combination WFLY2
      1. Use Chou-Ng estimation procedure or Box-Tiao
      2. If cointegration estimation fails for you – use a linear regression of levels instead (regressing 5Y yield on [3Y,7Y] yields)
3. Compute Half-Life & ADF statistic for WFLY2 using Sample1
   1. Note that you are using time series of levels, not daily differences
4. Repeat Step #3 out-of-sample: using 3m, 6m out of sample periods
   1. How do out-of-sample results compare across periods and combinations?

**Independent reading assignment**:

1. What is CCA (Canonical Correlation Analysis)? Find 2 examples unrelated to cointegration example we studied in class
2. Read first 3 chapters of d’Aspremont’s paper
   1. How can we use CCA to construct the most mean-reverting portfolios? How can we use CCA to construct the least mean-reverting portfolios?
   2. What is the benefit of sparse decomposition described in Chapter 3?