Homework 9: Patterns of Exchange

36-402, Advanced Data Analysis

Due at the start of class on Tuesday, 5 April 2011

There are many variables which influence the rate at which one currency is exchanged against another: changes in interest rates, changes in inflation rates, changes in country's trade balances, the actions of speculators and central banks, etc.[1] Often these variables act in almost the same way at the same time in countries which occupy similar positions in the global economy, so we might expect substantial correlation in movements of exchange rates. This suggests that it might be useful to try to reduce the large number of exchange rates to a smaller number of components.

The data set for this week, `fx.csv` on the class website, contains the sequence of exchange rates, against the Swiss franc[2], of some commercially important and widely traded currencies: the US dollar, the Japanese yen, the Euro, and nineteen others. The data set runs from early 1995 to early 2010, with dates given as row names[3]. We will use principal components to analyze these exchange rate histories.

1. *Exploration and basic understanding* (10 points) Load the data file. It should have 3779 rows (trading days) and 22 columns (currencies).
   (a) (3 points) The column names are abbreviated three-letter codes standing for currencies. What are the names of the countries and of the currencies? (*Hint:* use a search engine and/or the library.)
   (b) (2 points) The exchange rate of dollars to yen on 28 February 2003 was approximately 118.22 yen to the dollar. Explain how to calculate this from the data. (*Hint:* The data give exchange rates for both in terms of Swiss francs.)
   (c) (5 points) Produce basic numerical or graphical summaries for each currency. Do they seem to have the same range and distribution? Should they?

[1] See Barry Eichengreen's *Globalizing Capital: A History of the International Monetary System* (Princeton University Press, 1996) for an excellent introduction and history.

[2] We could give prices for all the currencies in terms of any one of them, and it won't matter which one is the "numeraire". This week's data is a cleaned-up part of `FXRatesCHF` from the `fxregime` package, whose authors are Swiss.

[3] For the first few years of the data, the "Euro" is really the old German deutsche mark.

2. *Data transformation* (10 points) Economic theory for exchange rates is mostly about changes to rates, and in fact mostly about the proportional changes, as measured by $\log \frac{r_t}{r_{t-1}}$.

   (a) (5 points) Create a new data frame where each column gives the logarithmic change in the exchange rate for the corresponding currency in the original data. *Hint:* use `diff` and `apply`.

   (b) (5 points) Produce the same sort of summaries for currencies as in problem 1c. Are they now more similar to one another, or less?

3. *Covariance matrix* (20 points)

   (a) (8 points) Calculate the matrix of covariances between changes in exchange rates. This should be a 22×22 matrix. Provide it as a table, with at most three decimal places for each entry.

   (b) (8 points) Make a visual display of the matrix, using grey-scale, contour plots, color, or a perspective plot. Make sure the result is clearly legible when printed.

   (c) (4 points) Are there any groups of currencies which seem particularly strongly correlated?

4. *PCA* (25 points)

   (a) (4 points) Perform PCA on the *transpose* of the rate-of-change data frame, so that countries are rows and dates are columns. (*Hint:* use the `t()` function.) Include your code.

   (b) (8 points) Make a plot of how much variance is account for by the first k components, for k from 1 to 22. How much variance do the first two components, taken together, retain? How many components are needed to capture 50% of the variance? To capture 90%?

   (c) (9 points) Make a figure where the three-letter symbol for each currency is plotted according to the currency's score on the first two principal components. *Hint:* see examples in the notes for the lecture on PCA.

   (d) (4 points) Describe any clusters or patterns you see in the plot from the previous part. Do these match the strongly correlated currencies from the previous question? Should they?

5. *Examining the components* (15 points)

   (a) (5 points) Plot the sequence of weights (or "rotation") for the first principal component. Using the row names of the data set, and the `axis` command, add regularly spaced dates to the horizontal axis.

   (b) (5 points) Can you relate this time series to major economic events since 1995?

(c) (5 points) Repeat parts (a) and (b) for the second principal component.

6. *Reconstruction error* (10 points)

   (a) (2 points) Plot the logarithmic changes for the US dollar as a function of time (with dates as in part (a) of the previous problem).

   (b) (3 points) Plot the approximation to the dollar's history based on the first three principal components. Describe the similarities and differences.

   (c) (5 points) Calculate the mean squared error between the actual history of the dollar and the approximation based on the first q principal components, for q from 1 to 22.

EXTRA CREDIT (20 points): Fit a factor analysis model. Use hypothesis testing to determine the number of factors. Produce a figure like that of problem 5a for the leading factor. Comment on how it resembles or differs from the figure in 5a.