Introduction to Matlab for Windows

Januj Juneja

November 2008

1 Getting Started

The objective of this document is to familiarize the reader with some applications of Matlab software that are relevant to learning about variation in the structure of interest rates. Specifically, using principal components analysis in matlab to extract the factors driving variation in term structure behavior. Using Matlab will allow us to understand more about the factor structure that is driving interest rate behavior. We already gained some exposure to these factors in our discussion of "spread" trades, in that one of the underlying pieces of motivation was that we might want to bet on changes in the "slope" factor and not the "level". As these factors are unobservable, one of the learning outcomes for this topic is a better understanding of what we mean when we say "level" factor or "slope" factor. Let's start with an overview of the data that we will use to learn about these factors.

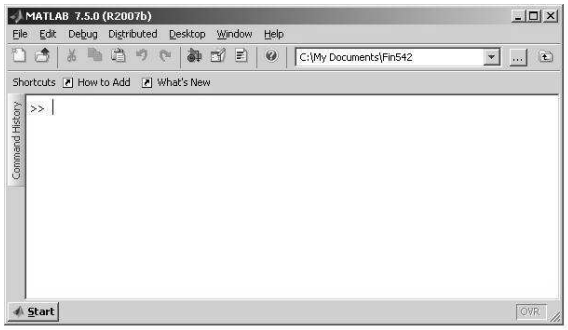
1.1 Data Cleaning Process

The data that we are going to use in class can be obtained from Dr. Lamoureux's website. On the course website, the data can be obtained by clicking on the link "Data file with 0-Coupon Treasury Yields". The name of the file is

"extbonds.txt". It is a text file, but before we will be able to apply principal components analysis, we will have to make the data ready to conduct the data analysis. We do this by opening the file in excel. After downloading the data to the appropriate directory (Let's call it C:/My Documents/Fin542), open it in excel. A "text import wizard" window will pop up and what you do is select "fixed width" in step 1, hit next until it finishes. This process allows you to import a text file into excel and sets it up as a spreadsheet for further manipulation of the data. Then, the data will be in excel and you will need to further manipulate the data in 2 ways. The first way deals with the fact that after August 2001, the Fed stopped issuing the long term bonds, so the data for the 15 year and the 25 year is the same. Hence, we will only run our data for as long as we have unique data on 5 bonds. The period you want to perform the analysis over is June 15, 1989 until August 23, 2001, this changes the sample size from n=821 observations to n=639. The other data cleaning aspect of the process involves creating a separate file in which you keep only the yield data for the 5 bonds; the 3 month, 6 month, 5 year, 15 year, and 25 year. So, you will have 5 columns of ONLY the interest rate data and not the date or discount factors or the SMT variable. Save this in your working directory (The C:/My Documents/Fin542 directory) as a .csv file and we will read it into matlab.

1.2 Getting Started with Matlab

When you open matlab you will see the following screen:



At the prompt, please type the following

$>> cd\ 'C:/MyDocuments/Fin542';$

Notice that the current directory address in the dialog box changes to point to that address. This directory should contain the data file, so that you are in the directory that contains the file which we will import from excel. Now to see what files you could import, type ls (this is shorthand for list). To read in the csv (csv means comma-separated variable, essentially the file is a comma delimted) file, type in the following at the prompt:

$>> X = csvread('XXX.csv');$

where XXX.csv is the name of your file which you store into the variable X. To see what variables are currently stored in your workspace, just type in whos at the prompt and it will tell you the size of the variable, how much space its taking up in memory, and its name; just some basic info about the variable. Once it is in the workspace, you data can be used as an input to the file, even if the file is in a different directory from the file that you run PCA from. In other words, once the yield data is in the workspace (the temporary area in memory that matlab uses to store files temporarily, to clear this workspace you should type clear at the prompt. However, you will need to download the data again to the workspace) Now, We will use the PCA code that I wrote, which uses the eigenvector decomposition function (eig.m function built into matlab) to compute the Principal Component factors, in the file titled PCA.m. You can download this file from Dr. Lamoureux's website, but please remember to save it in the directory in which your data is stored. Also, since it is saved as a text file (it has a .txt extension) please remember to right click on the link and save it with a .m extension so that the matlab compiler can recognize it as a matlab file. In this case, that would refer to the current directory which is C:/My Documents/Fin542. To view the source code for this file, just type:

$>> typePCA$

at the prompt.

To run the code type the following,

$>> Y = PCA(X, k)$

where k is the number of factors you want to use to compute the total explained variation in interest rates.

2 Writing Functions

If you want to write your own functions in matlab, just as I did in the above file called PCA.m, just click the new m-file button, the button with the piece of paper on it and a little explosion in the top left corner. and then you begin the function with the function keyword. The following is an example.

$functionY = myFunc(X)$

$placeyourcodehere$

end

This function takes in an input called X and it returns an output called Y. We could store it as myFunc.m (.m stands for matlab file). I could place several functions in one file and call each subfunction from the main function. For instance,

$functionY = myFunc(X)$

$Y = myFunc1(X);$

end

$functionY1 = myFunc1(X1)$

$Y1 = 3 + X1;$

end