SMF Problem 4. 17.05.2013

The purpose of this problem is to perform a Principal Components Analysis (PCA) in R on the US treasury rates with different maturities between Jan. 2 2003 and Apr. 11 2013. The corresponding data can be downloaded on the SMF tab of my webpage:
"http://www2.imperial.ac.uk/~bin06/".

(i) *Loading the data:* Use the function **read.csv** to load the data contained in it and the command **data.frame** to convert it to numerical values.

(ii) *Treating the missing values:* Replace the missing values (i.e. equal to "NA") e.g. by replacing them by the average between of previous and next value. (Hint: For all the maturities, the missing values happen to be on the same days)

(iii) *Taking the centred returns:* Create the matrix of the centred returns (i.e. for each series, take the first order difference –the returns– minus the mean of the returns) and, using the **pairs** function, inspect for correlation between the different variables.

(iv) *PCA:* Use the **princomp** function to run a PCA on the data. (Hint: the **summary** function gives the components weights, the **loadings** functions gives the components...) and comment on what components to retain.

(v) *Conclusion:* Interpret the meaning of the significant components.

NHB