18.3 PCA Example – Term-Structure Models "how many factors are enough?"

This Section considers the PCA and term-structure methods/models. The first Section provides a simple and straightforward PCA analysis of term-structure data. The Sections thereafter, continue to with the PCA considerations, but tie the analysis to modelling and issues relating more closely to trading. Notably, trading the yield curve on an outright, slope/spread, and curvature/butterfly basis is identical to an orhtogonalised PCA term-structure model.

18.3.1 PCA of Term Structure Data: Example Results

PCA analysis of a term-structure dataset is used to illustrate the basic aspects of the methodology. The calculations are performed on a 1-year history of US Treasury data. The data is used in its "raw" form.

> Caveat: while there is a temptation to interpolate the raw data to produce a "denser" spectrum of yields (e.g. the market may quote the 5-year and the 10-year and you wish to interpolate for the 6-, 7-, etc years) .. don't (unless you really know what you are doing[262]). Interpolation is just a linear combination of the raw data, and thus is "perfectly correlated", and so it is "perfectly redundant" in a PCA context. That is, the PCA will just need to spend extract time eliminating those as candidates.

The data has been "pruned" and only a selection of the maturity spectrum is used to illustrate the PCA methodology. An indicative listing of the data is shown in Figure 18.3 – 1.

	'3m'	'1y'	'2y'	'3y'	'5y'	'7y'	'10y'	'20y'
19-Nov-03	0.95	1.32	1.9	2.42	3.22	3.75	4.24	5.12
20-Nov-03	0.95	1.28	1.83	2.35	3.14	3.67	4.16	5.06
21-Nov-03	0.94	1.29	1.84	2.35	3.15	3.67	4.15	5.05
24-Nov-03	0.96	1.33	1.94	2.44	3.24	3.76	4.23	5.11
25-Nov-03	0.95	1.32	1.89	2.4	3.2	3.72	4.19	5.07
26-Nov-03	0.94	1.35	1.97	2.48	3.27	3.79	4.25	5.12
28-Nov-03	0.93	1.39	2.06	2.56	3.38	3.89	4.34	5.2
01-Dec-03	0.95	1.41	2.12	2.64	3.46	3.98	4.4	5.23
02-Dec-03	0.94	1.38	2.07	2.6	3.43	3.95	4.38	5.22
03-Dec-03	0.94	1.39	2.09	2.63	3.46	3.98	4.41	5.25
:	:	:	:	:	:	:	:	:

Figure 18.3 – 1. First few rows of US treasury PCA data.

[262] In some cases, expert knowledge and implementation can be used to test various curvature effects (see discussion of 2- vs. 3-factors in Section 18.3.2).

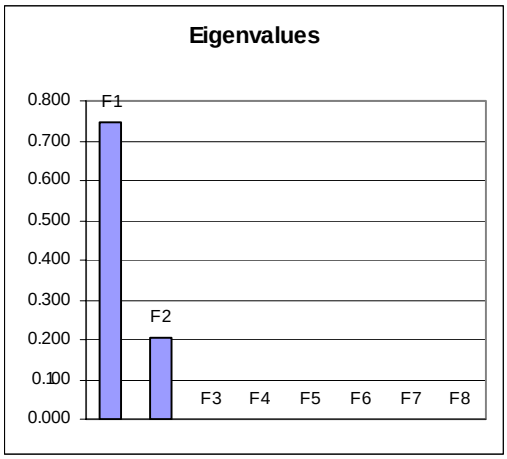
The analysis is performed using XLSTAT ® Addin [58] with MS Excel ®. XLSTAT ® offers the opportunity to analyse the data to any of the standard "PCA flavours". Here, the "covariance flavour" is used. XLSTAT ® produces a wide range of basic and advanced results, though here the focus is on the determination of the "minimum number of factors" that a term-structure model may require.

Figure 18.3 – 2 shows the eigenvalues for the data.

	F1	F2	F3	F4	F5	F6	F7	F8
Eigenvalue	0.749	0.204	0.002	0.001	0.000	0.000	0.000	0.000
% variance	78.333	21.308	0.213	0.098	0.024	0.015	0.006	0.004
Cumulative	78.333	99.641	99.854	99.952	99.976	99.990	99.996	100.000

Figure 18.3 – 2. XLSTAT ® eigenvalue results table for US yield curve PCA.

The image to right shows a plot of the eigenvalues. These figures illustrate strong evidence that the process driving the data can be explained by just two factors. That is, the degree of correlation is so high amongst the input yields that they look like just 2 factors, although 8 were supplied (i.e. the "angle between the streets" in Section 18.1 is effectively zero for most of these inputs).



Notice that this DOES NOT mean that the 3-month and the 1-year rates are the "important factors". It means that "2 factors" can explain 99.64% of the variability in the data history.

Those two factors are the principal components in the "eigenspace", and not in the "yield curve space". What those two factors are, from a trading perspective, have yet to be determined.

There is much analysis that is possible with PCA results, and post-processed PCA data. For example, the image to the right shows "XLSTAT ® Factors Plot". These types of representations can sometimes offer "hints" for further development. In this example, that the inputs are following an ascending progression, and from a positive quadrant to a negative quadrant is a "hint".