Recall that in FA we have an unobservability issue concerning the structural part of each observation, and thus the structural component of the var/cov matrix

\[ X_{i,o} = \Delta F_i = (\Delta \Theta) (\Theta' F_i) , \quad \Delta \Delta' = (\Delta \Theta)(\Theta' \Delta') \]

Last time we considered methods to produce

\[ \hat{\Delta} , \quad \hat{\Theta} , \quad \hat{\Delta}^* = \hat{\Delta} \hat{\Theta} \]

and select \( K \) (dimension of the \( F \)'s space; number of columns in \( \hat{\Delta} \)), but there is then the need to produce

\[ \hat{F}_i , \quad \hat{F}_i^* = \hat{\Theta'} \hat{F}_i , \quad i = 1...N \]

which represents our reduced data cloud, in \( K \) dimensions (regression techniques, see book chapter). In principle we could then pass to

\[ \hat{X}_{i,o} = \hat{\Delta} \hat{F}_i = \hat{\Delta} \hat{F}_i^* \leftarrow P_{\text{Span}(V_{i,k})} X_i \quad i = 1...N \]
\[ \hat{\Delta} = V_o \Lambda_o^{1/2} \]

Even when using pc method!

In PCA the reduced data cloud is simply

\[ P_{\text{Span}(V_{i,k})} X_i \leftrightarrow W_i = \begin{pmatrix} W_{i,1} \\ \vdots \\ W_{i,K} \end{pmatrix} \quad i = 1...N \]

and although the basis we are using in the \( W \)'s space makes very obvious sense (natural variability directions of the data), we might still want to change coordinate system for interpretation, passing to \( \Theta W_i \) for some appropriate rotation matrix.

The important thing to understand is that the subspace we identify through a dimension reduction exercise

- A low-dimensional subspace capturing most of the variability (PCA)
- A low dimensional subspace capturing the structural parts of the observations (FA)

is one, but the ways to span it are infinite!
On trial data, N=100, T=10

**Minitab**

Stat > Multivariate > Principal components (var/cov, looks like K=3 would work?)

**Variables**: x1 x2 … x10

**Number of factors to extract**: the already chosen K, or 10

**Type of matrix**: Correlation or Covariance (i.e. var/cov)

**Graphs**:
- Eigenvalue (scree) plot
- Score plot for first two components (i.e. projection of 100 points on the plane spanned by the first two eigenvectors)

**Storage**:
- Coefficients: v1 v2 … v10 (10 rows); 10 original coordinates of the 10 eigenvectors – all together, the V matrix
- Scores: w1 w2… w10 (100 rows); 10 coordinates of the 100 points in the eigenbasis

**Printed on screen**: eigenvalues, proportions, again the eigenvectors in original coord’s.

(Help option)

**Graph > Plot** (or 3D plot, or Matrix plot), of the w-columns, to visualize projections

**Graph > Time series plot**, of the v-columns, to visualize characteristic profiles given by PCA (superimpose on same graph, use color)

For interpretation, can run linear regressions of original coordinates on principal components

**Stat > Regression**

Response: x1 (or x2… etc)
Predictors: w1, w2, w3

R-sq in output, unusual observations
Stat > Multivariate > Factor Analysis (var/cov, K=3, both extraction methods)

Variables: x1 x2 … x10

Number of factors to extract: the already chosen K, or 10

Method of extraction: Principal components, or Max likelihood

Type of rotation: various

Type of matrix: Correlation or Covariance (i.e. var/cov)

Options:
Choice of correlation or var/cov matrix (calculated on data), or inputed matrix
Loadings for initial solution (i.e. solution before rotation); calculate or inputed matrix
Max likelihood extraction; option to use inputed initial communalities, options for number of iteration and convergence (see help)

Graphs:
Eigenvalue (scree) plot
Score plot for first two factors (i.e. projection of 100 points on the plane generated by the first two factors)
Loading plot for first two factors (i.e. plot of the first 2 entries in the K=3) rows of DeltaTheta matrix – rotated, vector drawing)

Storage:
Loadings: delta1 delta2 delta 3 (10 rows); the three columns of the DeltaTheta matrix (structural part of the original variables, in terms of rotated factors)
Coefficients: coeff1 coeff2 coeff3 (10 rows); factor scores coefficients?
Scores: f1 f2… f10 (100 rows); 3 rotated factor coordinates of the 100 points – Theta’Fi’s
Rotation matrix: 3 by 3, Theta
Residual matrix: S-DeltaDelta’
Eigenvalues (lambda’s) and eigenvector matrix (V); spectral decomposition when using the Principal components extraction method

Results:
What to print on the screen

(Graph > Plot (or 3D plot, or Matrix plot), of the f-columns, to visualize the reduced data cloud.
Graph > Time series plot, of the delta-columns, to visualize characteristic profiles given by FA (superimpose on same graph, use color).
Maniϕ > Display data, to view matrices.)
Chapter 13 from Methods of Multivariate Analysis. A.C. Rencher. Factor Analysis (again pay attention to different notation)

Papers on “Optimal” approaches to Factor Analysis (very technical, for those who are interested):


Della Riccia G., Shapiro A. Minimum rank and minimum trace of covariance matrices. Psychometrika, v. 47 n. 4, 1982

A couple of papers on 2D projection pursuing and touring

Posse C. Tools for two-dimensional exploratory projection pursuit. JCGS, v.4 n.2. 1995

Buja A., Cook D., Swayne D.F. Interactive high dimensional data visualization. JCGS, v.5 n. 1. 1996

(see “historic” references therein)

A paper with a general view of dimension reduction and the exhaustiveness/structure retention issue:


(brief discussion of projection pursuit and tours if time allows; exploratory approach, one, or a whole sequence, of 2D projections chosen according to a criterion – looking at the high-dimensional data cloud from a sequence of “viewpoints” that ought to be structurally informative)