UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Financial Econometrics

Topic 3:
Principal Component Analysis and Factor Analysis

Lecturer: Nico Katzke
nicokatzke@sun.ac.za

Department of Economics

DEPARTMENT OF ECONOMICS

What we will be discussing today

- We are looking today at using Principal Component Analysis techniques to reduce the dimensionality of data, as well as Exploratory Factor Analysis.
- Useful to econometricians, it provides a way of understanding **how correlated** the variables in a system are.
- Factor models, related to PCA but slightly different in application, offer parsimonious explanations of the underlying processes driving the data.

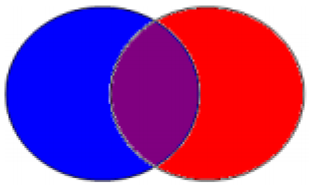
- (PCA and FA should not be confused – they are similar in that both reduce dimensionality and are factor analytic techniques, but differ in the mechanics and application and use different linear estimations).

Intuitively

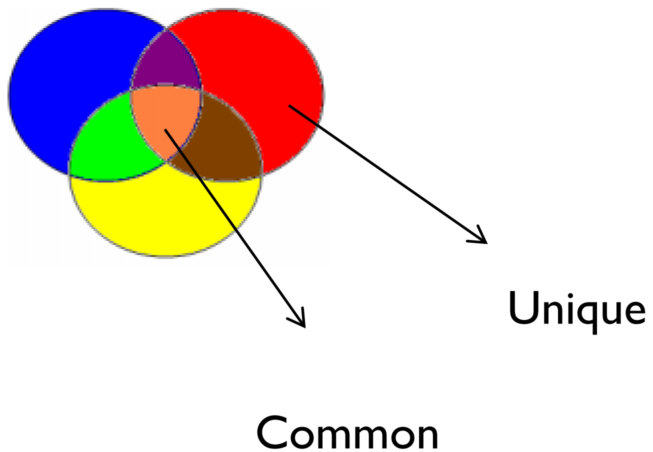
- Intuitively, we can think of PCA as a parsimonious means of analysing groups of correlated variables by finding optimal ways of combining such variables into smaller subsets that explain the variation.
- Factor analysis is, in turn, used to identify the structure underlying such variables – and seek to measure these latent variables (factors) themselves.
- Factors typically are related to real world features (such as productivity estimates, reading ability, etc), while for PCA the components are simply geometric abstracts that may / may not easily map into real world factors.

Graphically

- **PCA**: Extracts all components underlying a variable set.
- Amount of Components = amount of variables in the system.
- Thus the sum of all components' contribution to variance is 100%



- **Factor Analysis**: Analyses only the shared variance.
- This common variance is then distributed in **orthogonal factors**
- These factors are considered as causing variable variation.



PCA

- The goal of Principal Component Analysis is to reduce the dimensionality of the sources of variation in a dataset.
- Effectively it reorganizes the k-variables in the dataset into k −components, each being:
  - Independent
  - Orthogonalizes the variance contribution
    - Thus decreasing in the variance contribution as a result
- It therefore provides a mathematical basis for constructing a new set of orthogonal components that explain, in decreasing order, the source of variation.

PCA

- Components identified in PCA analysis are linear combinations of variables in the data set.
  - Variable combinations are based on calculated weights (eigenvectors)
- One problem with analysing data using PCA is that there is no clear statistical criterion for identifying significance.
- Rules of thumb are used for assessing what amount of components to consider (e.g. scree plots and using Kaizer's rule).

PCA

- Using PCA is useful for application to financial data as we get an idea of the commonality of variance between the included series of interest.
- PCA also does not require knowledge of the source of variation – and requires no ex ante specifications made on the data.
- Results are, however, at times interpreted as providing some deeper insight into the source of commonality (e.g. broadly defining sources as market, sector or global in nature) – which might be exaggerated.
- In this sense, it lets the data speak for itself.

Motivation

- Consider a dataset with 5 sector returns series, r_i.
- Mathematically, we now want to transform the covariance matrix of all the series, $\sum$, in such a way so as to identify components that are orthogonal and independent.
- Thus we want to maximize the diagonal elements (variance), and set to zero the off-diagonal elements (covariances) - much like a Cholesky decomposition approach.
- This is done by using Eigenvalue decompositioning.

Motivation

- Suppose we transform the data so that:

$$X = R.P$$

Where P is an orthonormal set (i.e. has orthogonal vectors) and X then the transformed data.

The Covariance set of the transformed data series, X, can then be written as:

$$\sum X = \frac{1}{n}.X'X = \frac{1}{n}(R.P)'(R.P) \;\rightarrow cov(X)$$

Now knowing that any covariance matrix is by definition symmetric, and knowing that any symmetric matrix can be represented as the product of a matrix of **eigenvectors** and a matrix of **eigenvalues** on the diagonal (D)…

Motivation

- We thus have:

$$\sum X = P'PDP^{-1}P = (P'P)D(P^{-1}P) = D$$

Using the fact that $P^{-1} = P'$ for the orthogonal matrix of eigenvectors and D is the diagonal matrix of eigenvalues (and zero off diagonals – so that source of variation is isolated).

This implies we can rewrite any variance matrix in terms of orthogonal sources of variation (or components).

The columns of P is thus the dimension along which we maximize the variance (called the eigenvectors).

Using correlation matrix instead.

- Of course there are practical problems to using the variance-covariance matrix – as scale and relative magnitude (and heterogeneity) would affect in the results.
- E.g. if one sector's return is highly volatile, this would bias the factors loadings toward explaining the highly volatile asset.
- One way of dealing with this is to use a standardized form, such as the **correlation** matrix as input.
- Using the correlation matrix as input – the eigenvalues on the diagonal will **sum to the number of variables** (can you see why?).

How many components to consider?

- Using the correlation matrix as opposed to the covariance matrix places all the variables on equal footing i.t.o. system variance contribution – i.e. as if each variable has variance = 1
- Using the Covariance matrix would imply variables with highest variation would dominate the first component – regardless of correlation with other variables.

- As eigenvalues sum to k, we could consider only those components that have values exceeding 1.

…How it is calculated

- Of course in practice the covariance matrix $\sum$ and correlation matrix ρ are unknown, yet it can be consistently estimated using the sample statistics as:

$$\widehat{\sum} = [\hat{\sigma}] = \frac{1}{T-1}.(r_t - \bar{r})(r_t - \bar{r})', \qquad \bar{r} = \frac{1}{\mathrm{T}}\sum r_t$$

$$\hat{\rho} = \hat{S}^{-1}\widehat{\sum}\,\hat{S}^{-1}, \qquad \hat{S} = diag(\sqrt{\widehat{\sigma_{11}}}, \ldots \sqrt{\widehat{\sigma_{kk}}})$$

How many components to consider?

- As there are no set of rules for deciding on how many components to regard as significant in contributing to variation, we can (e.g.) decide on the **amount of variation** that we want to explain (consider then components up to when the proportion of variance explained exceeds 80%), include all PC's larger than 1 (Kaizer) or we could define a set amount to add to meaning (e.g. consider only three PCs: 1st being market effects, 2nd being labelled as sector specific, 3rd labelled as global effects) ; another is to use the Broken Stick appraoch, Elbow approach, Kaizer-Guttman, etc.

Eigenvalues: (Sum = 7, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	5.045559	4.445853	0.7208	5.045559	0.7208
2	0.599706	0.200541	0.0857	5.645264	0.8065
3	0.399164	0.047710	0.0570	6.044429	0.8635
4	0.351454	0.118654	0.0502	6.395883	0.9137
5	0.232801	0.024959	0.0333	6.628684	0.9470
6	0.207842	0.044367	0.0297	6.836525	0.9766
7	0.163475	---	0.0234	7.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7
DLZACD	0.399653	-0.225599	-0.164710	-0.131873	0.133209	-0.730104	0.440530
DLZACS	0.393238	-0.205058	-0.275500	-0.139782	0.649176	0.529885	0.075283
DLZAEN	0.334756	0.690980	0.301130	-0.563787	-0.002749	0.025023	0.036283
DLZAFN	0.407591	-0.202602	-0.095173	-0.119991	-0.113909	-0.202294	-0.845853
DLZAIND	0.396247	-0.128436	-0.286717	-0.038236	-0.734807	0.358582	0.272584
DLZAMT	0.351651	0.523858	-0.191863	0.742692	0.088820	-0.065774	-0.036026
DLZATC	0.356061	-0.312463	0.823538	0.278863	-0.003478	0.108230	0.088780

Ordinary correlations:

	DLZACD	DLZACS	DLZAEN	DLZAFN	DLZAIND	DLZAMT	DLZATC
DLZACD	1.000000						
DLZACS	0.790434	1.000000					
DLZAEN	0.586605	0.576588	1.000000				
DLZAFN	0.827376	0.800076	0.610818	1.000000			
DLZAIND	0.779450	0.767194	0.593116	0.809718	1.000000		
DLZAMT	0.626554	0.623640	0.640192	0.640889	0.652977	1.000000	
DLZATC	0.683055	0.653110	0.516745	0.710432	0.650553	0.541231	1.000000

Table

- Consider the table at the top: we can decide to only use the first 2 components in order, e.g., to forecast the covariance of the 7 sector returns series.
- This follows as the first two components explain over 80% of the variation in the sectors.
  - Note that this implies a very high level of commonality in returns – and by default: a high degree of comovement confirmed by the high levels of correlation at the bottom.
- Thus – we can ignore the other components without too much loss in generality (and thus reduce dimensionality from 7 to 2).

Table

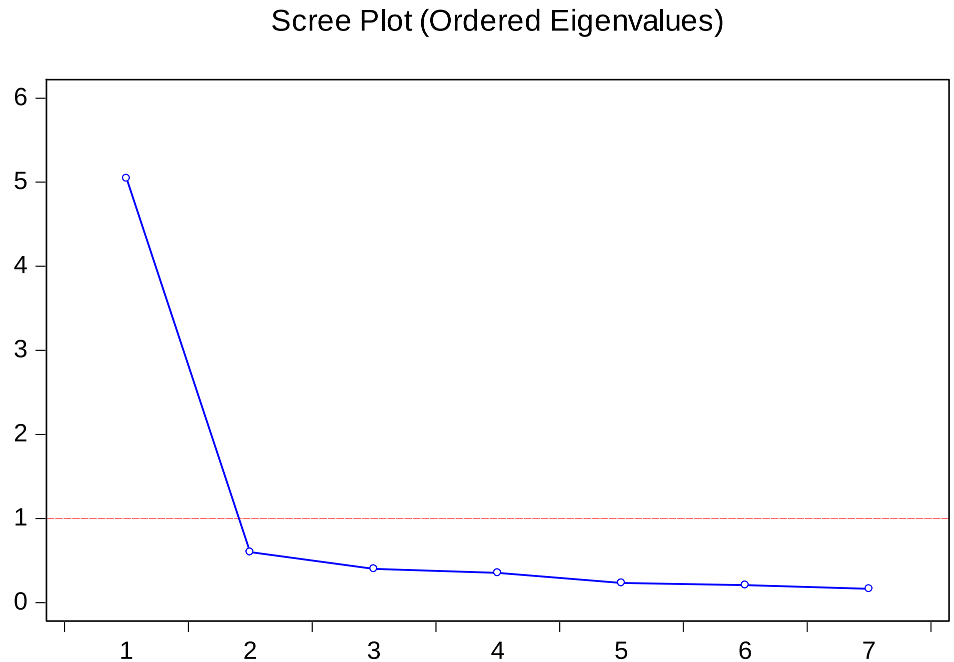
- From the table the proportion of variance explained is merely the Eigenvalue divided by the amount of variables (as we use the correlation measure) in the system ($\frac{5.045}{7}$ for $PC1$). Thus PC1 accounts for more than 5 times the variation of any single variable
- The second table reports eigenvectors, or loadings, which describes the linear combination of the variables that yield explain the variation measured. This can also be thought of as the **importance** of each variable in accounting for the component's variability
  - We see $\boldsymbol{PC}1$ having a roughly equal weighting contribution, which might be reasonably interpreted as a common market effect.
  - $\boldsymbol{PC}2$ shows that CD, CS, FN, IND and TC are loaded with the same sign (and roughly equal size), while EN & MT have a different sign and larger loadings (implying these sectors have a common factor loading for this PC and could be considered e.g. energy specific factors)…

Eigenvalues / Eigenvectors

- Conceptually, we can think of **EIGENVALUES** as the length (or relative strength) of an axis in the k −dimensional variable space (here: how much variation is explained).
- The **EIGENVECTORS** can then be thought of as determining the orientation of this axis – remember: eigenvectors are associated with the eigenvalues, and are not unique, and as such can be rotated.

Scree Plots

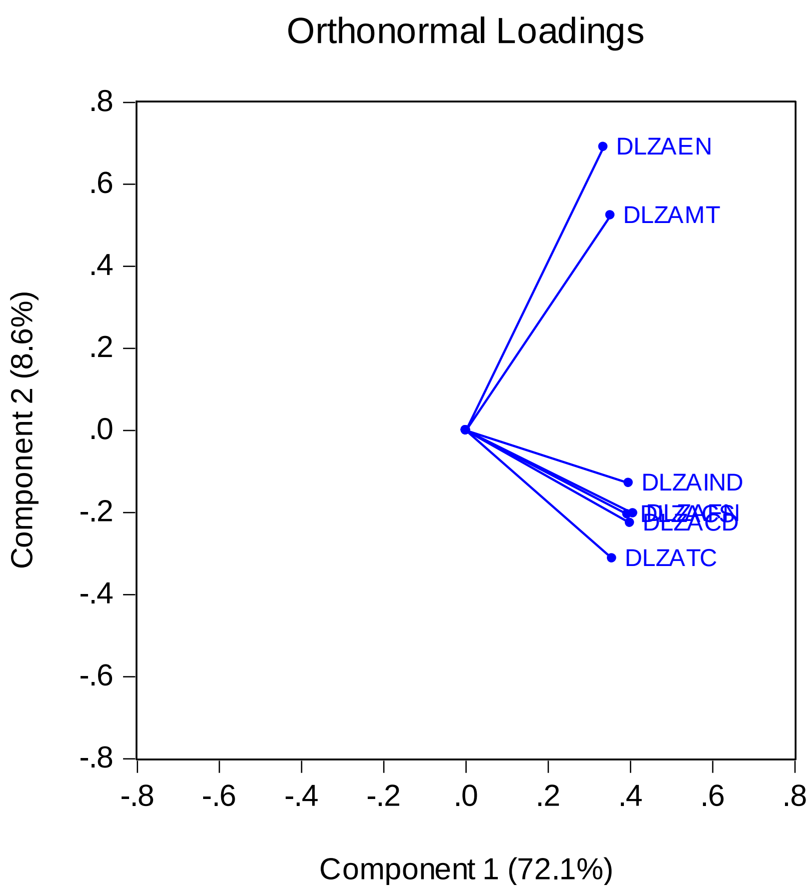
- We can also consider scree plots as a means of assessing how many components to regard as significant.
- By default, most statistical packages with built-in PCA abilities will show the ordered Eigenvalue plots as:



From this scree plot we can see that the PC2 – PC7 explain relatively little of the variation, and are roughly equal in contribution

Orthonormal loadings plot

- The orthonormal loadings plot can also be used to visually assess the loadings of the different variables for the respective PC's

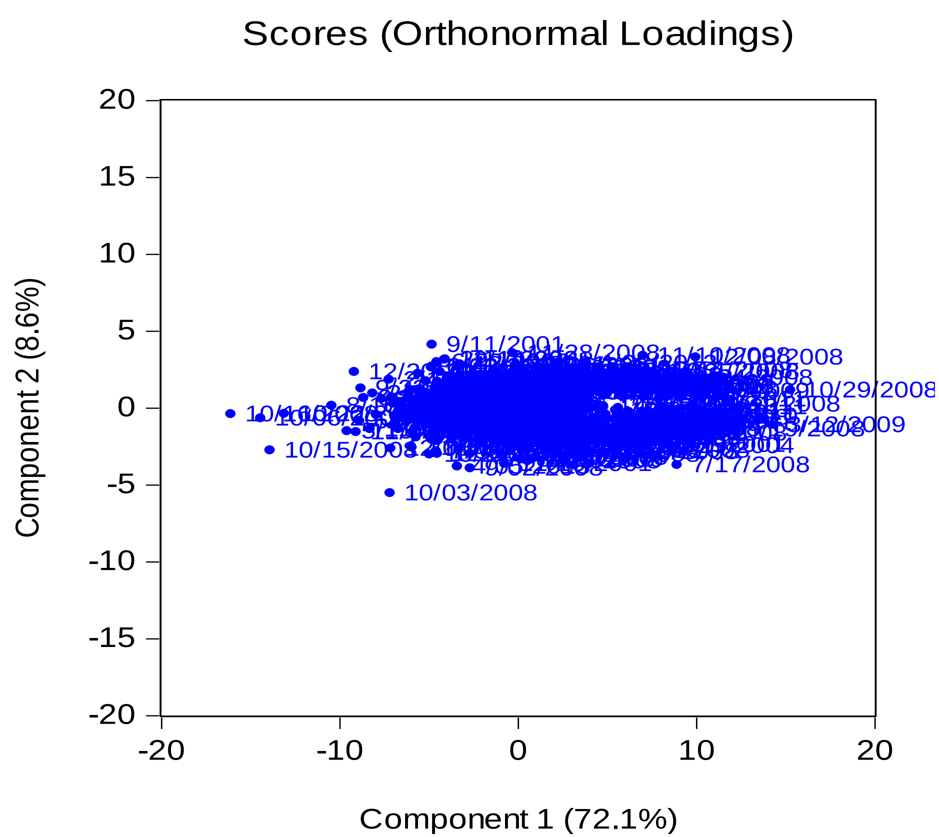


Note that for $PC1$, the loadings are roughly equal (around 0.4), while for $PC2$ – the loadings vary substantially.

Scores Plot

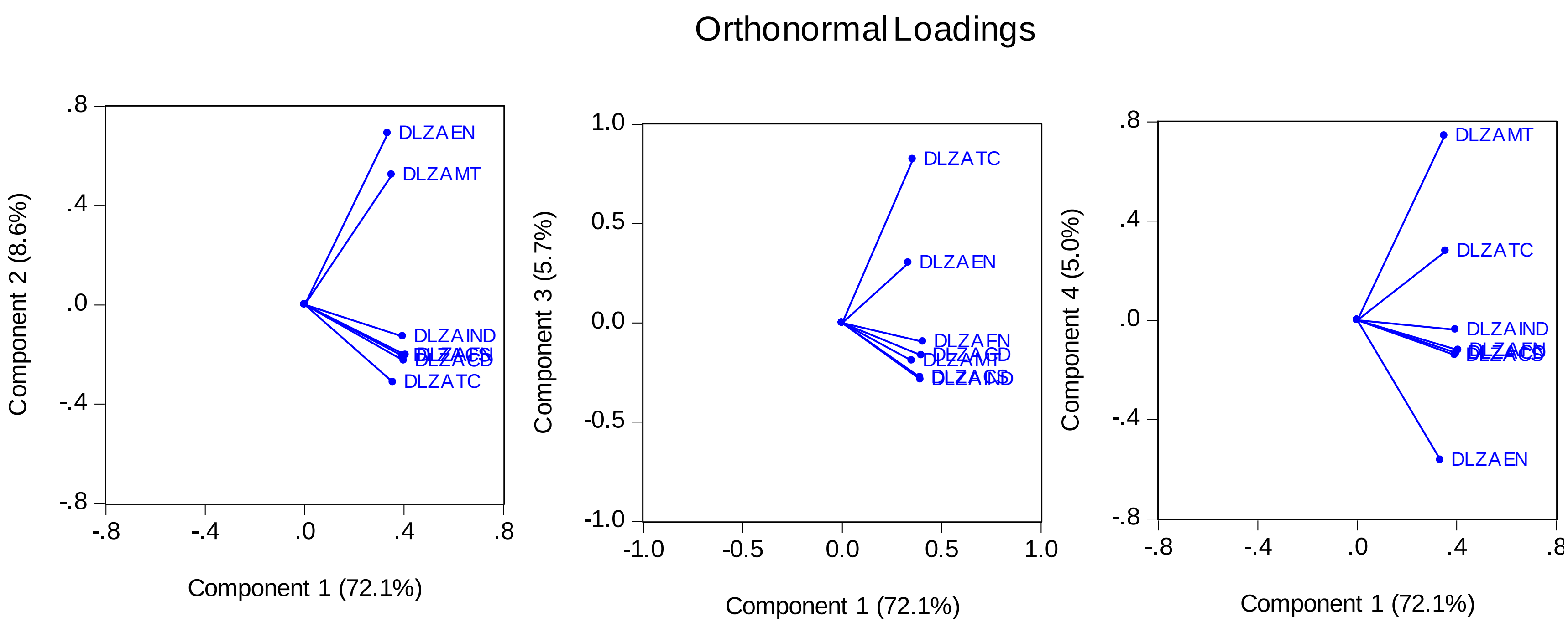
- This plots the actual values of the components for all the observations in the sample (good way to check for outliers which can strongly affect PCA results)

- From this we can clearly see that $PC1$ explains far more variation than PC2 does… (wider than it is high).
- **Outliers** are labelled by dates. We see most outliers being during 2008.



More PC plots

- To get an idea of the dispersion of the loadings across the PC's, we can plot any amount of loadings easily – e.g. select the first four and plot it in XY pairs



Usefulness

- Many applications of this approach have been suggested in the past, with e.g. Connor and Korajczyk (86, 88, 93) suggesting the use of PCA for creating APT mimicking portfolios using the eigenvector weights (see e.g. Goyal, et al (2008) extending this analysis).
  - The idea is that it allows investors to hold a weighted combination of assets whose return mimicks that of all the variables in the model.
- Other studies include Clarke, et al's (2006) construction of MVPs using PCA ; Meric, et al (2008) on co-movement of sectors ; Aziakpono, et al (2011) on interest rate analyses for African countries.

Considerations

- Using covariance / correlation as the dispersion matrix requires consideration of common variance (e.g. are we studying similarly volatile returns or e.g. bonds and equities).
- Are the variables included sufficiently correlated to start with – running PCA on uncorrelated variables does not provide us with much i.t.o. reduction in dimensionality.
- Outliers also have a significant impact on outcomes.
- If a variable has **low loadings** for the first few components – it indicates the variable is not contributing much to the variation in the system: Thus it can serve as a means of assessing which variables explain a significant proportion of system variation

UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Factor Analysis

Factor Analysis

- Factor analysis is a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of ***unobserved*** variables called factors.
  - Different from **PCA**, FA is a correlation-focused approach seeking to reproduce the inter-correlations among variables, **<u>in which the factors represent the *common* variance of variables. The part not explained is then defined as *asset specific* variance</u>**
- Factor analysis is also often used in data reduction to identify a small number of factors that explain most of the **variance** that is observed in a much larger number of variables.

FA - Intuitively

- Here follows a simple discussion of FA, as we will not be using this further in any great detail.
- Let’s consider how FA is used by referring to the oft cited example of Holzinger and Swineford (1939)’s psychological test results.
- The authors used psychological testing results from children in the form of 24 scores, including visuals, cubes, paragraphs, sentence, word meaning, etc.
- Their results are then aggregated in a Factor Analysis framework, where the first two significant factors explaining the variation constitute **verbal** and **spatial** attributes, e.g.

FA - Intuitively

- Suppose we define this example rather as 300 school students, and their scores in 4 tests (math, languages, science, biology) are M_1, M_2, M_3, M_4 respectively.
- Suppose we want to understand something about the students' **cognitive** and **verbal** abilities.
- We can then define two linearly unrelated factors, and <u>see what proportion of variance is ascribed</u> to it:

$$M_1 = \alpha_1 + \beta_{11}(comprehensive) + \beta_{12}(memory) + e_1$$

$$\dots$$

$$M_4 = \alpha_4 + \beta_{14}(comprehensive) + \beta_{14}(memory) + e_4$$

FA - Intuitively

- Comprehensive and Memory are then the two ***factors*** of interest.
- The Beta parameters above are then regarded as the ***loadings*** as with PCA.
- Suppose we find the loadings on the two unobserved (but specified by us) factors to be as follows:

Variable	Factor 1 Loading	Factor 2 Loading
M_1	+	0
M_2	0	+
M_3	0	+
M_4	+	0

Where 0 means small (insignificant) loading and + being of same sign and significant. This is then a plausible outcome, and we can interpret the importance of these abilities in explaining students' marks

FA - Intuitively

- In this simple example, as comprehensive and memory abilities are not **directly** observed, we have to infer such abilities on the variation explained by the factors.
- This requires assuming (in the most simple form) that:
  - The **errors** are **student specific factors** and are not correlated
  - The **common factors** are independent (**cognitive** and **memory** abilities are thus unrelated)
  - (More advanced models can see these assumptions relaxed).
- Making these assumptions, we then split the variation in all the variables between that which is **common**, and that which is **variable specific (idiosyncratic) – i.e. not accounted for by the two Factors**.

Types of Factor models with Financial Application

- Asset returns can e.g. be studied in a Factor model framework using (see Connor (1995)):
- **Macroeconomic Factor models** (using variables such as GDP growth rates, interest rates, inflation, etc. to describe commonality of variation in asset returns)
  - Here, factors are observable and the model estimated using linear regression analysis (so its not truly PCA type)
- **Fundamental Factor Models** (using firm specific factors, incl. firm size, Price/Book Values, etc.)
- **Statistical Factor Models** (similar to that which has been described in PCA – i.e. treating the common factors as unobservable and estimating it using a mathematical construct)

Asset Returns Factor Models

- Studying asset return factor models can thus be set up as follows: Let r_{it} be returns for asset i at time t

$$r_{it} = \alpha_i + \beta_{i1}.f_{1t} + \cdots + \beta_{im}.f_{mt} + \epsilon_{it}\ , \qquad \forall i, t$$

- With:

$$\beta_{i1} = \text{Factor Loading for asset } i'\text{s return to factor } f_1$$

$$f_{1t}$$

$$= \text{Factor 1, which is part of the } m$$

$$- \text{common factors driving the share returns.}$$

$$\epsilon_{it} = \text{the asset specific factors (assumed WN and uncorrelated with } \mathrm{f_m}$$

As mentioned, some factor models can relax the assumption of uncorrelated common factors.

Examples of Different Factor Models

- **<u>Macroeconomic Factor models:</u>**
  - Single Factor Model – See Sharpe (1970) in the market model (also defined as the CAPM model)
  - Multi-Factor model – see Roll, et all (1986) where the authors look at a Multifactor model for stock returns, explained by log differences of CPI and Employment figures.
- **<u>Fundamental Factor models:</u>**
  - Fama-French approach (1992) using 3 fundamentals:
  - Overall market return ; SMB (size) ; HML (Value vs Growth)

Examples of Different Factor Models

- **Statistical Factor Analysis:**
  - As defining factors that explain common behaviour might be of interest in certain fields – financial economists are typically not as excited by the prospects in explaining asset returns.
    - This is due in no small part to the immensely difficult task of identifying factors that clearly explain a significant part of the variation of asset returns.
  - Statistical FA therefore allows the modeller to infer from the data what factors drive common returns, without imposing such factors (e.g. CPI, GDP) or trying to interpret it *ex ante*.
  - We therefore treat the factors as UNOBSERVABLE

Using PCA to interpret Components as Factors

- As with our PCA example, we can interpret (loosely) the first two components as being, e.g., Market and Energy factors respectively.
- This implies doing Statistical FA using the Principal Component method.
  - This method does not require the assumption of Normality of the data or having factors prespecified / calculated up front.

- We can also use **MLE** methods (that use normal density and requires prespecification for the number of common factors) to calculate values for β and D directly…)

Factor Rotation

- Another useful tool in Factor Analysis is that of factor rotation.
- Consider the following Factor model on detrended returns series:

$$r_t - \mu = \beta . f_t + \epsilon_t$$

- It can easily be shown and intuitively understood that for any orthogonal matrix P, we have for:

$$\beta^* = \beta . P \;\; \& \;\; f_t^* = P' f_t$$

That:

$$r_t - \mu = \beta . f_t + \epsilon_t = \beta^* f_t^* + \epsilon_t$$

- Thus we can rotate factors without changing the results.

Factor Rotation

- This implies that we can rotate the common factor loadings in the m-dimensional space so that we have clearer insight into the strength of the respective loadings.
- Kaiser's **Varimax** criterion is often used as a means of rotation and works well in many applications.
  - The Varimax procedure selects the orthogonal rotation matrix P, such that it spreads out the squares of the loadings on each factor as much as possible.
- Basically, this aids in interpreting the loadings of common factors by more clearly spreading it out.

Fitting FA

- <u>We can fit factors to a Group as follows:</u>
- Remember that we want to detrend the series in our model and then split the dispersion matrix (e.g. typically the correlation matrix) into the **common variance** part and the **unique variance** part.
- This implies the returns series r_i to be written as:

$$r_t - \mu = \beta . f_t + \epsilon_t$$

With m-standardized factors, f_t, and the loadings contained in the matrix β. The residual will be labelled as the unique factors.

Fitting FA

- Remember the idea is that the Factor loading matrix, β, links the unobserved common factors to the observed data.
- The j^{th} row then loads the j^{th} variable's common factors – i.e. being the coefficients to link it to the linear factor model (acting like the eigenvectors in PCA).
- <u>We can then impose the following restrictions:</u>

$$E(f_i) = 0 \;\; ; \;\; E(\varepsilon_i) = 0 \;\; ; E(f_i . \varepsilon_i) = 0$$

$$E(f_i f_i{}') = \phi \;\; ; \;\; E(\varepsilon_i \varepsilon_i) = \sigma = Diagonal\ matrix\ of\ unique\ variance.$$

- <u>Then by construction we have:</u>

$$Var(r_i) = L\phi \mathrm{L}' + \sigma = common + unique\ variance$$

Fitting FA

- Then we can also calculate the **<u>factor structure matrix</u>** (which is the correlations between the variables and factors) :

$$Var(r,f) = E[(r_i - \mu).f_i'] = E[(\beta f_i + \varepsilon_i)F_i'] = \beta.\phi$$

The next step is to split the dispersion matrix into the common and unique components, is that we need to specify the amount of factors it needs to consider.

This is now similar to our discussion with the PCAs and deciding how many components to include in the analysis (and is arguably the single most important decision to make in FA.)

Fitting FA

- There exists then several methods to determining the optimal amount of factors to include:
  - Kaiser-Guttman (most commonly used): chooses factors based on the amount of eigenvalues calculated on the dispersion matrix (correlation matrix in our case) that have values larger than 1 (or exceeds the average if using covariance matrix).
  - Fraction of total variance explained – retain as many factors required for the sum of the $m -$ eigenvalues to exceed a threshold (say 80% / 95%)

Fitting FA

- MAP (minimum average partial) – computes average of the squared partial correlations after partialling out $m -$ components, choosing amount of factors that minimize this average.
- Others include Broken Stick (Jackson, 1993), S.E. Scree plot, Parallel Analysis (Humphreys & Montanelli, 1975), etc.

Fitting FA

- Estimation procedure
  - The next choice involves how to estimate the FA procedure which calculates the factor loadings and specific variances. These include (amongst others):
  - Principal Factors approach (see Gorsuch, 1993), Maximum Likelihood estimation of the Loadings directly, PACE, etc. can be chosen (see documentation for details).

Fitting FA

- Rotation
  - We then also choose between several measures for rotating the factor loadings to facilitate the interpretation thereof.
  - This follows as loadings are never unique, and can be rotated without changing the results.
- Scoring
  - The next consideration regards scoring. As factors are unobserved and implied from the observed data – its calculation depends on estimates of loadings. Factor score estimates thus can be used in further diagnostic analyses

And get…

	Loadings F1	Communalit...	Uniqueness
DLBRFNDM	0.583417	0.340375	0.659625
DLINFNDM	0.568124	0.322765	0.677235
DLZAFNDM	0.613676	0.376599	0.623401
DLCNFNDM	0.555513	0.308594	0.691406

Factor	Variance	Cumulative	Difference	Proportion	Cumulative
F1	1.348333	1.348333	---	1.000000	1.000000
Total	1.348333	1.348333		1.000000	

	Model	Independenc...	Saturated
Discrepancy	0.024964	0.819938	0.000000
Parameters	8	4	10
Degrees-of-freedo...	2	6	---

Control for partial effect of US-fn:

	Loadings F1	Communalit...	Uniqueness
DLBRFNDM	0.485095	0.235317	0.764689
DLINFNDM	0.576163	0.331964	0.668033
DLZAFNDM	0.548496	0.300848	0.699166
DLCNFNDM	0.583607	0.340597	0.659404

Factor	Variance	Cumulative	Difference	Proportion	Cumulative
F1	1.208726	1.208726	---	1.000000	1.000000
Total	1.208726	1.208726		1.000000	

	Model	Independenc...	Saturated
Discrepancy	0.033806	0.467683	0.000000
Chi-square statisti...	110.6463	1530.726	---
Chi-square prob.	0.0000	0.0000	---
Bartlett chi-square	110.5506	1529.713	---
Bartlett probability	0.0000	0.0000	---
Parameters	8	4	10
Degrees-of-freedo...	2	6	---

Running a PCA estimate on this group

Eigenvalues: (Sum = 4, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulativ... Proportion
1	2.094560	1.294227	0.5236	2.094560	0.5236
2	0.800333	0.220101	0.2001	2.894893	0.7237
3	0.580231	0.055355	0.1451	3.475124	0.8688
4	0.524876	---	0.1312	4.000000	1.0000

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3	PC 4
DLBRFNDM	0.499676	-0.543445	0.183324	0.649141
DLINFNDM	0.494689	0.485371	-0.686678	0.219479
DLZAFNDM	0.518061	-0.435939	-0.164209	-0.717360
DLCNFNDM	0.487052	0.528242	0.684033	-0.125855

Ordinary correlations:

	DLBRFNDM	DLINFNDM	DLZAFNDM	DLCNFNDM
DLBRFNDM	1.000000			
DLINFNDM	0.308375	1.000000		
DLZAFNDM	0.469924	0.350234	1.000000	
DLCNFNDM	0.309877	0.422823	0.326416	1.000000

Which to use?

PCA or FA?

Do not use both

- You should not be using both estimation procedures on the same data, as they say something completely different.
- **PCA** – weighted linear combination of correlated variables, seeking to explain the maximal amount of variation by specifying a number of components which reduces rank
- **FA** – constructs unobserved (latent) factors that explain the variation of the observed variables (by indirectly measuring its influence on the observed variables), and are influenced by measurement error which we define as **unique** (asset specific / idiosyncratic) **factors**.

Sum

- If we seek to explain the variation in a dataset – we use FA to specify the factor of interest and then use it to explain which part of the variation is explained by this (unobserved, but implied) variable.
- The components of the PCA do not have interpretable value – it merely states the amount of variance that is attributable to the specific linear component calculated.

PCA model:

$$r_i = \beta . Y$$
$$\beta = eigenvectors$$
$$Y = scores\ on\ components\ (Eigenvalues)$$

Exploratory FA Model:

$$r_i = \beta . Y + E$$
$$Y = common\ factors$$
$$\beta = Weights \setminus Factor\ Loadings$$
$$E = Unique\ (idiosyncratic) factors$$

Conclusion: Keep in Mind…

- When studying Statistical FA or PCA's – we assume **no serial autocorrelation.** This therefore has to be controlled for first if present (we can fit a VARMA model and use the residuals, e.g.).

- Also, note that the series need to be detrended – we will use the scale function in R.