Financial Times Series

Lecture 11

PCA and Factor models

- When we consider a portfolio containing many assets it may be the case that the returns can be described by just a few linear combinations of the individual asset returns

- It may also be the case that there are exogenous factors or macroeconomic variables, observable or non-observable firm specific or non-firm specific factors that can be use to describe return behaviour

Formally

- Assuming that we have k assets and T time periods, a factor model may be written

$$r_{it} = \alpha_i + \beta_{i1} f_{1t} + \cdots + \beta_{im} f_{mt} + \varepsilon_{it}$$

for $t = 1, \ldots, T$ and $i = 1, \ldots, k$, where α_i may be considered an intercept, $f_{1t}, \ldots, f_{mt}$ are factors common to all asset returns with loadings $\beta_{i1}, \ldots, \beta_{im}$ for asset i and ε_{it} is the asset specific factor

Properties and assumptions

- Writing $\boldsymbol{f_t} = (f_{1t}, \dots, f_{mt})'$, $\{\boldsymbol{f_t}\}$ is assumed to be a weakly stationary process with

$$E(\boldsymbol{f_t}) = \boldsymbol{\mu_f} \text{ and } Cov(\boldsymbol{f_t}) = \boldsymbol{\Sigma_f}$$

- We also assume that

$E(\varepsilon_{it}) = 0$ for all i and t, $Cov\left(f_{jt}, \varepsilon_{is}\right) = 0$ for all j, t, i and s and $Cov\left(\varepsilon_{jt}, \varepsilon_{is}\right) = \sigma^2{}_i$ if $j = i$ and $s = t$ and $Cov\left(\varepsilon_{jt}, \varepsilon_{is}\right) = 0$ otherwise

Properties

- Using vector notation the model may be written

$$\boldsymbol{r}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\boldsymbol{f}_t + \boldsymbol{\varepsilon}_t$$

from which it follows that

$$Cov(\boldsymbol{r}_t) = \boldsymbol{\beta}\boldsymbol{\Sigma}_f\boldsymbol{\beta}' + \boldsymbol{D}$$

where $\boldsymbol{D}$ is a diagonal matrix with elements $\sigma^2{}_1, \dots, \sigma^2{}_k$

Properties

- We may treat the factor model as a time series and write

$$\boldsymbol{R}_i = \alpha_i \mathbf{1}_T + \boldsymbol{F}\boldsymbol{\beta}'_i + \boldsymbol{E}_i$$

where $\boldsymbol{R}_i = (r_{i1}, \dots, r_{iT})'$, $\mathbf{1}_T$ is a T-dimensional vector of ones, $\boldsymbol{F}$ is a $T \times m$ matrix whose tth row is $\boldsymbol{f}'_t$ and $\boldsymbol{E}_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})'$

- Writing $\boldsymbol{\xi} = [\boldsymbol{\alpha}, \boldsymbol{\beta}]$ and $\boldsymbol{g}_t = (\mathbf{1}, \boldsymbol{f}'_t)'$ we may write

$$\boldsymbol{r}_t = \boldsymbol{\xi}\boldsymbol{g}_t + \boldsymbol{\varepsilon}_t$$

Properties

- Stacking all data we may write

$$\boldsymbol{R} = \boldsymbol{G}\boldsymbol{\xi}' + \boldsymbol{E}$$

where $\boldsymbol{R}$ is a $T \times k$ matrix whose tth row is $\boldsymbol{r}'_t$ or equivalently whose ith column is $\boldsymbol{R}_i$, $\boldsymbol{G}$ is a $T \times (m+1)$ matrix whose tth row is $\boldsymbol{g}'_t$ and $\boldsymbol{E}$ is a $T \times k$ matrix whose tth row is $\boldsymbol{\varepsilon}'_t$

Estimation if the factors are observable

- Least squares method gives that

$$\widehat{\boldsymbol{\xi}}' = (\boldsymbol{G}'\boldsymbol{G})^{-1}(\boldsymbol{G}'\boldsymbol{R})$$

- Residuals are given by

$$\widehat{\boldsymbol{E}} = \boldsymbol{R} - \boldsymbol{G}\widehat{\boldsymbol{\xi}}'$$

- The covariance of $\boldsymbol{\varepsilon}_t$ is estimated by

$$\widehat{\boldsymbol{D}} = \widehat{\boldsymbol{E}}'\widehat{\boldsymbol{E}}/(T - m - 1)$$

- The coefficients of determination for each asset are given by

$$R_i{}^2 = 1 - \frac{\left[\widehat{\boldsymbol{E}}'\widehat{\boldsymbol{E}}\right]_{i,i}}{[\boldsymbol{R}'\boldsymbol{R}]_{i,i}}$$

Example, market model

- Introduced by Sharpe (1970) and given by

$$r_{it} = \alpha_i + \beta_i r_{mt} + \varepsilon_{it}$$

where r_{it} is the excess return of stock i and r_{mt} is the excess return of the market, where we by excess mean relative to a risk-free investment

- So way may choose some stocks $i = 1, \dots, k$ and use some plausible index to describe the market

- To get the excess returns we may use 3-month treasury bill rates or something similar

Example

- We choose to use the same monthly observations (Jan ’90-Dec ’03 stocks as Tsay uses on page 471 in the 3rd ed)

Tick	Company	Tick	Company
AA	Alcoa	KMB	Kimberley-Clark
AGE	A.G. Edwards	MEL	Mellon Financial
CAT	Caterpillar	NYT	New York Times
F	Ford Motor	PG	Procter & Gamble
FDX	FedEx	TRB	Chicago Tribune
GM	General Motors	TXN	Texas Instruments
HPQ	Hewlett-Packard	SP5	S&P 500 Index

Example

- Of special interest are of course ”the betas”, ”the sigmas” and ”the R-squares”

Tick	$\hat{\beta}$	$\hat{\sigma}$	R^2
AA	1.2916	7.6941	0.3556
AGE	1.5141	7.8075	0.4252
CAT	0.9407	7.7245	0.2340
F	1.2192	8.2408	0.2991
FDX	0.8051	8.8539	0.1472
GM	1.0457	8.1301	0.2414
HPQ	1.6280	9.4693	0.3665
KMB	0.5498	6.0701	0.1462
MEL	1.1229	6.1200	0.4063
NYT	0.7706	6.5904	0.2146
PG	0.4688	6.4589	0.1134
TRB	0.7179	7.2151	0.1697
TXN	1.7964	11.4740	0.3330

Example

- For a good model fit the off-diagonal elements of the estimated residual correlation matrix should be small

Tick	AA	AGE	CAT	F	FDX	GM	HPQ	KMB	MEL	NYT	PG	TRB	TXN
AA	1,00	-0,13	0,45	0,22	0,00	0,14	0,24	0,16	-0,02	0,13	-0,15	0,12	0,19
AGE	-0,13	1,00	-0,03	-0,01	0,14	-0,09	-0,13	0,05	0,06	0,10	-0,02	-0,02	-0,17
CAT	0,45	-0,03	1,00	0,23	0,05	0,15	-0,07	0,18	0,09	0,07	-0,01	0,25	0,09
F	0,22	-0,01	0,23	1,00	0,07	0,48	0,00	0,05	0,10	0,19	-0,07	0,15	-0,02
FDX	0,00	0,14	0,05	0,07	1,00	0,02	0,08	0,13	-0,01	0,10	-0,02	0,18	0,00
GM	0,14	-0,09	0,15	0,48	0,02	1,00	0,01	0,10	0,09	-0,05	-0,01	0,11	0,09
HPQ	0,24	-0,13	-0,07	0,00	0,08	0,01	1,00	-0,18	-0,06	0,09	-0,13	-0,02	0,33
KMB	0,16	0,05	0,18	0,05	0,13	0,10	-0,18	1,00	0,17	0,09	0,26	0,14	-0,18
MEL	-0,02	0,06	0,09	0,10	-0,01	0,09	-0,06	0,17	1,00	-0,04	0,25	0,08	-0,01
NYT	0,13	0,10	0,07	0,19	0,10	-0,05	0,09	0,09	-0,04	1,00	0,12	0,39	-0,05
PG	-0,15	-0,02	-0,01	-0,07	-0,02	-0,01	-0,13	0,26	0,25	0,12	1,00	0,25	-0,04
TRB	0,12	-0,02	0,25	0,15	0,18	0,11	-0,02	0,14	0,08	0,39	0,25	1,00	-0,08
TXN	0,19	-0,17	0,09	-0,02	0,00	0,09	0,33	-0,18	-0,01	-0,05	-0,04	-0,08	1,00

Example; minimum variance portfolio

- For a collection of assets, like the ones used in the above example, we may construct a a portfolio with minimum variance/risk by choosing optimal weights

- If we let Σ be the covariance of the excess returns the optimal wieghts, ω, are given by

$$\omega = \frac{\Sigma^{-1}1}{1'\Sigma 1}$$

Example; minimum variance portfolio

- We could also use the covariance for the excess returns predicted by the fitted model

- The matrix to be used is given by

$$\widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\beta}}'Var(r_{mt}) + diag\left(\widehat{\boldsymbol{D}}\right)$$

Example; minimum variance portfolio

- For the assets in the above example the optimal weights are

Ticker	Raw excess return weights	Fitted model weights
AA	-0.0073	0.0117
AGE	-0.0085	-0.0306
CAT	0.0866	0.0792
F	-0.0232	0.0225
FDX	0.0943	0.0802
GM	0.0916	0.0533
HPQ	0.0344	-0.0354
KMB	0.2296	0.2503
MEL	0.0495	0.0703
NYT	0.1790	0.1539
PG	0.2651	0.2434
TRB	0.0168	0.1400
TXN	-0.0080	-0.0388

More factors

- As seen above in the market model there were som low coefficients of determination and some high correlations in the residuals

- For a better model we may try to incorporate other macro economic factors such as GDP, CPI, unemployment, etc, as described in section 9.2.2 of Tsay 3rd ed

Fundamental Factor Models

- Models that use observable firm specific factors such as industrial classification, market cap, book to market value or… are referred to as fundamental factor models

- The most famous such models are the BARRA model and the Fama-French model (For FF see p.482-3 of Tsay 3rd ed)

BARRA model

- The BARRA model is given by

$$\tilde{\boldsymbol{r}}_t = \boldsymbol{\beta} \boldsymbol{f}_t + \boldsymbol{\varepsilon}_t$$

where $\tilde{\boldsymbol{r}}_t$ is the vector of mean corrected excess returns

- Here $\boldsymbol{\beta}$ denotes the firm specific factors and $\boldsymbol{f}_t$ gives the weights for the factors or factor realizations

BARRA model

- To fit the model we use weighted least squares (WLS) and

$$\hat{\boldsymbol{f}}_t = (\boldsymbol{\beta}'\boldsymbol{D}^{-1}\boldsymbol{\beta})^{-1}(\boldsymbol{\beta}'\boldsymbol{D}^{-1}\tilde{\boldsymbol{r}}_t)$$

- However, $\boldsymbol{D}$ is unknown in practice so we need a two-step approach

BARRA model

- In the first step, for each t, we use a preliminary OLS estimate

$$\hat{\boldsymbol{f}}_{t,o} = (\boldsymbol{\beta}'\boldsymbol{\beta})^{-1}(\boldsymbol{\beta}'\tilde{\boldsymbol{r}}_t)$$

- The residuals are given by

$$\boldsymbol{e}_{t,o} = \tilde{\boldsymbol{r}}_t - \boldsymbol{\beta}\hat{\boldsymbol{f}}_{t,o}$$

BARRA model

- From the residuals we get and estimate of $\boldsymbol{D}$

$$\widehat{\boldsymbol{D}}_o = diag\left\{\frac{1}{T-1}\sum_{t=1}^{T}\boldsymbol{e}_{t,o}\boldsymbol{e}'_{t,o}\right\}$$

- The second step is to use $\widehat{\boldsymbol{D}}_o$ to get

$$\widehat{\boldsymbol{f}}_{t,g} = \left(\boldsymbol{\beta}'\widehat{\boldsymbol{D}}_o^{-1}\boldsymbol{\beta}\right)^{-1}\left(\boldsymbol{\beta}'\widehat{\boldsymbol{D}}_o^{-1}\tilde{\boldsymbol{r}}_t\right)$$

BARRA model

- The residuals are given by

$$\boldsymbol{e}_{t,g} = \tilde{\boldsymbol{r}}_t - \boldsymbol{\beta}\widehat{\boldsymbol{f}}_{t,g}$$

- In turn we get the residual covariance matrix as

$$\widehat{\boldsymbol{D}}_g = diag\left\{\frac{1}{T-1}\sum_{t=1}^{T}\boldsymbol{e}_{t,g}\boldsymbol{e}'_{t,g}\right\}$$

BARRA model

- We get the estimated covariance for the factor realizations from

$$\widehat{\boldsymbol{\Sigma}}_f = \frac{1}{T-1}\sum_{t=1}^{T}(\widehat{\boldsymbol{f}}_{t,g} - \widehat{\boldsymbol{f}}_g)(\widehat{\boldsymbol{f}}_{t,g} - \widehat{\boldsymbol{f}}_g)'$$

where $\widehat{\boldsymbol{f}}_g = \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{f}}_{t,g}$

- And the covariance of the excess returns for the fitted BARRA model is

$$Cov(\boldsymbol{r}_t) = \boldsymbol{\beta}\widehat{\boldsymbol{\Sigma}}_f\boldsymbol{\beta}' + \widehat{\boldsymbol{D}}_g$$

Example BARRA, Industry factor

- We use industry class as the specific factor for the ten assets; A.G. Edwards (AGE), Citigroup (C), Morgan Stanley (MWD), Merrill Lynch (MER), Dell Inc. (DELL), Hewlett-Packard (HPQ), Int. Bus. Machines (IBM), Alcoa (AA), Caterpillar (CAT) and Procter and Gamble (PG)
- AGE, C, MDW and MER belong to the financial services class
- DELL, HPQ and IBM belong to the computer class
- AA, CAT and PG belong to the high-tech industry class
- Data is collected monthly from Jan ’90 to Dec ’03

Example BARRA, Industry factor

- For each of the assets $i = 1, \ldots, 10$ the model is given by

$$\tilde{r}_{it} = \beta_{i1} f_{1t} + \beta_{i2} f_{2t} + \beta_{i3} f_{3t} + \varepsilon_{it}$$

where $\beta_{ij} = 1$ if an assets belong to the industry class j and $\beta_{ij} = 0$ otherwise

- So for, e.g., IBM we have $\boldsymbol{\beta}_i = (0,1,0)'$ and for Alcoa we have $\boldsymbol{\beta}_i = (0,0,1)'$

Example BARRA, Industry factor

- Since ”the betas” are indicator variables, we get

$$\hat{\boldsymbol{f}}_{t,o} = \begin{bmatrix} \dfrac{AGE_t + C_t + MDW_t + MER_t}{4} \\ \dfrac{DELL_t + HPQ_t + IBM_t}{3} \\ \dfrac{AA_t + CAT_t + PG_t}{3} \end{bmatrix}$$

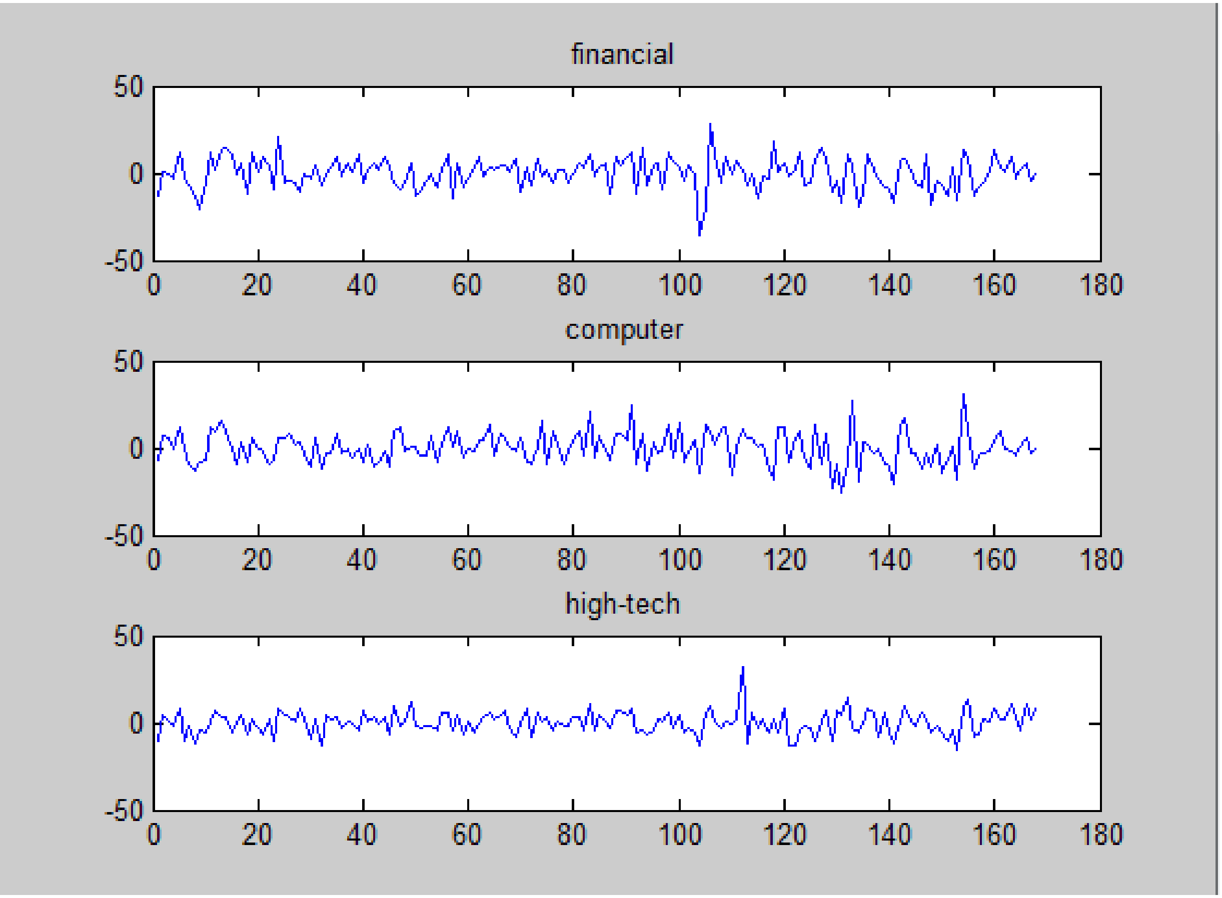
Example BARRA, Industry factor

- Proceeding in the above two-step fashion we get

$$\hat{\boldsymbol{f}}_{t,g} = \begin{bmatrix} 0.19AGE_t + 0.25C_t + 0.26MDW_t + 0.30MER_t \\ 0.23DELL_t + 0.40HPQ_t + 0.37IBM_t \\ 0.33AA_t + 0.43CAT_t + 0.24PG_t \end{bmatrix}$$

Example BARRA, Industry factor

- The estimated factor realistations:



PCA

- Principal component analysis is concerned with finding a few factors, referred to as principal components, that describe most of the variation in a high-dimensional vector or matrix of observations or random variables

- In financial applications it may be used to simplify the interpretation of the covariance structure in a portfolio

PCA

- PCA applies either to the covariance or the correlation matrix
- For a random vector $\boldsymbol{r} = (r_1, \dots, r_k)$, and some wieights $\boldsymbol{w}_i = (w_{i1}, \dots, w_{ik})'$ we may define linear combinations

$$y_i = \boldsymbol{w}'_i \boldsymbol{r} = \sum_{j=1}^{k} w_{ij} r_j$$

- We are only interested in the how the weights are proportionally allocated and will assume that the squared weights sum to one

PCA

- If $Cov(\boldsymbol{r}) = \boldsymbol{\Sigma}$ we have

$$Var(y_i) = \boldsymbol{w}'_i \boldsymbol{\Sigma} \boldsymbol{w}_i$$

$$Cov(y_i, y_j) = \boldsymbol{w}'_i \boldsymbol{\Sigma} \boldsymbol{w}_j$$

- The idea of PCA is to find weights so that the variances $Var(y_i)$ are as large as possible while $Cov(y_i, y_j) = 0$ for $i \neq j$

PCA

- The first principal component is found by maximizing $Var(y_1)$ while $\boldsymbol{w'}_1\boldsymbol{w}_1 = 1$

- The second principal component is found by maximizing $Var(y_2)$ while $\boldsymbol{w'}_2\boldsymbol{w}_2 = 1$ and $Cov(y_1, y_2) = 0$

- The ith principal component is found by maximizing $Var(y_i)$ while $\boldsymbol{w'}_i\boldsymbol{w_i} = 1$ and $Cov\left(y_j, y_i\right) = 0$ for $j = i - 1, i - 2, \ldots, 1$

PCA

- Since covariance matrix $\boldsymbol{\Sigma}$ is non-negative definite it has an eigenvalue/eigenvector decomposition $(\lambda_1, \boldsymbol{e}_1), \ldots, (\lambda_k, \boldsymbol{e}_k)$

- It can be shown that $y_i = \boldsymbol{e}'_i \boldsymbol{r}$ and that

$$Var(y_i) = \boldsymbol{e}'_i \boldsymbol{\Sigma} \boldsymbol{e}_i = \lambda_i$$

$$Cov(y_i, y_j) = \boldsymbol{e}'_i \boldsymbol{\Sigma} \boldsymbol{e}_j = 0$$

- Note that the proportion of the total variance of $\boldsymbol{r}$ explained by y_i is

$$\frac{\lambda_i}{\lambda_1 + \cdots + \lambda_k}$$

Application of PCA

- In practice the covariance and correlation matrices of the return vector $\boldsymbol{r}_t$ have to be estimated using

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{T-1}\sum_{t=1}^{T}(\boldsymbol{r}_t - \bar{\boldsymbol{r}})(\boldsymbol{r}_t - \bar{\boldsymbol{r}})'$$

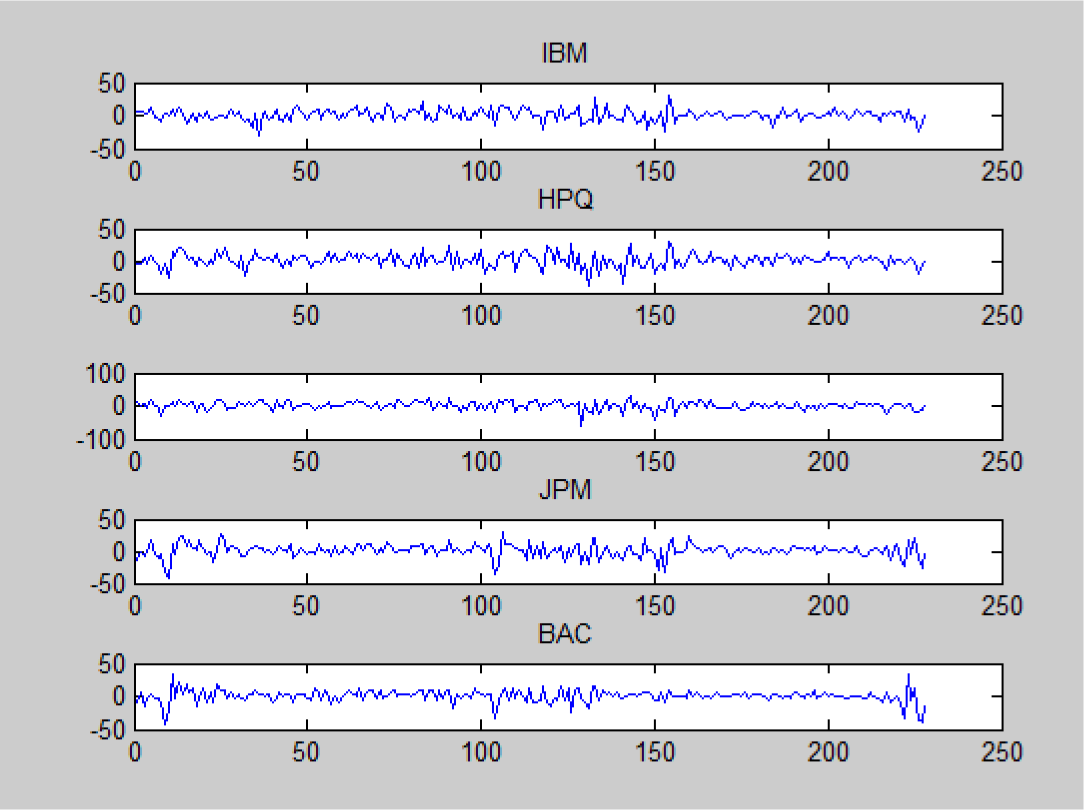
where $\bar{\boldsymbol{r}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{r}_t$ and

$$\widehat{\boldsymbol{\rho}} = \widehat{\boldsymbol{S}}^{-1}\widehat{\boldsymbol{\Sigma}}\widehat{\boldsymbol{S}}^{-1}$$

where $\widehat{\boldsymbol{S}}$ is the diagonal matrix of standard deviations of $\boldsymbol{r}_t$

Application of PCA

- We use monthly log returns of IBM, HPQ, INTC, JPM and BAC fron jan '90 to Dec '08



Application of PCA

- We use monthly log returns of IBM, HPQ, INTC, JPM and BAC fron jan ’90 to Dec ’08

- The sample covaariance is

	IBM	HPQ	INTC	JPM	BAC
IBM	74,64	42,28	48,03	30,10	21,07
HPQ	42,28	112,22	70,45	42,42	26,30
INTC	48,03	70,45	146,50	44,59	29,24
JPM	30,10	42,42	44,59	106,04	67,45
BAC	21,07	26,30	29,24	67,45	91,83

Application of PCA

- The eigenvalues, propotions of total variance and eigenvectors are

Eigenval	284.17	112.93	57.44	46.81	29.87
Prop	0.535	0.213	0.108	0.088	0.056
Cumul	0.535	0.748	0.856	0.944	1.000
eigenvec	0.330	-0.139	-0.264	0.895	0.014
	0.483	-0.279	-0.701	-0.430	0.116
	0.581	-0.478	0.652	-0.096	0.016
	0.448	0.550	0.013	-0.064	-0.702
	0.347	0.610	0.119	-0.009	0.702

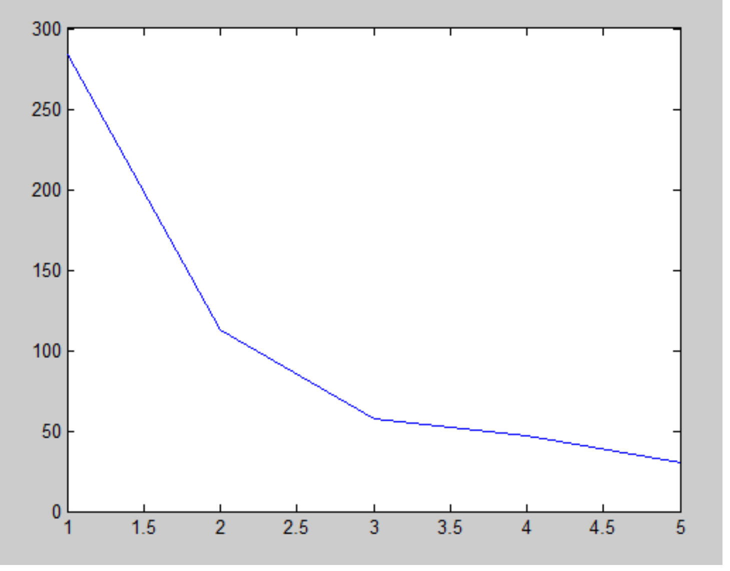
Application of PCA

- We see that the first two eigenvalues account for 75% of the variance
- The intepretation is that the first component is a market component, describing the overall behavoiur of the market
- The second component accounts for differences between the assets, which in this case may be an industrial factor (note the minus signs for ”IT” and plus signs for ”Bank”)

Scree plot

- A simple but effective graphical tool used to determine the number of relevant components
- Just a plot of the eigenvalues sorted from large to small
- We look for an ”elbow”

to the right of which the eigenvalues are small and approximately equal in size



Statistical Factor Analysis

- Here ”the setup” is (again)

$$\tilde{\boldsymbol{r}}_t = \boldsymbol{\beta} \boldsymbol{f}_t + \boldsymbol{\varepsilon}_t$$

where $\tilde{\boldsymbol{r}}_t$ is the mean corrected $\boldsymbol{r}_t$ and $\boldsymbol{\beta}$ is the $k \times m$ matrix of factor loadings

- The process defined by $\boldsymbol{r}_t$ is assumed to be weakly stationary

Statistical Factor Analysis

- The big difference compared to previously covered models is that the factors $\boldsymbol{f}_t$ are unobservable random variables

- We assume that

$$E(\boldsymbol{f}_t) = 0 \text{ and } Cov(\boldsymbol{f}_t) = \boldsymbol{I}_m$$
$$E(\boldsymbol{\varepsilon}_t) = 0 \text{ and } Cov(\boldsymbol{\varepsilon}_t) = \boldsymbol{D} = diag({\sigma^2}_1, \dots, {\sigma^2}_k)$$
$$\boldsymbol{f}_t \text{ and } \boldsymbol{\varepsilon}_t \text{ are independent so that } Cov(\boldsymbol{f}_t, \boldsymbol{\varepsilon}_t) = \mathbf{0}_{m\times k}$$

Statistical Factor Analysis

- It is an exercise to show that

$$\Sigma_{\boldsymbol{r}} = \boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{D}$$

and

$$Cov(\boldsymbol{f}_t, \boldsymbol{r}_t) = \boldsymbol{\beta}$$

- We then have that

$$Var(r_{it}) = \beta^2{}_{i1} + \cdots + \beta^2{}_{im} + \sigma^2{}_i$$

and

$$Cov(r_{it}, r_{jt}) = \beta_{i1}\beta_{j1} + \cdots + \beta_{im}\beta_{jm}$$

- The variance $c^2{}_i = \beta^2{}_{i1} + \cdots + \beta^2{}_{im}$ accounted for by the common factors is called the communality and the remaining variance $\sigma^2{}_i$ is called unique or (asset) specific

Statistical Factor Analysis

- In practice the covariance matrix (and hence $\boldsymbol{r}_t$) need not have an othogonal factor decomposition (representation)

- Also, if the representation exists it is not unique, since for a matrix $\boldsymbol{P}$ such that $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{I}$ we may let $\boldsymbol{\beta}^{\,*} = \boldsymbol{\beta}\boldsymbol{P}$ and $\boldsymbol{f}^{*}{}_{t} = \boldsymbol{P}'\boldsymbol{f}_t$ so that

$$\tilde{\boldsymbol{r}}_t = \boldsymbol{\beta}\boldsymbol{f}_t + \boldsymbol{\varepsilon}_t = \boldsymbol{\beta}\boldsymbol{P}\boldsymbol{P}'\boldsymbol{f}_t + \boldsymbol{\varepsilon}_t = \boldsymbol{\beta}^{\,*}\boldsymbol{f}^{*}{}_{t} + \boldsymbol{\varepsilon}_t$$

- The non-uniqueness is not necessarily a disadvantage since it lets us rotate factors which may facilitate intepretation

Estimation

- We may use the above described PCA for the coavariance (or correlation) matrix $\Sigma_{\boldsymbol{r}}$ for which the matrix of factor loadings is given by

$$\widehat{\boldsymbol{\beta}} = \left[\sqrt{\lambda_1}\,\hat{\boldsymbol{e}}_1 \cdots \sqrt{\lambda_m}\,\hat{\boldsymbol{e}}_m\right]$$

where $m < k$ and the eigenvalues are sorted in descending order of size

- The estimated asset specific variances are the diagonal elements of $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{r}} - \widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\beta}}'$ i.e. $\widehat{\boldsymbol{D}} = diag\{\hat{\sigma}^2{}_1, \dots, \hat{\sigma}^2{}_k\}$

- The estimated communalities are

$$\hat{c}^2{}_i = \hat{\beta}^2{}_{i1} + \cdots + \hat{\beta}^2{}_{im}$$

- The estimation error matrix is

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{r}} - \left(\widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\beta}}' - \widehat{\boldsymbol{D}}\right)$$

Estimation

- We may also use ML under normality assumptions on $\boldsymbol{f}_t$ and $\boldsymbol{\varepsilon}_t$

- Using ML is briefly described on p.491-2 of Tsay 3rd ed

- Matlab ”factoran” uses ML as does ”factanal” in R

Factor Rotation

- As mentioned above factor rotations may help in interpreting the factor
- There are many different rotations in the literature
- We will focus on varimax (see p.492 of Tsay 3rd ed for criterion) for which the idea is to spread the squares of the factor loadings as much as possible

Example

- Using two factors in matlab’s ”factoran” for the IBM, HPQ, INTC, JPM, BAC-data, we get

Variable	f_1	f_2	f_1 rot	f_2 rot	communalities
IBM	0.327	0.530	0.594	0.187	0.388
HPQ	0.349	0.668	0.733	0.174	0.568
INTC	0.337	0.647	0.709	0.168	0.532
JPM	0.734	0.186	0.361	0.666	0.573
BAC	0.960	-0.111	0.127	0.967	0.934
Variance	1.801	1.192	1.540	1.454	2.994
Proportion	0.360	0.239	0.308	0.291	0.599

Interpretation

- So two factors account for 60% of the variance in the returns of five stocks
- The interpretations of the rotated factor loadings are that the ”IT” stocks load the first factor heavily while the ”Bank” stocks load the second factor heavily
- So the varimax rotation separates the industrial sectors
- The relatively low communality of IBM indicates that IBM has some ”own” special features which we may want to investigate