Factor Analysis for Stock Performance

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHINIC INSTITUTE

In partial fulfillment of the requirements for the

Professional Masters Degree in

Financial Mathematics

By

Wei Cheng

May 2, 2005

Approved:

Professor Dalin Tang, Thesis Advisor

Professor Eugene Yablonski, Co-Advisor

Professor Bogdan Vernescu, Department Head

Abstract

Factor models are very useful and popular models in finance. In this project, factor models are used to examine hidden patterns of relationships for a set of stocks. We calculate the weekly rates of return and analyze the correlation among those variables. We propose to use Principal Factor Analysis (PFA) and Maximum-likelihood Factor Analysis (MLFA) as a data mining tool to recover the hidden factors and the corresponding sensitivities. Prior to applying PFA and MLFA, we use the Scree Test and the Proportion of Variance Method for determining the optimal number of common factors. Then, rotation for PFA and MLFA were performed to improve the first order approximations. PFA and MLFA were used to extract three underlying factors. It was determined that the MLFA provided a more accurate estimation for weekly rates of return.

Acknowledgements

First and foremost, I would like to sincerely thank Professor Dalin Tang for his great help, patience, and continuous guidance and his ability to motivate me.

I would also like to gratefully acknowledge Professor Eugene Yablonski for his guidance, enthusiasm and encouragement in this project, especially in my life.

I also want to thank all Mathematical Sciences Department. all the faculty and stuff who taught me and helped me during my MS's study.

Special thanks to Ms. Ann Shi for her endless support and also helping get this project up and running.

Finally, I would also like to recognize my boyfriend, Yuefeng Kang, for being there for his understanding, support in helping me throughout this project.

At last, I would like to thank God, family, and friends for their love and never-ending support.

Table of Contents

1. Introduction

A major objective of scientific activities is to summarize, by theoretical formulations, the empirical relationships among a given set of events and discover natural laws that are hidden in hundreds and thousands of random events. It could be stated that scientists analyze the relationships among a set of variables, while these relationships are evaluated across a set of individuals under specified conditions. The variables are the characteristic being measured and could be anything that can be objectively identified or scored.

There are mainly reasons that make a factor model attractive for portfolio management. First, a factor model dramatically reduces the work for variable estimate. Secondly, it gives a clearer picture of the major source of the portfolio risk. A Factor model relates the returns of securities to a set of factors. The factors can be system (market) factors or non-system (individual) factors [13]. Finding the factors for the model is not an easy task to researchers, as the factors are hidden and not necessary directly related to the fundamental factors, such as GDP or interest rates. Some companies have developed a few factor models to predict interest rates and credit spreads based on the market technical and economic data. In fact, there is a lot of research on predicting stock market returns using such factors as momentum, size, style, and other factors.

1.1 Objective of the Project

The general purpose of factor analytic techniques is to find a way of condensing the information contained in a number of original variables into smaller set of new composite factors with a minimum loss of information. That is, the objective is to summarize the interrelationships among the variables in a concise but accurate manner as an aid in conceptualization. Parsimoniously describing data, factor analysts explicitly recognize that any relationship is limited to a particular area of applicability. Areas qualitatively different, that is, areas where relatively little generalization can be made from one area to another, are referred to as separate factors. Each factor represents an area of generalization that is qualitatively distinct from that represented by any other factor. Within an area where data can be summarized, factor analysts first represent that area by a factor and then seek to make the degree of generalization between each variable and the factor explicit.

In this project, we describe factor analysis and use this technique to examine the underlying patterns of relationships for eight stocks, and determine if the information can be condensed or summarized in a smaller set of factors or components.

1.2 Overview of Models and Methods

Factor model is a fundamental model in finance. Many theories are established based on it, for examples, Modern Portfolio Theory and Arbitrage Pricing Theory. Factor model serves as an efficient and common model for the return generating process. Furthermore, factor model is also the foundation of Arbitrage Pricing Theory. Arbitrage Pricing Theory plays an important role in modern finance and it analyzes the capital asset pricing in finance [11].

For the estimation of factor models, there are many methods available: Principal Component Method, Principal Factor Method, Iterated Principal Factor Method and Maximum-likelihood Method. In this paper, we apply Principal Factor Analysis and Maximum- likelihood Factor Analysis, and the modern signal processing methods to recover the hidden factors and the corresponding sensitivities.

1.3 Outline of the Project

In this project, we apply Principal Factor Analysis and Maximum-likelihood Factor Analysis to construct the underlying factors and obtain the corresponding sensitivities for a financial factor model. Chapter 2 reviews the backgrounds of factor model and also analyzes the Principal Factor Analysis and Maximum-likelihood Factor Model. In Chapter 3, we use two mentioned techniques to examine the underlying pattern of relationships for eight stocks (IBM, Dell, Apple, Sony, Novell, Microsoft, AMD, Intel) listed on the New York Stock Exchange and NASDA for the period from January 1998 through December 2004. Specifically, we calculate the weekly rates of return and analyze the correlation among those variables. Furthermore, we extract a set of hidden factors and analyze the sensitivity of the weekly return rates to those factors. Finally, it turned out that the Maximum-likelihood Factor Model provided a more accurate estimation for weekly rates of return.

2. Factor Analysis

Factor analysis is a class of multivariate statistical methods whose primary purpose is data reduction and summarization. Broadly speaking, it addresses the problem of analyzing the interrelationships among a large number of variables and then explaining these variables in terms of their common, underlying factors [18].

Factor Analysis has provoked rather turbulent controversy throughout its history. The basic ideas of Factor Analysis were suggested around the turn of the century by Francis Galton and Charles Spearman among others, and originated mainly from the efforts of psychologists to gain a better understanding of 'intelligence'. Intelligence tests customarily contain a large variety of questions that depend to a greater or lesser extent on verbal ability, mathematical ability, memory, etc. Because of this early associated with constructs such as intelligence, Factor Analysis was developed to analyze these test scores so as to determine if 'intelligence' is made up of a single underlying general factor or of several more limited factors measuring attributes like 'mathematical ability'.

In Multivariate Analysis, one often has data or a large number of variables $v_1, v_2 ... v_p$, and it's tempting to think that there is a reduced list of underlying factors that determine the full set. A successful factor analysis is usually one in which different factors influence or "load onto" disjoint subsets of variables.

Factor Analysis aims to describe the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities called factors. Basically, the factor model is motivated by the following argument: Suppose variables can be grouped by their correlations. That is, suppose all variables within a particular group are highly correlated among themselves, but have relatively small correlations with variables in a different group. Then it is conceivable that each group of variables represents a single underlying construct or factor that is responsible for the observed correlations.

2.1 What is the Main Problem of Factor Analysis

Factor Analysis is a very powerful mathematical tool which can be used to examine a wide range of data sets. Most applications of Factor Analysis have been in psychology, chemistry, sociology, social sciences and economics. A statistical procedure that gives both qualitative and quantitative distinctions can be quite useful. Some of the purposes for which factor analysis can be used are as follows:

1. Through factor-analytic techniques, the number of variables for further research can be minimized while also maximizing the amount of information in the analysis. The original set of variables is reduced to a much smaller set that accounts for most of the reliable variance of the initial variable pool. The smaller set of variables can be used as operational representatives of the constructs underlying the complete set of variables.
2. Factor analysis can be used to search data for possible qualitative and quantitative distinctions, and is particularly useful when the sheer amount of available data

exceeds comprehensibility. Out of this exploratory work can arise new constructs and hypotheses for further theory and research. The contribution of exploratory research to science is, of course, completely dependent upon adequately pursuing the results in further research studies so as to confirm or reject the hypotheses developed.

3. If a domain of data can be hypothesized to have certain qualitative and quantitative distinctions, then this hypothesis can be tested by factor analysis. If the hypotheses are tenable, the various factors will represent the theoretically derived qualitative distinctions. If one variable is hypothesized to be more related to one factor than another, this quantitative distinction can also be checked.

We have known the essential purpose of factor analysis is to describe the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities. Therefore, the primary question in factor analysis is whether the data are consistent with a prescribed structure. Suppose variables can be grouped by their correlations. That is, suppose all variables within a particular group are highly correlated among themselves, but have relatively small correlations with variables in a different group. That means each group of variables represents a single underlying construct, or factor. In portfolio management, one key difficulty is to estimate the covariance matrix. Factor model provides a good way to estimate it because it involves less variable, can relate the model variables with fundamental analysis (not for all factor models).

2.2 Basic Factor Model

2.2.1 Model Definition and Assumptions

In factor analysis we represent observed variables $v_1, v_2, ..., v_p$ as linear combinations of a small set of random variables $f_1, f_2, ..., f_m (m < p)$ called factors. The factors are underlying constructs or latent variables that "generate" the v's. Like the original variables, the factors vary from individual to individual; but unlike the variables, the factors cannot be measured or observed. If the original variables $v_1, v_2, ..., v_p$ are at least moderately correlated, the basic dimensionality of the system is less than p. The goal of factor analysis is to reduce the redundancy among the variables by using a smaller number of factors. Suppose the pattern of the high and low correlations in the correlation matrix are such that the variables in a particular subset have high correlations among themselves but low correlations with all the other variables. Then there may be single underlying factor that give rise to the variables in the subset. If the other variables can be similarly grouped into subsets with a like pattern of correlations, then a few factors can represent these groups of variables.

Suppose we make observations on p variables, $v_1, v_2, ..., v_p$, which have mean vectors μ and covariance matrix Σ. The Factor Analysis model expresses each variable as a linear combination of underlying common factors $f_1, f_2, ..., f_m$, with an accompanying residual term to account for that part of the variable that is unique. For $f_1, f_2, ..., f_p$ in any observation vector y, the model is as follows:

$$\begin{aligned} v_1 - \mu_1 &= \lambda_{11} f_1 + \lambda_{12} f_2 + \cdots + \lambda_{1m} f_m + e_1 \\ v_2 - \mu_2 &= \lambda_{21} f_1 + \lambda_{22} f_2 + \cdots + \lambda_{2m} f_m + e_2 \\ &\vdots \\ v_p - \mu_p &= \lambda_{p1} f_1 + \lambda_{p2} f_2 + \cdots + \lambda_{pm} f_m + e_p \end{aligned} \tag{2.2.1}$$

The number of factors m should be substantially smaller than p, otherwise, we don't achieve a parsimonious description of the variables as functions of a few underlying factors. The coefficient λ_{ij} is the weights usually called the factor loading, so that λ_{ij} is the loading of the ith variable on the jth factor. With appropriate assumptions, λ_{ij} indicates the importance of the *j*th factor f_i to the *i*th variable v_i and can be used in interpretation of f_j. The variable $\{e_i\}$ describes the residual variation specific to the ith variable. The factors f_j are often called the common factors while the residual variables e_i are often called the specific factors [4, 6, 13, 15].

The Factor Analyst usually makes a number of assumptions about model (2.2.1). The specific factors are assumed to be independent of one another ($\text{cov}(e_i, e_k) = 0, i \neq k$) and of the common factors ($\text{cov}(e_i, f_j) = 0$ for all i and j). It is also usually assumed that the common factors are independent of one another ($\text{cov}(f_j, f_k) = 0, j \neq k$), though this assumption is sometimes relaxed when the factors are later rotated. As we have assumed the **v**'s to have zero mean, it is also convenient to assume that the factors all have zero mean ($E(f_j) = 0, j = 1,2,\ldots,m$). We also assume the common factors to have zero mean($E(e_i)=0$). Looking at equation (2.1), we see that there is an arbitrary scale factor related to each common factor and so it is customary to choose the common factors so that each has unit variance ($\mathrm{var}(f_j)=1$). The variances of the specific factors may vary and we denote the variance of e_i by ψ_i ($\mathrm{var}(e_i)=\psi_i$) [4, 13]. It is also customary to assume that the common factors and specific factors each have a multivariate normal distribution. This implies that **v** is also multivariate normal, where $\mathrm{v}^T=\{v_1,v_2,...,v_p\}$. The large number of assumptions that have to be made in setting up the Factor Analysis model is one of the drawbacks to the method. Indeed, Lawley and Maxwell (1971, p. 38) stress that the model is 'useful only as an approximation to reality' and 'should not be taken too seriously'.

The model is usually written in matrix notation as

$$\mathbf{v} - \boldsymbol{\mu} = \mathbf{Lf} + \mathbf{e} \qquad (2.2.2)$$

Where

$$\mathbf{v} = (v_1, v_2, \ldots, v_p)', \qquad (2.2.3)$$

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)', \qquad (2.2.4)$$

$$\mathbf{f} = (f_1, f_2, \ldots, f_m)', \ \mathbf{e} = (e_1, e_2, \ldots, e_p)', \qquad (2.2.5)$$

and

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{pmatrix} \qquad (2.2.6)$$

We assume that $E(e_i) = 0$, $\mathrm{var}(e_i) = \psi_i$, and $\mathrm{cov}(e_i, e_k) = 0$, $i \neq k$. The assumption $\mathrm{cov}(e_i, e_k) = 0$ implies that the factors account for all the correlations among the **v**'s, that is, all that the **v**'s have in common.

From equation (2.2.2), we have

$$E(\mathbf{f}) = \mathbf{0}, \qquad \mathrm{cov}(\mathbf{f}) = \mathbf{I}, \tag{2.2.7}$$

$$E(\mathbf{e}) = \mathbf{0}, \qquad \mathrm{cov}(\mathbf{e}) = \Psi = \begin{pmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \psi_p \end{pmatrix} \tag{2.2.8}$$

$$\mathrm{cov}(\mathbf{f}, \mathbf{e}) = \mathbf{0} \tag{2.2.9}$$

$$E(\mathbf{x}) = \mu, \qquad \mathrm{cov}(\mathbf{x}) = \mathbf{LL}` + \psi \tag{2.2.10}$$

$$\mathrm{cov}(\mathbf{x}, \mathbf{f}) = \mathbf{L} \tag{2.2.11}$$

$$\sigma_{ii} = \mathrm{var}(y_i) = (\lambda_{i1}^2 + \lambda_{i2}^2 + \cdots + \lambda_{im}^2) + \psi_i$$

$$= h_i^2 + \psi_i = \text{communality} + \text{specific variance} \tag{2.2.12}$$

Where

Communality $= h_i^2 = \lambda_{i1}^2 + \lambda_{i2}^2 + \cdots + \lambda_{im}^2$

Specific variance $= \psi_i$

The communality is also referred to as common variance, and the specific variance has been called specificity, unique variance, or residual variance [4]. One advantage of the factor analysis model is that when it does not fit the data, the estimate of **L** clearly reflects this failure. In such cases, there are two problems in the estimates: (1) it is unclear how many factors there should be, and (2) it is unclear what the factors are.

The Factor Model assumes that the $\frac{p(p+1)}{2}$ parameters of Σ can be represented by p×m variances and the p specific variance. For example, if p is 10, m is 2. Σ has 55, Factor Model has 30. If we can represent Σ with m orthogonal factors, there are many fewer parameters to estimate.

2.2.2 Estimating the Parameters in the Factor Model (Loadings and Communalities)

The parameters of the Factor Analysis model, including the factor loadings and the error variances, are nearly always unknown and need to be estimated from the sample data. The sample covariance matrix is occasionally used, but it is much more common to work with the sample correlation matrix. Then the v variables in equation (2.2.1) and (2.2.2) refer to scaled variables having zero mean and unit variance [7].

In the early days of Factor Analysis, a variety of iterative methods were used to estimate the factor loadings. These involved subjective judgment, such as guessing the communalities, with the result that different researchers could analyze the same data and find entirely different factors. One popular method was the Principal Factor Method. This chooses the first factor so as to account for as much as possible of the communal variance, the second factor to account for as much as possible of the remaining communal variance, and so on. The method requires suitable estimates of the communalities. If they are chosen to be unity, then the method reduces to principal component analysis. Then in

1940, a major step forward was made by D. N. Lawley, who developed the maximum-likelihood equations. These are fairly complicated and difficult to solve, but recent computational advances, particularly by K. G. Jöreskog, have made maximum-likelihood estimation a practical proposition, and computer programs are now becoming widely available [7].

Till now, there are many methods available for estimation of factor models, such as Principal Component, Principal Factor, Maximum-likelihood, Iterated Principal Factor, etc. In this paper, we will discuss two of the most popular methods of parameter estimation, the Principal Factor Method and the Maximum-likelihood Factor Method. The solution from either method can be rotated in order to simplify the interpretation of factors.

2.2.2.1 Principal Factor Model (Principal Axis Model)

Principal Factor Analysis is performed on the reduced covariance matrix $\underset{\sim}{S}^*$, obtained by replacing the observed diagonal elements of $\underset{\sim}{S}$ with estimated communalities. Two frequently used estimates are:

(1) The square of the multiple correlation coefficient of the ith variable with all other variables.

(2) The largest of the absolute values of the correlation coefficients between the ith variable and one of the other variables.

Each of these estimates will give higher communality values when v_i is highly correlated with the other variables, which is what is required [12].

Let covariance matrix Σ have eigenvalue- eigenvector pairs ((λ_i, e_i) with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$. Then the spectral decomposition says [4],

$$\Sigma = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \cdots + \lambda_p e_p e_p' \qquad (2.2.13)$$

$$= \left[\sqrt{\lambda_1} e_1 \vdots \quad \sqrt{\lambda_2} e_2 \vdots \quad \cdots \quad \vdots \sqrt{\lambda_p} e_p \right] \begin{bmatrix} \sqrt{\lambda_1} e_1' \\ \cdots \\ \sqrt{\lambda_2} e_2' \\ \cdots \\ \sqrt{\lambda_3} e_3' \\ \cdots \\ \sqrt{\lambda_4} e_4' \end{bmatrix}$$

(2.2.14)

So if $L = \left[\sqrt{\lambda_1} e_1 \vdots \sqrt{\lambda_2} e_2 \vdots \cdots \vdots \sqrt{\lambda_p} e_p\right]$, then **Σ=LL'**

For the m<<p factor model, $L = \left[\sqrt{\lambda_1} e_1 \vdots \sqrt{\lambda_2} e_2 \vdots \cdots \vdots \sqrt{\lambda_m} e_m\right]$

So the principal factor estimate of K is

$$\tilde{L} = \left[\sqrt{\tilde{\lambda}_1}\tilde{e}_1 \vdots \sqrt{\tilde{\lambda}_2}\tilde{e}_2 \vdots \cdots \vdots \sqrt{\tilde{\lambda}_m}\tilde{e}_m\right] \tag{2.2.15}$$

Where $(\tilde{\lambda}_i, \tilde{e}_i)$ i=1, … , m are the eigenvalue-eigenvector of the sample covariance matrix S the sample covariance matrix.

Allowing for specific factors, we find the approximation becomes

Σ ≈ LL' + ψ

$$= \left[\sqrt{\lambda_1} e_1 \vdots \;\; \sqrt{\lambda_2} e_2 \vdots \;\; \cdots \;\; \vdots \sqrt{\lambda_m} e_m\right] \begin{bmatrix} \sqrt{\lambda_1} e_1' \\ \cdots \\ \sqrt{\lambda_2} e_2' \\ \cdots \\ \sqrt{\lambda_3} e_3' \\ \cdots \\ \sqrt{\lambda_4} e_4' \end{bmatrix} \begin{bmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_p \end{bmatrix} \tag{2.2.16}$$

Where

$$\psi_i = \sigma_{ii} - \sum_{j=1}^{m} l_{ij}^2 \;\; \text{for i=1, 2, … , p.} \tag{2.2.17}$$

And

$$S = \tilde{L}\tilde{L}' + \tilde{\psi} \tag{2.2.18}$$

$$\tilde{\psi} = diag(S - \tilde{L}\tilde{L}') \tag{2.2.19}$$

Estimated commonalities are

$$\tilde{h}_i = \tilde{l}_{i1}^{\,2} + \tilde{l}_{i2}^{\,2} + \cdots + \tilde{l}_{im}^{\,2} \tag{2.2.20}$$

$$\tilde{\psi} = S_{i1} - h_i^2 \tag{2.2.21}$$

Therefore,

$$Var(V_i) = S_{ii} = h_i^2 + \psi_i \tag{2.2.22}$$

2.2.2.2 Maximum-likelihood Factor Model

Maximum-likelihood method was first applied to factor analysis by Lawley (1940, 1941, 1943) but its routine use had to await the development of computers and suitable numerical optimization procedures.[13] This method is the procedure of finding the value of one or more parameters for a given statistic which makes the known likelihood distribution a maximum. A maximum likelihood estimator is a value of the parameter a such that the likelihood function is a maximum (Harris and Stocket 1998, p. 824).

If we assume that our raw data arise from a multivariate normal distribution then we then maximum-likelihood estimates of the factors loadings and specific variances can be obtained. When F_i and e_i are jointly normal, the observations $V_i - \mu = LF_i + e_i$ are then normal, the likelihood is

$$L(\mu, \Sigma) = (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{p}{2}} e^{-\frac{1}{2} tr\left[\Sigma^{-1}\left(\sum_{i=1}^{n}(v_i - \bar{v})(v_i - \bar{v})' + n(\bar{v} - \mu)(\bar{v} - \mu)'\right)\right]}$$

$$\begin{aligned} &= (2\pi)^{-\frac{(n-1)p}{2}} |\Sigma|^{-\frac{(n-1)}{2}} e^{-(\frac{1}{2}) tr\left[\Sigma^{-1}\left(\sum_{i=1}^{n}(v_i - \bar{v})(v_i - \bar{v})'\right)\right]} \\ &\times (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} e^{-(\frac{n}{2})(\bar{v} - \mu)'\Sigma^{-1}(\bar{v} - \mu)} \end{aligned} \tag{2.2.23}$$

Which depends on L and ψ through **Σ** = **LL**' + **ψ.** It is desirable to make L well defined by imposing the computationally convenient uniqueness condition $L^{'}\psi^{-1}L = \Delta$, which is a diagonal matrix [1, 4, 13].

$v_1, v_2, \ldots, v_n$ is a random sample from $N_P(\mu, \Sigma)$, where **Σ** = **LL**' + ψ is the covariance matrix for the m common factor model of (2.2.2). The maximum likelihood estimators $\tilde{L}, \tilde{\psi}$, and $\tilde{\mu} = \bar{v}$ maximize (2.2.23) subject to $\tilde{L}^{'}\tilde{\psi}^{-1}\tilde{L}$ being diagonal. The maximum likelihood estimates of the communalities are $\tilde{h}_i^2 = \tilde{l}_{i1}^1 + \tilde{l}_{i2}^2 + \cdots + \tilde{l}_{im}^2$ for i= 1, 2, … , p. (2.2.24). So proportion of total sample variance due to jth factor is

$$(\tilde{l}_{1j}^2 + \tilde{l}_{2j}^2 + \cdots + l_{pj}^2)/(s_{11} + s_{22} + \cdots + s_{pp})\,. \qquad (2.2.25)$$

Even if **V** is not normal, the Maximum-likelihood method might have nice interpretations.

2.2.3 Factor Rotation

Factor rotation is the process of manipulating or adjusting the factor axes in a clockwise direction to achieve a simple and theoretically more meaningful factor solution. When a set of factors has been derived, they are not always easy to interpret. Various methods have been proposed for rotating the factors to find new ones which may be easier to interpret. The usual aim of rotation methods is to make the loadings 'large' or 'small' so that most variables have a high loading on a small number of factors. It can help to achieve simple structure.

The factor loadings (row of Λ) in the population model are unique only up to multiplication by an orthogonal matrix that rotates the loadings. The rotated loadings reproduce the covariance matrix and satisfy all basic assumptions. The estimated loading matrix $\hat{\Lambda}$ can be rotated to obtain $\hat{\Lambda}^* = \hat{\Lambda}T$, where T is orthogonal. Since T'T=I by orthogonal matrix properties (CC'=I), the rotated loadings provide the same estimate of the covariance matrix as before: $S = \hat{\Lambda}^*\hat{\Lambda}^{*'} + \hat{\psi} = \hat{\Lambda}TT'\hat{\Lambda}' = \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$ [4. 18, 21].

The goal of rotation is to place the axes close to as many points as possible. If there are clusters of points, we seek to move the axes so as to pass though or near these clusters. This would associate each group of variables with a factor and make interpretation more objective. We are trying to extract the natural groupings in the variables. These groups of variables plot as clusters of pointes in the loading space. We attempt to move the axes as close to these clusters of points as possible. The resulting axes then represent the natural factors.

The initial factors will be rotated so that the factors meet criteria that make them more relevant to the purpose of the study. The variables are therefore selected so that the rotated factors have a greater likelihood of being meaningful. To rotate factors requires that a principal be utilized to determine the position of the factors. Most rotations are to analytic criteria and are carried out by computers.

In this project, we consider two types of rotation. The first, called orthogonal rotation, the original perpendicular axes are rotated rigidly and remain perpendicular. In an orthogonal rotation, angles and distances are preserved, communalities are unchanged, and the basic configuration of the points remains the same. Only the reference axes differ. The second method, called oblique rotation, makes no such restriction: the factor axes can be rotated independently, so that they are not necessarily perpendicular to one another after rotation. Orthogonal factor solutions are mathematically simple to handle, while oblique factor solutions are more flexible and more realistic, because the theoretically important underlying factors are not assumed to be uncorrelated to each other.

The choice of an orthogonal or oblique rotation should be made on the basis of the particular needs of a given research problem. If the goal of the research is to reduce the number of original variables regardless of how meaningful the resulting factors may be, then the appropriate approach would be an orthogonal rotation. Also, if the research wants to reduce the larger number of variables into a smaller set of uncorrelated variables for subsequent use in a regression or other prediction technique, then an orthogonal is the best. However, if the ultimate goal of the factor analysis is to obtain several theoretically meaningful factors, then an oblique method is appropriate.

2.2.3.1 Orthogonal Rotation

Orthogonal rotation is the process of extracting so that the factor axes are maintained at 90 degrees. There are three popular and readily algorithms are available: varimax rotation, quartiman rotation and equimax rotation. The algorithms differ in the definition of what constitutes simple structure. The variamx method is the most popular of these methods and is often used to rotate principal components solutions. The procedure seeks to rotate factors so that the variation of the squared factor loadings for a given factor is made large. This is accomplished by having large, medium, and small loadings within a particular factor. Normalized loadings are obtained by first dividing each variable's loading by the square root of its communality. By such a scaling all variable's are given equal weight in the rotation. The quartimax method rotates in such a fashion so as to accomplish, for a given variable, and only one major loading on a given factor. This usually cannot be accomplished. However, quartimax tries to get as close as possible to this criterion. An apparent undesirable property of quartumax is a tendency to generate a general factor with all or most of the variables having high loadings. The last method, equimax method, attempts to achieve simple structure with respect to both the rows and columns of the factor loading matrix [3, 4, 16].

The factors are extracted in such a way that the factor axes are maintained at 90 degrees, meaning that each factor is independent from all other factors. Therefore, the correlation between factors is arbitrarily determined to be zero. Any orthogonal rotation method will not alter the values of the communality estimates. However, the proportion of a variable's variance accounted for by a given factor will be different. Though the total amount of variance accounted for by the common factors does not change with orthogonal rotation, the percentage of variance accounted for by an individual factor will, in genera, be different. In the orthogonally rotated factor pattern matrix, no significance is attached to factor order. Obviously, the percentage of common variance accounted for by a common factor will also change after orthogonal rotation.

If there are only two factors, we can use a graphical rotation based on a visual inspection of a plot of factor loadings. If m is larger than 2 we can use the most popular method varimax rotation, which seeks rotated loadings that maximize the variance of the squared loadings in each column of $\hat{\Lambda}^*$. This method attempts to make the loadings either large or small to facilitate interpretation. But not all variables load highly on only one factor. In many cases, the points are not well closeted, and the axes simply cannot be rotated so as to be near all of them. If the loadings in a column were nearly equal, the variance would be close to 0. If the squared loadings approach 0 and 1 for factoring R, the variance will increase.

2.2.3.2 Oblique Rotation

Oblique rotation is the process of extracting the correlated factors rather than arbitrarily constraining the factor solution so the factors are orthogonally independent each other. There are five popular algorithms: oblimax rotation, quartimin rotation, covarimin rotation, biquartimin rotation and oblimin rotation. Oblimax methods seek to rotate the factors so that the numbers of high and low loading are increased by decreasing those in the middle range. Quartimin rotation minimized the sum of inner products of varimin is the structure loadings. Covarimin is the varimax analog of the oblique rotation methods. Biquartimin is a compromise algorithm falling somewhere between the quartimin and covarimin methods. Oblimin rotation is similar to the biquartimin method in that it combines the quartimin and covarimin methods, but in different combinations [3, 6, 8].

Oblique solutions assume the original variables are correlated to some extent, therefore the underlying factors must be similarly correlated.

With oblique rotation methods, communality estimates or variance accounted for are no longer applicable (the sums of the squared elements on the loading matrix are not invariant under oblique transformations. The connection between the two types of oblique axes: primary and reference is that the pattern of the primary axes is the structure of the reference axes.

This method uses a general nonsingular transformation matrix T so that $\Lambda^* = T'\Lambda$, by population properties cov(Ay)=AΣ'A, we know cov(Λ^*) $= T'IT = T'T = I$. Thus the new factors are correlated. Since distances and angles are not preserved, the communalities are changed [4, 6, 14].

2.2.4 Choosing the Number of Factors, m

How do we decide on how many factors to extract? When a large set of variables is factored, the analysis will extract the largest and best combinations of variables first, and then proceed to smaller, less understandable combinations. In deciding how many factors to extract, generally begin with four criteria.

1. *The percentage of variance criterion.* This method applies particularly to the principal component method. It also can be extended to the principal factor method, where prior estimates of communalities are used to form $S-\tilde{\psi}$ or $R-\tilde{\psi}$ .But $S-\tilde{\psi}$ or $R-\tilde{\psi}$ will often have some negative eigenvalues. Therefore, the values of m range from 1 to p, the cumulative proportion of eigenvalues will exceed 1 and reduce to 1. The better strategy is to choose m equal to the value for which the percentage first exceeds 100%. In General, consider a solution which accounts for 60 percent of the total variance as a satisfactory solution. That is, choose m equal to the number of factors necessary for the variance accounted for to achieve a predetermined percentage, say 60%, of the total variance tr(S) or tr(R)[6, 18].
2. *The latent root criterion***.** This rule is the commonly used criterion of long standing and performs well in practice. This method might be to use $R-\hat{\psi}$ and let m equal the number of positive eigenvalues. Choose m equal to the number of eigenvalues greater than the average eigenvalue. For R the average is 1: for S it is $\sum_{j=1}^{p}\frac{\theta_j}{p}$ [2, 6, 18].

3. *The scree test criterion*. The scree test was named after the geological term scree. It also results in practice well. This rule is derived by plotting the latent roots against the number of factors in their order of the extraction, and the shape pf the resulting curve is used to evaluate the cutoff point. Use the scree test based on a plot of the eigenvalues of S or R. If the graph drops sharply, followed by a straight line with much smaller slope, choose m equal to the number of eigenvalues before the straight line begins [2, 6, 18].

4. *The a priori criterion***.** This rule is most popular method. When applying the a priori criterion we already knows how many factors to extract before undertaking the factor analysis. We just test the hypothesis that m is the correct number of factors, H_0: **Σ=ΛΛ'+ Ψ.** Where **Λ** is p×m. That is, we wish to test $H_0 : \Sigma = LL^{'} + \psi$ vs $H_1 : \Sigma \neq LL^{'} + \psi$, where L is p×m. The test statistic, a function of the likelihood ratio, is $\left(n - \frac{2p + 4m + 11}{6}\right)\ln\left(\frac{\left|\tilde{L}\tilde{L}^{'} + \tilde{\psi}\right|}{|S|}\right)$ which is approximately χ^2_υ when H_0 is true, where ν=$\frac{1}{2}[(p-m)^2 - p - m]$ and $\hat{\Lambda}$ and $\hat{\psi}$ are the maximum likelihood estimators. Rejection of H_0 implies that m is too small and more factors are needed. When n is large, the test shows more factors to be significant than do the other three methods. This method is useful if we are attempting to replicate another researcher's work and extract exactly the same number of factors that was previously found [6, 18].

In practice, we seldom use a single criterion for selecting how many factors to extract. We should consider more criterions.

2.3 Drawbacks to Factor Analysis

Factor analysis does have some advantages over Principal Component Analysis, particularly in that maximum-likelihood estimation overcomes the scaling problem and that a proper statistical model, with an error structure, is involved. But there are many drawbacks to Factor Analysis:

1. A large number of assumptions have to be made in setting up the Factor Analysis model. These assumptions are not always realistic in practice.
2. An even more basic assumption in the Factor Analysis model is that the factors exist at all; The concept of a set of underlying unobservable variables is one which may be reasonable in some situation.
3. The Factor Analysis model also assumes knowledge of m, the number of factors. In practice, m is often unknown and different values may be tried sequentially, starting with m=1. But it is not easy to select the ‘correct ‘value of m. Although a test is available, it is rather complicated and depends on the model assumptions, so that external considerations are often used to select the value of m. It is somewhat disturbing to find that the form of the factors may change completely as m changes.
4. Even for a given value of m, the factors are not unique as different methods of rotation may produce set of factors which look quite different. Although rotation is mathematically respectable, the danger here is that the analyst may go on

trying different values of m and different methods of rotation until he gets the answer he is looking for. The lack of uniqueness introduced by allowing rotation also has the drawback that different investigators may use different rotations on the same set of data and get apparently different results. In addition, it is unusual to get repeatable results on replicate samples.

5. The Factor Analysis model in equation (3.2) has no obvious inverse and it is not so easy to estimating factor scores from observed data, though two methods of estimating factor scores are described by Lawley and Maxwell(1971,Chapter 8). This makes it more difficult to use Factor Analysis.

2.4 Factor Model Description in Finance

This multivariate statistical technique of factor analysis has found increase use during the past decade in the various fields of business related research, especially in marketing and personal management. A multifactor model is a general form of a factor model, and is the most popular model for the return generation process. The return r_i on the *i*th security is represented as,

$$r_i = \alpha_i + \sum_{m=1}^{k} \beta_{im} F_m + u_i \tag{2.5.1}$$

where k is the number of factors and it is a positive integer larger than zero,

$F_1, F_2, \cdots, F_K$ are the factors affecting the returns of *i*th security,

$\beta_{i1}, \beta_{i2}, \cdots, \beta_{ik}$ are the corresponding sensitivities,

α_i is regarded as "zero" factor that is invariant with time,

u_i is a zero mean random variable of *i*th security.

It is generally assumed that the covariance between u_i and factors F_i are zero. Also u_i and u_i for security i and j are independent if $i \neq j$. The simplest factor model is one-factor model, i.e., k=1. One-factor model with market index as the factor variable is called the market model. However, factor model does not restrict the factor to be the market index. Investigators use different approaches in factor model**.** The first one assumes some known fundamental factors are the factors that influence the security and $\beta's$ are evaluated accordingly**.** The second approach assumes the sensitivities to factors are known, and the factors are estimated from the security returns. The third approach is factor analysis. This one assumes neither factor values nor the security sensitivities is known. Under factor analysis approach, principle component analysis (PCA) was the most successful method. PCA was used to find the factors and their sensitivities. However it was also shown that the separated factors are not able to truly reflect the real case but also one meaningful factor, which corresponds to the market effect, is extracted. This is due to two limitations of PCA. First the separated principal components must be orthogonal to each other. Second, PCA uses only up to second order statistics, i.e. the covariance and correlation matrix[11].

Figure 2.1 shows the general steps followed in the application of factor analysis techniques [18].

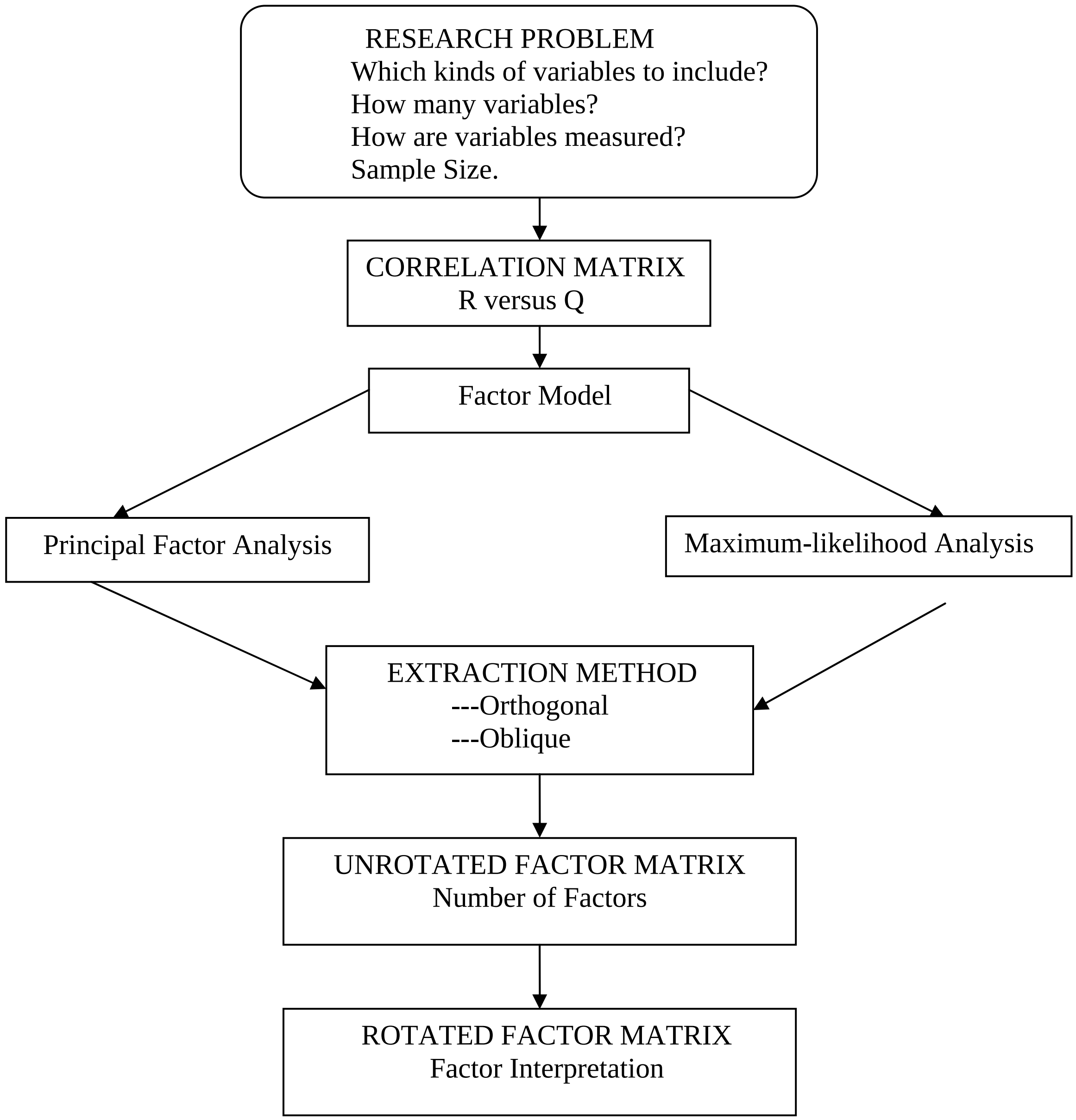

Figure 2.1 Factor Analysis Decision Diagram

3 Application of Factor Analysis to Financial Problems

We take the weekly rates of return for eight stocks (IBM, Dell, Apple, Sony, Novell, Microsoft, AMD, Intel) listed on the New York Stock Exchange and NASDA were determined for the period January 1998 through December 2004. The weekly rates of return are defined as (current Friday closing price-previous Friday closing price)/ (previous Friday price), adjusted for stock splits and dividends. The observations in 365 successive weeks appear to be independently distributed, but the rates of return across stocks are correlated, since, as one expects, stocks tend to move together in response to general economic conditions.

Till now, we have many choice of m, many choices of estimation methods and many choices of rotation. At the present time, factor analysis still maintains the flavor of an art, and no single strategy should yet be "chiseled into stone." In this project, we suggest and illustrate one reasonable option [4]:

1. Analysis of correlation matrix
2. Find out the perfect m
3. Perform a principal component factor analysis
4. Perform a maximum likelihood factor analysis
5. Testing for m-common factors
6. Compare the solutions obtained from the two factor analyses, PFA and MLFA
7. Try a varimax rotation for principal component factor analysis
8. Try a variamax rotation for maximum likelihood factor analysis
9. Compare the solutions obtained from the rotated two factor analyses

Let $x_1, x_2, \ldots, x_8$ denote observed weekly rates of return for IBM, Dell, Apple, Sony, Microsoft, Intel respectively. Then

$$\bar{x}' = [0.0014, 0.0019, 0.0070, -0.0001, -0.0016, 0.0003]$$

3.1 Correlation Analysis

The first decision in the application of factor analysis involves the calculation of the correlation matrix and analysis the correlation matrix.

The correlation matrix is presented next:

	X1	X2	X3	X4	X5	X6	X7	X8
X1	1.000	0.306	0.329	0.171	0.289	0.238	0.391	0.380
X2		1.000	0.328	0.188	0.239	0.270	0.347	0.404
X3			1.000	0.250	0.239	0.147	0.352	0.355
X4				1.000	0.328	0.237	0.225	0.241
X5					1.000	0.243	0.289	0.310
X6						1.000	0.260	0.345
X7							1.000	0.485
X8								1.000

Table 3.1 Correlation Matrix for Stock Price Data

The preceding correlation matrix consists of eight rows and eight columns. Where the row for one variable intersects with the column for a second variable, we can find the correlation for that pair of variables. For example, where the row for x_1 intersects with the column for x_2, we can see that the correlation between these variables is 0.306. Where the row for x_2 intersects with the column for x_3, we can see that the correlation between these variables is 0.328.

It is clear from the bold entries in the correlation matrix that variables 1, 2, 3 and 7 are correlated with one another, but these variables are not too much correlated with variables 4, 5 and 6. Similarly, variable 4 is correlated with variable 5. Let us analysis variable 8, it is correlated with every other variables except variable 4. Therefore, variables 1, 2, 3, 7 and 8 forms a group, variables 4 and 5 forms a group. Variable 8 also is correlated with variable 5 and 6. We analyzed the correlation matrix and got these results because we expect that the apparent linear relations between the variables can be explained in terms of, at most, three or four common factors.

3.2 Number of Common Factor

The final choice of m recommends combining all two in a structured sequence. First, perform a scree test and look for obvious breaks in the data. Because there will often be more than one break in the eigenvalue plot, it may be necessary to examine two or more possible solutions. Second, we should review the amount of common variance accounted for by each factor.

3.2.1 Scree Test

With the scree test (Cattell, 1966), we plot the eigenvalues associated with each factor, and look for a break between the factors with relatively large eigenvalues and those with smaller eigenvalues. The factors that appear before the break are assumed to be meaningful and are retained for rotation; those appearing after the break are assumed to be unimportant and are not retained.

Specifying the SCREE option in the PROC FACTOR statement causes the SAS System to print an eigenvalue plot as part of the output [9].

```
 3.5+
    |
    |
    |
    |
    |        1
  3.0 +
    |
    |
    |
    |
    |
  2.5 +
    |
    |
    |
E   |
i   |
g 2.0 +
e   |
n   |
v   |
a   |
l   |
u 1.5 +
e   |
s   |
    |
    |
    |
  1.0 +           2
    |                3
    |                   4
    |                      5
    |                         6
    |                            7
  0.5 +                             8
    |
    |
    |
    |
    |
  0.0 +
    ---------+--------+--------+--------+--------+--------+--------+--------+--------+------
        0     1     2     3     4     5     6     7     8    Number
```

Figure 3.1: Scree Plot of Eigenvalues (from SAS)

From Figure 3.1, we can see that the factor numbers are listed on the horizontal axis, while eigenvalues are listed on the vertical axis. This figure plots the first 8 factors extracted in this case. Starting with the first factor, the plot slopes steeply down initially and then slowly becomes an approximately horizontal line. The point at which the curve first begins to straighten out is considered to be the maximum number of factors to extract. With this plot, we also notice that there is relatively large break between factors 1 and 2 and a small break between 4 and 5 but that there is not a large break between factors 2 and 3, 3 and 4, 5 and 6, 6 and 7, 7 and 8. Because factor 5 through 8 have relatively small eigenvalues, and the data points for factors 5 through 8 could almost be fitted with a straight line, they can be assumed to be relatively unimportant factors. Because there is a relatively large break between factors1 and 2, factor 1 can be viewed as a relatively important factor. And factor 2, 3 and 4 can be viewed as relatively weakly important factors. Given the plot, a scree test would suggest that only factors 1, 2, 3 and 4 be retained. Factor 5 through 8 appear after the break, and thus will not be retained.

3.2.2 Proportion of Variance Accounted for

A second criterion in making the number of factors decision involves retaining a factor if it accounts for a certain proportion (or percentage) of the variance in the data set. This proportion can be calculated with a simple formula [2]:

$$\text{Proportion} = \frac{\textit{Eigenvalue for the factors}}{\textit{Total eigenvalues of the correlation matrix}}$$

In principal component analysis, the "total eigenvalues of the correlation matrix" was equal to the total number of variables being analyzed because each variable contributed one unit of variance to the data set. In common factor analysis, however, the total eigenvalues will be equal to the sum of the communalities that appear on the main diagonal of the matrix being analyzed.

.

Variable	Eigenvalue	Difference	Proportion	Cumulative
1. IBM	3.080	2.118	0.385	0.385
2. Dell	0.962	0.110	0.120	0.505
3. Apple	0.853	0.119	0.107	0.612
4. Sony	0.734	0.073	0.092	0.704
5. Novell	0.660	0.030	0.083	0.786
6. Microsoft	0.630	0.046	0.079	0.865
7. AMD	0.584	0.087	0.073	0.938
8. Intel	0.497	-	0.062	1.000

Table 3.2 Eigenvalues of the Reduced Correlation Matrix

From the "Proportion" row of the preceding eigenvalue table, we can see that the first factor alone accounts for 38.50% of the common variance, the second factor alone accounts for12.03%, and the third accounts for 10.66%, the forth factor accounts for 9.17%. If we were using, say, 10% as the criterion for deciding whether a factor should be retained, only factor 1, 2 and 3 would be retained in this case.

3.3 Perform a Principal Factor Analysis

We perform the principal factor analysis used with PROC FACTOR in SAS program for four factors separately. The PROC FACTOR statement begins the FACTOR procedure, and the number of options should be added [2, 9]:

METHOD=PRINCIPAL

to request that the principal factors method be used for the initial extraction. Although the principal factors methods is probably the most popular extraction method, some researchers prefer the maximum likelihood method because it provides a significance test for solving the " number of factors" problem, and it is also believe to provide better parameter estimates.

NFACT=n

Allows us to specify the number of factors to be retained, where n is the number of factors.

3.3.1 One Factor Solution for Principal Factor Analysis

PFA	One-factor Solution	
Variable	Estimated Factor Loadings F1	Specific Variances $\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1. IBM	0.639	0.39
2. Dell	0.631	0.60
3. Apple	0.609	0.63
4. Sony	0.496	0.75
5. Novell	0.579	0.67
6. Microsoft	0.532	0.72
7. AMD	0.701	0.51
8. Intel	0.739	0.45
Cumulative proportion of total (standardized) sample variance explained	0.385	

Table 3.3: One Factor Solution for Principal Factor Analysis

Let us analyze the first factor, F_1. For principal component factorization, all stocks have large positive loadings on the first factor, and the loadings for variables 1, 2, 3 are very closer, almost 0.60. The variables 5 and 6 are much closer, around 0.53 and the variables 7 and 8 are very closer, almost 0.70. This factor might be labeled general economic conditions and can be called a market factor.

The residual matrix corresponding to the solution for m=1 factors is

$$R-\hat{L}\hat{L}'-\hat{\psi}=\begin{bmatrix} 0.7983 & 0.4032 & 0.3892 & 0.3169 & 0.3700 & 0.3399 & 0.4479 & 0.4722 \\ 0.4032 & 0.9982 & 0.3843 & 0.3130 & 0.3653 & 0.3357 & 0.4423 & 0.4663 \\ 0.3892 & 0.3843 & 1.0009 & 0.3021 & 0.3526 & 0.3240 & 0.4269 & 0.4501 \\ 0.3169 & 0.3130 & 0.3021 & 0.9960 & 0.2872 & 0.2639 & 0.3477 & 0.3665 \\ 0.3700 & 0.3653 & 0.3526 & 0.2872 & 1.0052 & 0.3080 & 0.4059 & 0.4279 \\ 0.3399 & 0.3357 & 0.3240 & 0.2639 & 0.3080 & 1.0030 & 0.3729 & 0.3931 \\ 0.4479 & 0.4423 & 0.4269 & 0.3477 & 0.4059 & 0.3729 & 1.0014 & 0.5180 \\ 0.4722 & 0.4663 & 0.4501 & 0.3665 & 0.4279 & 0.3931 & 0.5180 & 0.9961 \end{bmatrix} \quad (3.1)$$

The elements of $R-\tilde{L}\tilde{L}'-\tilde{\psi}$ are not smaller. We don't prefer this approach for 1 factor.

3.3.2 Two-Factor Solution for Principal Factor Analysis

PFA	Two-factor Solution		
Variable	Estimated Factor Loadings		Specific Variances
	F1	F2	$\tilde{\psi}_i = 1-\tilde{h}_i^2$
1. IBM	0.639	-0.259	0.53
2. Dell	0.631	-0.242	0.54
3. Apple	0.609	-0.216	0.58
4. Sony	0.496	0.664	0.31
5. Novell	0.579	0.441	0.47
6. Microsoft	0.532	0.275	0.64
7. AMD	0.701	-0.231	0.46
8. Intel	0.739	-0.161	0.53

Cumulative proportion of total (standardized) sample variance explained	0.385	0.505	

Table 3.4: Two-Factor Solution for Principal Factor Analysis

The second factor contrasts the hardware stocks, the semiconductor stocks with software & programming stocks and audio & video equipment stocks. The hardware and the semiconductor stocks have negative loadings. The hardware stocks have relatively larger negative loadings than the semiconductor stocks. The software & programming stocks and audio & video equipment stocks have positive loadings. The audio & video equipment stocks have relatively larger positive loadings, on the factor. Thus, F2 seems to differentiate stocks in different industries and might be called an industry factor.

The residual matrix corresponding to the solution for m=2 factor is

$$R-\tilde{L}\tilde{L}'-\tilde{\psi}=\begin{bmatrix} -0.0054 & -0.1595 & -0.1159 & 0.0263 & 0.0335 & -0.0304 & -0.2168 & -0.1337 \\ -0.1595 & 0.0033 & -0.1089 & 0.0361 & -0.0192 & 0.0005 & -0.1515 & -0.1011 \\ -0.1159 & -0.1089 & 0.0025 & 0.0909 & -0.0186 & -0.1172 & -0.1244 & -0.1293 \\ 0.0263 & 0.0361 & 0.0909 & 0.0031 & -0.2520 & -0.2099 & 0.0311 & -0.0188 \\ 0.0355 & -0.0192 & -0.0186 & -0.2520 & 0.0003 & -0.1868 & -0.0155 & -0.0466 \\ -0.0304 & 0.0005 & -0.1172 & -0.2099 & -0.1868 & 0.0014 & -0.0497 & -0.0041 \\ -0.1168 & -0.1515 & -0.1244 & 0.0311 & -0.0155 & -0.0497 & -0.0048 & -0.0703 \\ -0.1337 & -0.1011 & -0.1293 & -0.0188 & -0.0466 & -0.0041 & -0.0703 & -0.1021 \end{bmatrix} \quad (3.2)$$

The elements of $R-\tilde{L}\tilde{L}'-\tilde{\psi}$ are much smaller than those of the residual matrix corresponding to the one factor. We still need to compare these elements with more factors.

3.3.3 Three-Factor Solution for Principal Factor Analysis

PFA	Three-factor Solution			
Variable	Estimated Factor Loadings			Specific Variances
	F1	F2	F3	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1. IBM	0.639	-0.259	0.058	0.53
2. Dell	0.631	-0.242	-0.133	0.53
3. Apple	0.609	-0.216	0.478	0.35
4. Sony	0.496	0.664	0.272	0.24
5. Novell	0.579	0.441	0.165	0.44
6. Microsoft	0.532	0.275	-0.690	0.17
7. AMD	0.701	-0.231	0.013	0.46
8. Intel	0.739	-0.161	-0.158	0.40
Cumulative proportion of total (standardized) sample variance explained	0.385	0.505	0.612	

Table 3.5: Three Factors Solution for Principal Factor Analysis

Focusing attention on the third factor loadings, we see that all of variables except the variable 4 have small loadings. The loadings for variable 2, 6 and 8 are negative

The residual matrix corresponding to the solution for m=3 factor is

$$R - \tilde{L}\tilde{L}' - \tilde{\psi} = \begin{bmatrix} 0.0088 & -0.1517 & -0.1437 & 0.0105 & 0.0239 & 0.0097 & -0.2176 & -0.1245 \\ -0.1517 & -0.0044 & -0.0454 & 0.0722 & 0.0027 & -0.0913 & -0.1498 & -0.1221 \\ -0.1437 & -0.0454 & 0.0040 & -0.0391 & -0.0975 & 0.2127 & -0.1306 & -0.0538 \\ 0.0105 & 0.0722 & -0.0391 & -0.0009 & -0.2969 & -0.0222 & 0.0275 & 0.0242 \\ 0.0239 & 0.0027 & -0.0975 & -0.2969 & 0.0031 & -0.0729 & -0.0176 & -0.0206 \\ 0.0097 & -0.0913 & 0.2127 & -0.0222 & -0.0729 & -0.0047 & -0.0408 & -0.1131 \\ -0.1176 & -0.1498 & -0.1306 & 0.0275 & -0.0176 & -0.0408 & -0.0049 & -0.0683 \\ -0.1245 & -0.1221 & -0.0538 & 0.0242 & -0.0206 & -0.1131 & -0.0683 & 0.0030 \end{bmatrix} \quad (3.3)$$

3.3.4 Four-Factor Solution for Principal Factor Analysis

PFA	Four- factor Solution				
Variable	Estimated Factor Loadings				Specific Variances
	F1	F2	F3	F4	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1. IBM	0.639	-0.259	0.058	-0.458	0.31
2. Dell	0.631	-0.242	-0.133	0.396	0.37
3. Apple	0.609	-0.216	0.478	0.297	0.27
4. Sony	0.496	0.664	0.272	0.270	0.17
5. Novell	0.579	0.441	0.165	-0.427	0.26
6. Microsoft	0.532	0.275	-0.690	0.089	0.16
7. AMD	0.701	-0.231	0.013	-0.123	0.44
8. Intel	0.739	-0.161	-0.158	0.019	0.40
Cumulative proportion of total (standardized) sample variance explained	0.385	0.505	0.612	0.704	

Table 3.6: Four-Factor Solution for Principal Factor Analysis

The residual matrix corresponding to the solution for m=4 factors is smaller. Thus, on a purely descriptive basis, we would judge a four-factor model with the factor loadings as providing a good fit to the data.

The proportion of the total variance explained by the four-factor solution is appreciably larger than that for the one-factor, two-factor, three-factor solutions.

The residual matrix corresponding to the solution for m=4 factor is

$$R-\tilde{L}\tilde{L}'-\tilde{\psi}=\begin{bmatrix} 0.0015 & 0.0296 & -0.0076 & 0.1342 & -0.1717 & 0.0504 & -0.2739 & 0.1158 \\ 0.0296 & -0.0012 & -0.1630 & -0.0347 & 0.1718 & -0.1265 & -0.1011 & -0.1296 \\ -0.0076 & -0.1630 & -0.0042 & -0.1193 & 0.0294 & 0.1862 & -0.0940 & -0.0595 \\ 0.1342 & -0.0347 & -0.1193 & -0.0038 & -0.1816 & -0.0462 & 0.0607 & 0.0190 \\ -0.1717 & 0.1718 & 0.0294 & -0.1816 & 0.0007 & -0.0349 & -0.0701 & -0.0125 \\ 0.0504 & -0.1265 & 0.1862 & -0.0462 & -0.0349 & -0.0027 & -0.0298 & -0.1148 \\ -0.1739 & -0.1011 & -0.0940 & 0.0607 & -0.0701 & -0.0298 & -0.0001 & -0.0659 \\ -0.1158 & -0.1296 & -0.0595 & 0.0190 & -0.0125 & -0.1148 & -0.0659 & 0.0026 \end{bmatrix} \quad (3.4)$$

We use the following line chart for comparing the residual matrix for 4 factors

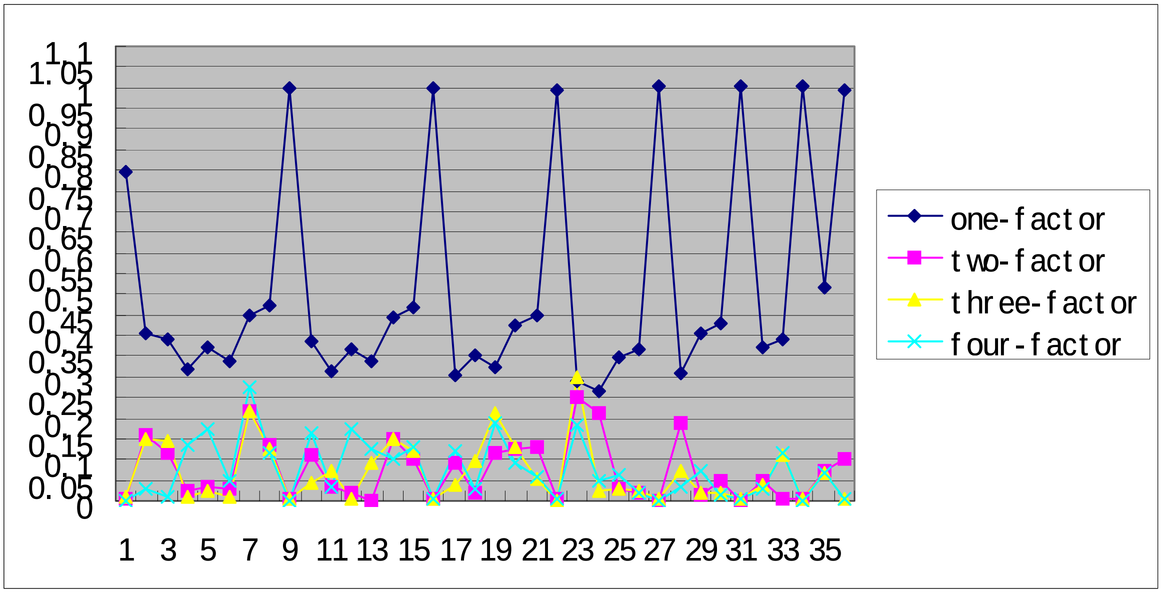

Fig. 3.2 Residual Matrix for one-factor, two-factor, three-factor and four-factor PFA

We find the residual matrix for m=1 is largest and the residual matrix for m=3 are smallest. We prefer using m is 3.

3.4 Perform a Maximum-likelihood Factor Analysis

We want to use maximum-likelihood factor analyses for one, two, three and four factors. It is already apparent from the principal factor analysis that the best number of common factors is almost certainly three. The one factor, two factor and four-factor maximum-likelihood solutions reinforce this conclusion and illustrate some of the numerical problems that can occur.

3.4.1 One Factor for Maximum-likelihood Factor Analysis

Maximum-likelihood	One-factor Solution	
Variable	Estimated Factor Loadings F1	Specific Variances $\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1. IBM	0.565	0.68
2. Dell	0.556	0.69
3. Apple	0.529	0.72
4. Sony	0.393	0.85
5. Novell	0.479	0.77
6. Microsoft	0.445	0.80
7. AMD	0.653	0.57
8. Intel	0.703	0.51

Table 3.7 Maximum-likelihood Factor Analysis with One Factor

```
          Iteration   Criterion Ridge    Change                   Communalities

     1         0.0875196  0.0000  0.1175  0.31830  0.30810  0.27977  0.16028  0.23581  0.19782
                                          0.42048  0.48937
     2         0.0874237  0.0000  0.0058  0.31867  0.30912  0.27964  0.15516  0.23003  0.19791
                                          0.42601  0.49330
     3         0.0874230  0.0000  0.0005  0.31871  0.30919  0.27957  0.15470  0.22961  0.19779
                                          0.42633  0.49377


                           Convergence criterion satisfied.




                       Significance Tests Based on 364 Observations

                                                                 Pr >
                          Test                     DF    Chi-Square     ChiSq

                 H0: No common factors             28      538.0624    <.0001
                 HA: At least one common factor
                 H0: 1 Factor is sufficient        20       31.3703    0.0505
                 HA: More factors are needed
```

Table 3.8 Output from SAS for the analysis of one factor

$$\hat{L}\hat{L}^{'} + \hat{\psi} = \begin{bmatrix} 0.9987 & 0.3139 & 0.2985 & 0.2220 & 0.2705 & 0.2510 & 0.3686 & 0.3967 \\ 0.3139 & 0.9992 & 0.2940 & 0.2187 & 0.2664 & 0.2473 & 0.3631 & 0.3908 \\ 0.2985 & 0.2940 & 0.9995 & 0.2079 & 0.2533 & 0.2351 & 0.3452 & 0.3715 \\ 0.2220 & 0.2187 & 0.2079 & 1.0047 & 0.1884 & 0.1749 & 0.2568 & 0.2764 \\ 0.2705 & 0.2664 & 0.2533 & 0.1884 & 0.9995 & 0.2131 & 0.3129 & 0.3367 \\ 0.2510 & 0.2473 & 0.2351 & 0.1749 & 0.2131 & 0.9978 & 0.2904 & 0.3125 \\ 0.3686 & 0.3631 & 0.3452 & 0.2568 & 0.3129 & 0.2904 & 0.9964 & 0.4589 \\ 0.3967 & 0.3908 & 0.3715 & 0.2764 & 0.3367 & 0.3125 & 0.4568 & 1.0038 \end{bmatrix} \quad (3.5)$$

$$R - \hat{L}\hat{L}^{'} - \hat{\psi} = \begin{bmatrix} 0.0013 & -0.0075 & 0.0307 & -0.0508 & 0.0188 & -0.0127 & 0.0223 & -0.0164 \\ -0.0075 & 0.0008 & 0.0336 & -0.0304 & -0.0270 & 0.0224 & -0.0164 & 0.0134 \\ 0.0307 & 0.0336 & 0.0005 & 0.0416 & -0.0145 & -0.0877 & 0.0072 & -0.0160 \\ -0.0508 & -0.0304 & 0.0416 & -0.0047 & 0.1396 & 0.0617 & -0.0314 & -0.0356 \\ 0.0188 & -0.0270 & -0.0145 & 0.1396 & 0.0005 & 0.0295 & -0.0243 & -0.0264 \\ -0.0127 & 0.0224 & -0.0877 & 0.0617 & 0.0295 & 0.0022 & -0.0307 & 0.0323 \\ 0.0223 & -0.0164 & 0.0072 & -0.0314 & -0.0243 & -0.0307 & 0.0036 & 0.0261 \\ -0.0164 & 0.0134 & -0.0160 & -0.0356 & -0.0264 & 0.0323 & 0.0261 & -0.0038 \end{bmatrix} \quad (3.6)$$

Table 3.8 displays the results of the analysis with one factor. The solution on the third iteration is so close to the optimum that PROC FACTOR cannot find a better solution, hence we can get this message: Convergence criterion satisfied.

3.4.2 Two-Factor for Maximum-likelihood Factor Analysis

Maximum likelihood	Two-factor solution		
Variable	Estimated factor loadings		Specific variances
	F1	F2	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1. IBM	0.171	0.547	0.67
2. Dell	0.188	0.527	0.69
3. Apple	0.250	0.463	0.72
4. Sony	1.000	0.000	0.00
5. Novell	0.328	0.362	0.76
6. Microsoft	0.237	0.373	0.81
7. AMD	0.225	0.622	0.56
8. Intel	0.739	0.670	0.49

Table 3.9: Maximum-likelihood Factor Analysis with Two-Factor

```
          Iteration   Criterion   Ridge  Change                      Communalities

   1       0.0394961  0.0000  0.2078  0.32596  0.31363  0.28101  0.37357  0.30658  0.20255
                                      0.44428  0.49840
   2       0.0373806  0.0000  0.1762  0.32402  0.31227  0.27785  0.54973  0.25041  0.19989
                                      0.44142  0.50727
   3       0.0356287  0.0000  0.3434  0.32517  0.31271  0.27594  0.89308  0.22374  0.19238
                                      0.43840  0.50674
   4       0.0353177  0.0000  0.1069  0.32515  0.31275  0.27564  1.00000  0.22140  0.19187
                                      0.43817  0.50664
   5       0.0350850  0.0000  0.0168  0.32843  0.31361  0.27668  1.00000  0.23818  0.19510
                                      0.43712  0.50735
   6       0.0350849  0.0000  0.0002  0.32863  0.31360  0.27662  1.00000  0.23825  0.19523
                                      0.43706  0.50713


        Convergence criterion satisfied.

       Significance Tests Based on 364 Observations

                                                              Pr >
                Test                      DF    Chi-Square     ChiSq

        H0: No common factors             28      538.0624    <.0001
        HA: At least one common factor
         H0: 2 Factors are sufficient      13       12.5662     0.4818


                  Final Community Estimates and Variable Weights
          Total Communality: Weighted = 3.689455   Unweighted = 3.296527

                Variable    Communality       Weight

                x1           0.32863875    1.48949479
                x2           0.31359938    1.45687961
                x3           0.27662033    1.38239279
                x4           1.00000000         Infty
                x5           0.23826002    1.31277321
                x6           0.19522642    1.24259353
                x7           0.43705484    1.77638549
```

Table 3.10 Output from SAS for the analysis of two factors

$$\hat{L}\hat{L}' + \hat{\psi} = \begin{bmatrix} 0.9987 & 0.3208 & 0.2960 & 0.1713 & 0.2540 & 0.2447 & 0.3786 & 0.4933 \\ 0.3208 & 1.0036 & 0.2912 & 0.1884 & 0.2525 & 0.2414 & 0.3702 & 0.4927 \\ 0.2960 & 0.2912 & 0.9966 & 0.2495 & 0.2492 & 0.2318 & 0.3440 & 0.4947 \\ 0.1713 & 0.1884 & 0.2495 & 1.0000 & 0.3280 & 0.2366 & 0.2254 & 0.7390 \\ 0.2540 & 0.2525 & 0.2492 & 0.3280 & 0.9983 & 0.2125 & 0.2986 & 0.4847 \\ 0.2447 & 0.2414 & 0.2318 & 0.2366 & 0.2125 & 1.0053 & 0.2853 & 0.4250 \\ 0.3786 & 0.3702 & 0.3440 & 0.2254 & 0.2986 & 0.2853 & 0.9971 & 0.5831 \\ 0.4933 & 0.4927 & 0.4947 & 0.7390 & 0.4847 & 0.4250 & 0.5831 & 1.4853 \end{bmatrix} \tag{3.7}$$

$$R - \hat{L}\hat{L} - \hat{\psi} = \begin{bmatrix} 0.0013 & -0.0144 & 0.0331 & -0.0000 & 0.0353 & -0.0063 & 0.0123 & -0.1130 \\ -0.0144 & -0.0036 & 0.0364 & -0.0001 & -0.0130 & 0.0283 & -0.0236 & -0.0885 \\ 0.0331 & 0.0364 & 0.0034 & 0.0000 & -0.0104 & -0.0844 & 0.0084 & -0.1392 \\ -0.0000 & -0.0001 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.4982 \\ 0.0353 & -0.0130 & -0.0104 & 0.0000 & 0.0017 & 0.0300 & -0.0101 & -0.1744 \\ -0.0063 & 0.0283 & -0.0844 & 0.0000 & -0.0300 & -0.0053 & -0.0256 & -0.0801 \\ 0.0123 & -0.0236 & 0.0084 & 0.0000 & -0.0101 & -0.0256 & 0.0029 & -0.0982 \\ -0.1130 & -0.0885 & -0.1392 & -0.4982 & -0.1744 & -0.0801 & -0.0982 & -0.4853 \end{bmatrix} \tag{3.8}$$

Table 3.10 displays the results of the analysis using two factors. The analysis converges without incident. The variable IBM has a communality estimate as a missing/infinite value. The first eigenvalue is also infinite. Infinite values are ignored in computing the total of the eigenvalues and the total final communality. The IBM variable is a Heywood case.

3.4.3 Three-Factor for Maximum-likelihood Factor Analysis+

Maximum likelihood	Three-factor solution			
Variable	Estimated factor loadings			Specific variances
	F1	F2	F3	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1. IBM	0.171	0.504	0.223	0.67
2. Dell	0.188	0.499	0.162	0.69
3. Apple	0.250	0.398	0.306	0.69
4. Sony	1.000	0.000	0.000	0.00
5. Novell	0.328	0.347	0.094	0.76
6. Microsoft	0.237	0.614	0.511	0.31
7. AMD	0.225	0.566	0.274	0.55
8. Intel	0.241	0.631	0.193	0.51

Table 3.11: Maximum-likelihood Factor Analysis with Three Factors

```
Iteration   Criterion   Ridge  Change                     Communalities

    1       0.0163415  0.0000  0.3351  0.32531  0.30779  0.36562  0.50087  0.22485  0.36481
                                       0.43645  0.52141
    2       0.0135414  0.0000  0.2235  0.32939  0.31284  0.35630  0.72441  0.24240  0.36936
                                       0.43930  0.51030
    3       0.0129628  0.0000  0.2756  0.33152  0.31265  0.33877  1.00000  0.22925  0.40585
                                       0.44008  0.50411
    4       0.0127418  0.0000  0.0614  0.33415  0.31242  0.33081  1.00000  0.23797  0.46721
                                       0.44179  0.49925
    5       0.0126677  0.0000  0.0649  0.33404  0.31177  0.32360  1.00000  0.23750  0.53208
                                       0.44376  0.49649
    6       0.0126354  0.0000  0.0604  0.33375  0.31122  0.31933  1.00000  0.23713  0.59247
                                       0.44498  0.49503
    7       0.0126246  0.0000  0.0459  0.33352  0.31085  0.31693  1.00000  0.23689  0.63838
                                       0.44566  0.49426
    8       0.0126218  0.0000  0.0286  0.33338  0.31064  0.31563  1.00000  0.23676  0.66693
                                       0.44602  0.49388
    9       0.0126211  0.0000  0.0151  0.33330  0.31053  0.31498  1.00000  0.23669  0.68202
                                       0.44620  0.49370
   10       0.0126210  0.0000  0.0072  0.33326  0.31048  0.31467  1.00000  0.23666  0.68920
                                       0.44629  0.49361
   11       0.0126210  0.0000  0.0032  0.33324  0.31045  0.31452  1.00000  0.23665  0.69242
                                  Convergence criterion satisfied.
                          Significance Tests Based on 364 Observations (continued to next page)

                                                                       Pr >
                      Test                        DF    Chi-Square     ChiSq

                HO: No common factors              28       538.0624    <.0001
                HA: At least one common factor
                HO: 3 Factors are sufficient        7         4.5120    0.7193

                    Final Communality Estimates and Variable Weights
                 Total Communality: Weighted = 5.772038   Unweighted = 3.829078
```

Variable	Communality	Weight
x1	0.33322778	1.49976561
x2	0.31043359	1.45018841
x3	0.31443510	1.45865210
x4	1.00000000	Infty
x5	0.23663569	1.30999051
x6	0.69444362	3.27271768
x7	0.44635174	1.80619614
x8	0.49355028	1.97452746

Table 3.12: Output from SAS for the analysis of three factors

$$\hat{L}\hat{L}^{'} + \hat{\psi} = \begin{bmatrix} 1.0032 & 0.3198 & 0.3118 & 0.1713 & 0.2519 & 0.2362 & 0.3851 & 0.4025 \\ 0.3198 & 1.0004 & 0.2952 & 0.1883 & 0.2499 & 0.2682 & 0.3692 & 0.3914 \\ 0.3118 & 0.2952 & 1.0045 & 0.2495 & 0.2487 & 0.1474 & 0.3655 & 0.3706 \\ 0.1713 & 0.1883 & 0.2495 & 1.0000 & 0.3280 & 0.2366 & 0.2254 & 0.2408 \\ 0.2519 & 0.2499 & 0.2487 & 0.3280 & 0.9966 & 0.2427 & 0.2959 & 0.3159 \\ 0.2362 & 0.2682 & 0.1474 & 0.2366 & 0.2427 & 1.0044 & 0.2610 & 0.3458 \\ 0.3851 & 0.3692 & 0.3655 & 0.2254 & 0.2959 & 0.2610 & 0.9964 & 0.4645 \\ 0.4025 & 0.3914 & 0.3706 & 0.2408 & 0.3159 & 0.3458 & 0.4645 & 1.0036 \end{bmatrix} \tag{3.9}$$

$$R - \hat{L}\hat{L}^{'} - \hat{\psi} = \begin{bmatrix} -0.0002 & -0.0134 & 0.0174 & 0 & 0.0373 & 0.0022 & -0.0942 & -0.0222 \\ -0.0134 & -0.0004 & 0.0324 & 0 & -0.0105 & 0.0015 & -0.0225 & 0.0128 \\ 0.0174 & 0.0324 & -0.0005 & 0 & -0.0099 & 0.0000 & -0.0131 & -0.0151 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0373 & -0.0105 & -0.0099 & 0 & 0.0004 & -0.0001 & -0.0074 & -0.0057 \\ 0.0022 & 0.0015 & 0.0000 & 0 & -0.0001 & -0.0004 & -0.0013 & -0.0010 \\ 0.0058 & -0.0225 & -0.0131 & 0 & -0.0074 & -0.0013 & 0.0004 & 0.0205 \\ -0.0222 & 0.0128 & -0.0151 & 0 & -0.0057 & -0.0010 & 0.0205 & -0.0006 \end{bmatrix} \tag{3.10}$$

Table 3.12 displayed the results of the analysis using three-factor. The variable Sony has a communality estimate as a missing/infinite value. This time, The Sony variable is a Heywood case [9, 19].

The probability levels for the chi-square test are less than 0.0001 for the hypothesis of no common factors, 0.0505 for one common factor, and 0.4818 for two common factors, and

0.7193 for three common factors. Therefore, the three-factor model seems to be an adequate representation.

We also can use a comparison line chart for comparing the residual matrix for one-factor, two-factor and three-factor.

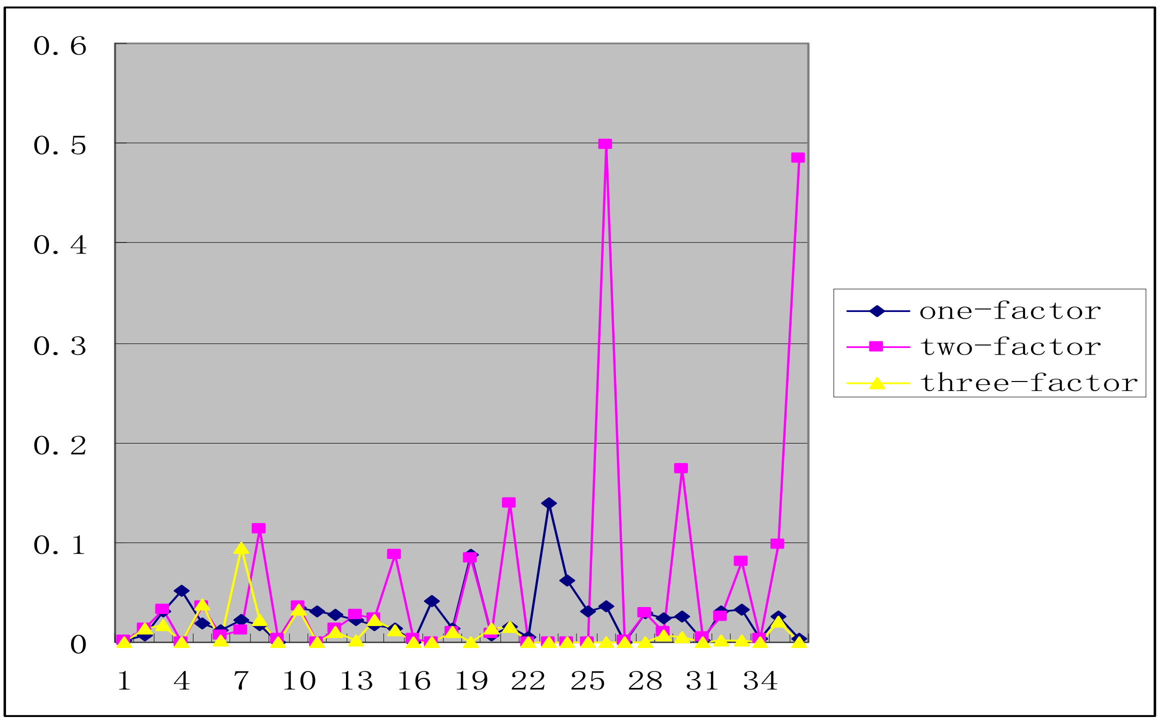

Fig. 3.3 Residual Matrix for one-factor, two-factor, and three-factor MLFA

We find the yellow line is very close to the x axis. The value in this line is much smaller than the pink line and the black like. That means the elements of $R-\hat{L}\hat{L}'-\hat{\psi}$ are much smaller than those of the residual matrix corresponding to the one-factor and two-factor for maximum-likelihood. Therefore, three-factor for maximum-likelihood factor model is the best model for this case.

3.5 Large Sample Test for Three Common Factors

Test the hypothesis [4, 13, 17] $H_0 : \Sigma = LL^{'} + \psi$, with m=2, at the test statistic in

$$(n-1-(2p+4m+5)/6)\ln\frac{\left|\hat{L}\hat{L}^{'}+\hat{\psi}\right|}{\left|S_n\right|} > \chi^2_{[(p-m)^2-p-m]/2}(\alpha) \tag{3.5.1}$$

is based on the ratio of generalized variances

$$\frac{\left|\hat{\Sigma}\right|}{\left|S_n\right|}=\frac{\left|\hat{L}\hat{L}^{'}+\hat{\psi}\right|}{\left|S_n\right|} \tag{3.5.2}$$

Let $\hat{V}^{-1/2}$ be the diagonal matrix such that $\hat{V}^{-1/2}S_n\hat{V}^{-1/2}=R$. By the properties of determinants

$$\left|\hat{V}^{-1/2}\right|\left|\hat{L}\hat{L}^{'}+\hat{\psi}\right|\left|\hat{V}^{-1/2}\right|=\left|\hat{V}^{-1/2}\hat{L}\hat{L}^{'}\hat{V}^{-1/2}+\hat{V}^{-1/2}\hat{\Psi}\hat{V}^{-1/2}\right| \tag{3.5.3}$$

and

$$\left|\hat{V}^{-1/2}\right|\left|S_n\right|\left|\hat{V}^{-1/2}\right|=\left|\hat{V}^{-1/2}S_n\hat{V}^{-1/2}\right| \tag{3.5.4}$$

Consequently,

$$\begin{aligned}\frac{\left|\hat{\Sigma}\right|}{\left|S_n\right|}&=\frac{\left|\hat{V}^{-1/2}\right|}{\left|\hat{V}^{-1/2}\right|}\frac{\left|\hat{L}\hat{L}^{'}+\hat{\psi}\right|}{\left|S_n\right|}\frac{\left|\hat{V}^{-1/2}\right|}{\left|\hat{V}^{-1/2}\right|}\\&=\frac{\left|\hat{V}^{-1/2}\hat{L}\hat{L}^{'}\hat{V}^{-1/2}+\hat{V}^{-1/2}\hat{\psi}\hat{V}^{-1/2}\right|}{\left|\hat{V}^{-1/2}S_n\hat{V}^{-1/2}\right|}\\&=\frac{\left|\hat{L}_Z\hat{L}_Z^{'}+\hat{\psi}_Z\right|}{\left|R\right|}\end{aligned} \tag{3.5.5}$$

By $\hat{\rho} = (\hat{V}^{-1/2}\hat{L})(\hat{V}^{-1/2}\hat{L}^{'}) + \hat{V}^{-1/2}\hat{\psi}\hat{V}^{-1/2} = \hat{L}_Z\hat{L}_Z^{'} + \hat{\psi}_Z$, we determine

$$\frac{\left|\hat{L}\hat{L}^{'} + \hat{\psi}\right|}{|R|} = \frac{0.2294}{0.2295} = 0.9908$$

Using Bartlett's correction, we get the test statistic in

$$(n-1-(2p+4m+5)/6)\ln\frac{\left|\hat{L}\hat{L}^{'} + \hat{\psi}\right|}{|S_n|} > \chi^2_{[(p-m)^2-p-m]/2}(\alpha)$$

$$[n-1-(2p+4m+5)/6]\ln\frac{\left|\hat{L}\hat{L}^{'} + \hat{\psi}\right|}{|S_n|} = [364-1-(16+12+5)/6]\ln\frac{0.2274}{0.2295} = \text{-}3.289$$

Since $\frac{1}{2}[(p-m)^2 - p - m] = \frac{1}{2}[(8-3)^2 - 8 - 3] = 7$, the 5% critical value

$\chi^2_7(0.05) = 14.07$ is not exceeded, and we fail to reject H_0. We conclude that the data do

not contradict a three-factor model.

3.6 Compare the Solutions Obtained from PFA and MLFA Using Three-Factor

	Maximum likelihood				Principal Factor			
Variable	Estimated factor Loadings			Specific Variance	Estimated Factor Loadings			Specific Variance
	F_1	F_2	F_3	$\hat{\psi}_i = 1-\hat{h}$	F_1	F_2	F_3	$\hat{\psi}_i = 1-\hat{h}$
1. IBM	0.171	0.504	0.223	0.67	0.639	-0.259	0.058	0.53
2. Dell	0.188	0.49872	0.162	0.69	0.631	-0.242	0.133	0.53
3. Apple	0.250	0.398	0.306	0.69	0.609	-0.216	0.478	0.35
4. Sony	1.000	0.000	-0.000	0.00	0.496	0.664	0.272	0.24
5. Novell	0.328	0.347	0.094	0.76	0.579	0.441	0.165	0.44
6. Microsoft	0.237	0.614	-0.511	0.31	0.532	0.275	0.690	0.17
7. AMD	0.225	0.566	0.274	0.55	0.701	-0.231	0.013	0.46
8. Intel	0.241	0.631	0.193	0.51	0.739	-0.161	0.158	0.40

Table 3.13 Solutions obtained from PFA and MLFA using three-factor

The residual matrix for maximum likelihood method is

$$R-\hat{L}\hat{L}'-\hat{\psi}=\begin{bmatrix} -0.0002 & -0.0134 & 0.0174 & 0 & 0.0373 & 0.0022 & -0.0942 & -0.0222 \\ -0.0134 & -0.0004 & 0.0324 & 0 & -0.0105 & 0.0015 & -0.0225 & 0.0128 \\ 0.0174 & 0.0324 & -0.0005 & 0 & -0.0099 & 0.0000 & -0.0131 & -0.0151 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0373 & -0.0105 & -0.0099 & 0 & 0.0004 & -0.0001 & -0.0074 & -0.0057 \\ 0.0022 & 0.0015 & 0.0000 & 0 & -0.0001 & -0.0004 & -0.0013 & -0.0010 \\ 0.0058 & -0.0225 & -0.0131 & 0 & -0.0074 & -0.0013 & 0.0004 & 0.0205 \\ -0.0222 & 0.0128 & -0.0151 & 0 & -0.0057 & -0.0010 & 0.0205 & -0.0006 \end{bmatrix} \qquad (3.11)$$

The residual matrix for PFA is

$$R-\tilde{L}\tilde{L}'-\tilde{\psi}=\begin{bmatrix} 0.0088 & -0.1517 & -0.1437 & 0.0105 & 0.0239 & 0.0097 & -0.2176 & -0.1245 \\ -0.1517 & -0.0044 & -0.0454 & 0.0722 & 0.0027 & -0.0913 & -0.1498 & -0.1221 \\ -0.1437 & -0.0454 & 0.0040 & -0.0391 & -0.0975 & 0.2127 & -0.1306 & -0.0538 \\ 0.0105 & 0.0722 & -0.0391 & -0.0009 & -0.2969 & -0.0222 & 0.0275 & 0.0242 \\ 0.0239 & 0.0027 & -0.0975 & -0.2969 & 0.0031 & -0.0729 & -0.0176 & -0.0206 \\ 0.0097 & -0.0913 & 0.2127 & -0.0222 & -0.0729 & -0.0047 & -0.0408 & -0.1131 \\ -0.1176 & -0.1498 & -0.1306 & 0.0275 & -0.0176 & -0.0408 & -0.0049 & -0.0683 \\ -0.1245 & -0.1221 & -0.0538 & 0.0242 & -0.0206 & -0.1131 & -0.0683 & 0.0030 \end{bmatrix} \tag{3.12}$$

Picture for comparing maximum-likelihood method with principal component factor method is showing on the following line chart:

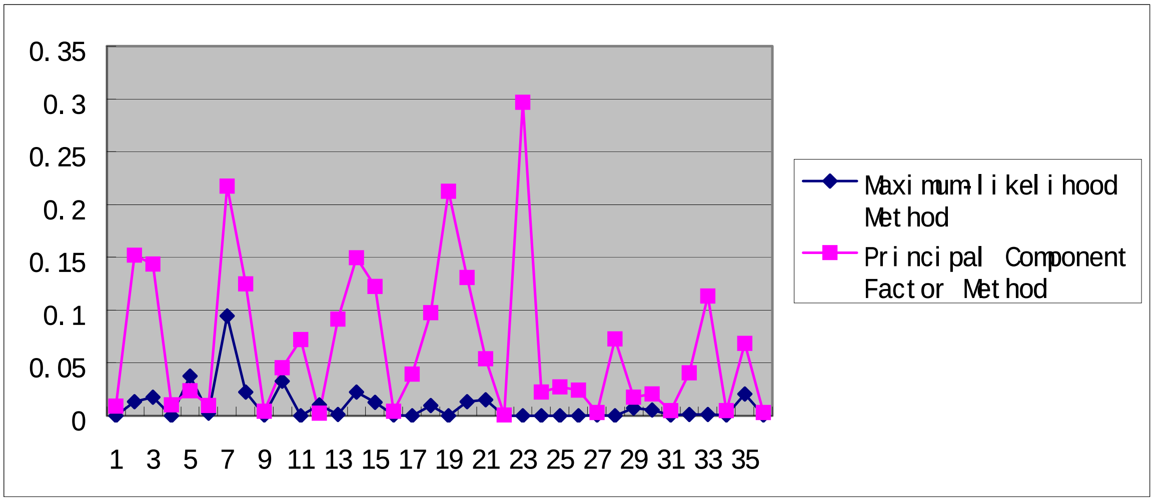

Figure 3.4 Comparing three-factor MLFA with three-factor PFA

Obviously from the line chart 3.6.4, the elements $R-\hat{L}\hat{L}'-\hat{\psi}$ are much smaller than those of the residual matrix corresponding to the principal component factoring of R. On this basis, we prefer the maximum likelihood approach and typically feature it in subsequent steps.

Focusing attention on the maximum likelihood solution, we see that all variables have the positive loadings on F_1**.** They are very close except x_4, Sony variable, which is from the audio & video equipment market. We call this factor the market factor, as we did in the principal component solution. The interpretation of the second factor, however, is not as clear as it appeared to be in the principal component solution.

3.7 Perform Rotation for Principal Factor Analysis

After extracting the initial factors, we also have gotten an unrotated factor pattern matrix. In the factor pattern matrix, the observed variables are assumed to be linear combinations of the common factors, and the factor loadings are standardized regression coefficients for predicting the variables from the factors.

When more than one factor has been retained, an unrotated factor pattern is usually difficult to interpret. Factor patterns are easiest to interpret when some of the variables in the analysis have very high loadings on a given factor, and the remaining variables have near-zero loadings on that factor. Unrotated factor patterns often fail to display this type of pattern. For example, consider the ladings in the column headed"F1" in Table 3.3. Notice that variables1, 2, 3, 7 and 8 do display fairly high loadings for this factor, which is good. Unfortunately, however, variables 4, 5 and 6 do not display near-zero loadings for this factor; the loadings for these three variables range from 0.4 to 0.58, which is to say that they are of moderate size. For reasons that will be made clear shortly,this would make it difficult to interpret factor 1.

To make interpretation easier, we will normally perform a linear transformation on the factor solution called a rotation. We also know an orthogonal rotation. It was explained that orthogonal rotation result in components (or factors) that are uncorrelated with one another.

Next, we analyze the promax rotation, which is a specific type of oblique rotation. Oblique rotations generally result in factors (or components) that are correlated with one another.

A promax rotation is actually conducted in two steps. The first step involves an orthogonal varimax prerotation. At this point in the analysis, the extracted factors are still uncorrelated. During the second steps (the promax rotation), the orthogonality of the factors is relaxed, and they are allowed to become correlated. The interpretation of an oblique solution is more complicated than the interpretation of an orthogonal solution, although oblique rotations often provide better results (at least in those situations in which the actual, underlying factors truly are correlated) [2, 3, 9].

3.7.1 Step 1: Varimax Rotation

During the prerotation step, the SAS system produces a rotation factor pattern similar to that which would be produced if we had specified ROTATE=VARIMAX. The matrix appears on the following output.

PFA	Three-factor Solution			
Variable	Rotated Estimated FactorLoadings			Specific Variances
	F_1^*	F_2^*	F_3^*	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1: IBM	0.678	0.120	0.066	0.52
2: Dell	0.642	0.056	0.243	0.53
3: Apple	0.674	0.304	-0.315	0.35
4:Sony	0.076	0.865	0.080	0.24
5: Novell	0.256	0.686	0.142	0.44
6: Microsoft	0.218	0.199	0.865	0.16
7: AMD	0.710	0.154	0.133	0.45
8: Intel	0.685	0.160	0.320	0.40

Table 3.14: Varimax Rotation Factor Pattern for PFA

3.7.2 Step 2: Oblique Rotation

Inter-factor Correlations			
	Factor 1	Factor 2	Factor 3
Factor 1	1.000	0.406	0.229
Factor 2	0.406	1.000	0.156
Factor 3	0.229	0.156	1.000

Table 3.15: Inter-factor Correlations for PFA

In Table 3.15, look at the section headed “Inter-factor correlations”. Where the row headed “Factor 1” intersects with the column headed “Factor 2” we can find a correlation coefficient of 0.406. This means that there is a correlation of 0.406 between factor 1 and factor 2. At this point in the analysis, we do not know exactly what this correlation means because we have not yet interpreted the meaning of the factors themselves. We will therefore return to this correlation after the interpretation of the factors has been completed.

Let us interpret the nature of the given factor. We begin with the high loadings on that factor. A high loading means that the variable is “measuring “that factor. We review all of the variables with high loadings on factor 1, and determine variable 1, 2, 3 and variables 7, 8 have in common.

Let us look at the rotated factor pattern

PFA	Three-factor Solution			
Variable	Rotated Estimated Factor Loadings (Std Reg Coefs)			Specific Variances
	F_1^*	F_2^*	F_3^*	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1: IBM	0.702	-0.028	0.001	0.52
2: Dell	0.657	-0.100	0.188	0.53
3: Apple	0.702	0.194	-0.395	0.35
4:Sony	-0.115	0.905	0.047	0.23
5: Novell	0.112	0.669	0.099	0.44
6: Microsoft	0.094	0.109	0.860	0.16
7: AMD	0.722	-0.003	0.066	0.45
8: Intel	0.672	-0.002	0.260	0.40

Table 3.16: Promax(oblique) Rotated Factor Pattern for PFA

In table 3.16, the abbreviation “Std Reg Coefs” stands for “standardized regression coefficients” [9]. The loadings that appear in this factor pattern are regression coefficients of the variables on the factors. In common factor analysis, the observed variables are viewed as linear combinations of the factors, and the elements of the factor pattern are regression weights associated with each factors in the prediction of these variables. The loadings in this matrix represent the unique contribution that each factor makes to the variance of the observed variables. Notice that the pattern loading for variable 4 on factor 1 is only –0.115, and for variable 1,2,3,6,7 and 8 on factor 2, and variable 1, 2, 4, 5, and 7 on factor 3 also very small. It does not mean those variables and factors are completely unrelated. Because this is a pattern loading, its small value merely means that factors make a very small unique contribution to the variance in those variables.

The following is a guideline for structured procedure to follow in interpreting a rotated factor pattern.

PFA	Three-factor Solution			
Variable	Reference Structure			Specific Variances
	F_1^*	F_2^*	F_3^*	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1: IBM	0.631	-0.026	0.001	0.52
2: Dell	0.590	-0.091	0.183	0.53
3: Apple	0.631	0.177	-0.384	0.35
4:Sony	-0.103	0.825	0.045	0.24
5: Novell	0.101	0.610	0.096	0.44
6: Microsoft	0.085	0.100	0.835	0.16
7: AMD	0.649	-0.003	0.064	0.45
8: Intel	0.603	-0.002	0.252	0.40

Table 3.17: Reference Structure (Semipartial Correlations) for PFA

The coefficients in this matrix represent the semipartial correlations “between variables and common factors, removing from each common factor the effects of other common factors”

We notice that the size of the loadings in the current reference structure is very closer to those in the rotated factor pattern. It is clear that interpretation of factors as was obtained using the rotated factor pattern.

PFA	Three-factor Solution			
Variable	Factor Structure (Correlations)			Specific Variances
	F_1^*	F_2^*	F_3^*	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1: IBM	0.691	0.257	0.158	0.52
2: Dell	0.659	0.196	0.323	0.53
3: Apple	0.690	0.417	-0.204	0.35
4:Sony	0.263	0.866	0.162	0.24
5: Novell	0.406	0.730	0.229	0.44
6: Microsoft	0.335	0.282	0.899	0.16
7: AMD	0.736	0.300	0.231	0.45
8: Intel	0.730	0.311	0.413	0.40

Table 3.18 Factor Structure (Correlations) for PFA

The structure loadings that it contains represent the product-moment correlations between the variables and common factors. The correlation between items and factor 1 are high except item 4. While the correlation between items and factor 2 is lower except 4. For factor 3, the correlation between item 3 and factor 3 is negative. Now we understand the big picture of how the 8 variables are really related to the three factors.

The structure matrix is generally less useful for interpreting the meaning of the factors if we compared to the rotated pattern matrix because it often fails to demonstrate simple structure. But we sill need to review the structure matrix because the pattern matrix and the structure matrix provide different information about the relationships between the observed variables and the underlying factors. The factor pattern reveals the unique contribution that each factor makes to the variance of the variable. The pattern loadings in this matrix are essentially standardized regression coefficients comparable to those obtained in multiple regression.

3.8 Perform Variamax Rotation for Maximum-likelihood Factor Analysis

3.8.1 Step 1: Varimax Rotation

MLFA	Three-factor Solution			
Variable	Rotated Estimated FactorLoadings			Specific Variances
	F_1^*	F_2^*	F_3^*	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1: IBM	0.558	0.061	0.132	0.67
2: Dell	0.521	0.080	0.180	0.69
3: Apple	0.538	0.158	0.012	0.69
4:Sony	0.174	0.979	0.110	0.00
5: Novell	0.386	0.248	0.162	0.76
6: Microsoft	0.223	0.112	0.795	0.31
7: AMD	0.647	0.100	0.133	0.55
8: Intel	0.653	0.103	0.237	0.51

Table 3.19 Varimax Rotation Factor Pattern for MLFA

3.8.2 Step 2: Oblique rotation

Inter-Factor Correlations			
	Factor 1	Factor 2	Factor 3
Factor 1	1.000	0.340	0.474
Factor 2	0.340	1.000	0.262
Factor 3	0.474	0.262	1.000

Table 3.20: Inter-factor Correlations for MLFA

In Table 3.20, look at the section headed “Inter-factor correlations”. Where the row headed “Factor 1” intersects with the column headed “Factor 2” we can find a correlation coefficient of 0.340. This means that there is a correlation of 0.340 between factor 1 and factor 2. We begin with the high loadings on that factor. A high loading means that the variable is “measuring “that factor. We review all of the variables with high loadings on factor 1, and determine variable 1, 2, 3 and variables 7, 8 have in common.

Let us look at the rotated factor pattern

MLFA	Three-factor solution			
Variable	Rotated Estimated Factor Loadings (Std Reg Coefs)			Specific Variances
	F_1^*	F_2^*	F_3^*	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1: IBM	0.581	-0.035	0.014	0.67
2: Dell	0.522	-0.014	0.075	0.69
3: Apple	0.577	0.082	-0.122	0.69
4:Sony	0.006	0.996	0.006	0.00
5: Novell	0.347	0.187	0.073	0.76
6: Microsoft	0.002	0.011	0.829	0.31
7: AMD	0.674	-0.007	-0.008	0.55
8: Intel	0.650	-0.014	0.107	0.51

Table 3.21: Promax(oblique) Rotated Factor Pattern for MLFA

Notice that the pattern loading for variable 4 and variable 6 on factor 1 are 0.006 and 0.002 which are very small. Also variables 1, 2, 3, 6, 7 and 8 on factor 2 are small. For other variables 1,2,3,7 and 8 on factor 1, and variable 4, on factor 2 are large. While on factor 3, only variable 6 is 0.829 which is very larger. It does not mean those variables and factors are completely unrelated. Because this is a pattern loading, its small value merely means that factors make a very small unique contribution to the variance in those variables.

MLFA	Three-factor solution			
Variable	Reference Structure			SpecificVariances
	F_1^*	F_2^*	F_3^*	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1: IBM	0.495	-0.033	0.013	0.67
2: Dell	0.445	-0.013	0.066	0.69
3: Apple	0.491	0.076	-0.107	0.69
4:Sony	0.0055	0.930	0.006	0.00
5: Novell	0.296	0.175	0.064	0.76
6: Microsoft	0.002	0.011	0.725	0.31
7: AMD	0.574	-0.006	-0.007	0.55
8: Intel	0.554	-0.013	0.094	0.51

Table 3.22: Reference Structure (Semipartial Correlations) for PFA

We still notice that the size of the loadings in the current reference structure is very closer to those in the rotated factor pattern. It is clear that interpretation of factors as was obtained using the rotated factor pattern.

MLFA	Three-factor solution			
Variable	Factor Structure (Correlations)			SpecificVariances
	F_1^*	F_2^*	F_3^*	$\tilde{\psi}_i = 1 - \tilde{h}_i^2$
1: IBM	0.5763	0.166	0.281	0.67
2: Dell	0.553	0.183	0.319	0.69
3: Apple	0.547	0.246	0.173	0.69
4:Sony	0.348	1.000	0.270	0.00
5: Novell	0.446	0.326	0.287	0.76
6: Microsoft	0.399	0.230	0.833	0.31
7: AMD	0.668	0.220	0.310	0.55
8: Intel	0.696	0.235	0.412	0.51

Table 3.23: Factor Structure (Correlations) for PFA

We have known the structure loadings that it contains represent the product-moment correlations between the variables and common factors. From the Table 3.23, we get the correlation between items and factor 1 are high except item 4 and 6. While the correlation between items and factor 2 is lower except 4. For factor 3, the correlation between items and factor 3 are almost close except item 6.

3.9 Compare the Solutions Obtained from the Rotated MLFA and Rotated PFA

	MLFA				PFA			
Variable	Rotated Estimated Factor Loadings			Specific Variance	Rotated Estimated Factor Loadings			Specific Variance
	F_1	F_2	F_3	$\hat{\psi}_i = 1-\hat{h}$	F_1	F_2	F_3	$\hat{\psi}_i = 1-\hat{h}$
1. IBM	0.171	0.504	0.223	0.67	0.639	-0.259	0.058	0.53
2. Dell	0.188	0.499	0.162	0.69	0.631	-0.242	0.133	0.53
3. Apple	0.250	0.398	0.306	0.69	0.609	-0.216	0.478	0.35
4. Sony	1.000	0.000	0.000	0.00	0.496	0.664	0.272	0.24
5. Novell	0.328	0.347	0.094	0.76	0.579	0.441	0.165	0.44
6. Microsoft	0.237	0.614	-0.511	0.31	0.532	0.275	0.690	0.17
7. AMD	0.225	0.566	0.274	0.55	0.701	-0.231	0.013	0.46
8. Intel	0.241	0.631	0.193	0.51	0.739	-0.161	0.158	0.40

Table 3.24: Compare the solutions for MLFA and PFA

The residual matrix for rotated Maximum-likelihood Factor Analysis is

$$R-\tilde{L}\tilde{L}'-\tilde{\psi}=\begin{bmatrix} -0.0001 & -0.0134 & 0.0175 & -0.0000 & 0.0372 & 0.0023 & 0.0060 & -0.0221 \\ -0.0134 & -0.0006 & 0.0324 & 0.0003 & -0.0106 & 0.0015 & -0.0224 & 0.0128 \\ 0.0175 & 0.0324 & -0.0005 & -0.0005 & -0.0101 & 0.0002 & -0.0129 & -0.0150 \\ -0.0000 & 0.0003 & -0.0005 & 0.0000 & 0.0000 & -0.0004 & 0.0004 & -0.0002 \\ 0.0372 & -0.0106 & -0.0101 & 0.0000 & 0.0002 & -0.0002 & -0.0074 & -0.0059 \\ 0.0023 & 0.0015 & 0.0002 & -0.0004 & -0.0002 & -0.0003 & -0.0012 & -0.0011 \\ 0.0060 & -0.0224 & -0.0129 & 0.0004 & -0.0074 & -0.0012 & -0.0001 & 0.0207 \\ -0.0221 & 0.0128 & -0.0150 & -0.0002 & -0.0059 & -0.0011 & 0.0207 & -0.0005 \end{bmatrix} \qquad (3.13)$$

The residual matrix for rotated Principal Factor Analysis is

$$R-\tilde{L}\tilde{L}'-\tilde{\psi}=\begin{bmatrix} -0.0088 & -0.1517 & -0.1437 & 0.0105 & 0.0239 & 0.0097 & -0.1176 & -0.1245 \\ -0.1517 & -0.0044 & -0.0454 & 0.0722 & 0.0027 & -0.0913 & -0.1498 & -0.1221 \\ -0.1437 & -0.0454 & 0.0040 & -0.0391 & -0.0975 & 0.2127 & -0.1306 & -0.0538 \\ 0.0105 & 0.0722 & -0.0391 & -0.0009 & -0.2969 & -0.0222 & 0.0275 & 0.0242 \\ 0.0239 & 0.0027 & -0.0975 & 0.2969 & 0.0031 & -0.0729 & -0.0176 & -0.0206 \\ 0.0097 & -0.0913 & 0.2127 & -0.0222 & -0.0729 & -0.0047 & -0.0408 & -0.1131 \\ -0.1176 & -0.1498 & -0.1306 & 0.0275 & -0.0176 & -0.0408 & -0.0049 & -0.0683 \\ -0.1245 & -0.1221 & -0.0538 & 0.0242 & -0.0206 & -0.1131 & -0.0683 & 0.0030 \end{bmatrix} \qquad (3.14)$$

We also use the figure to compare the rotated MLFA and rotated PFA.

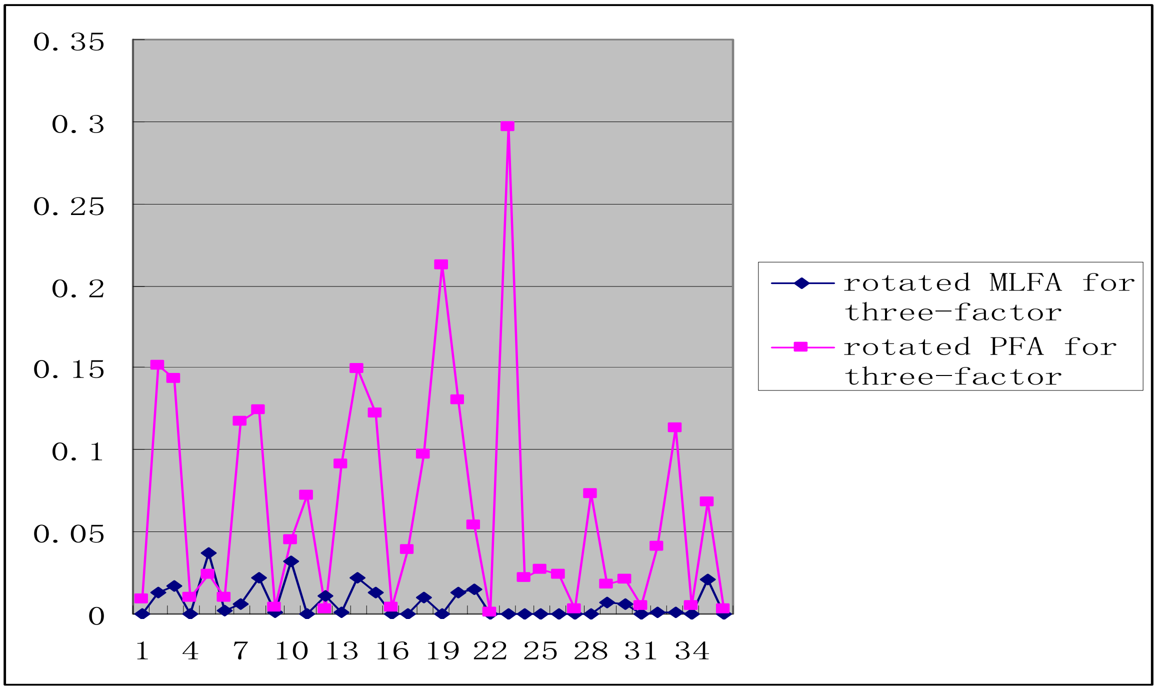

Fig. 3.5 Compare Three-factor Rotated MLFA with Three-factor Rotated PFA

Obviously, rotated MLFA model is better than rotated PFA model.

Last we still should compare the three-factor MLFA and the rotated MLFA for three-factor. The following is the comparison group:

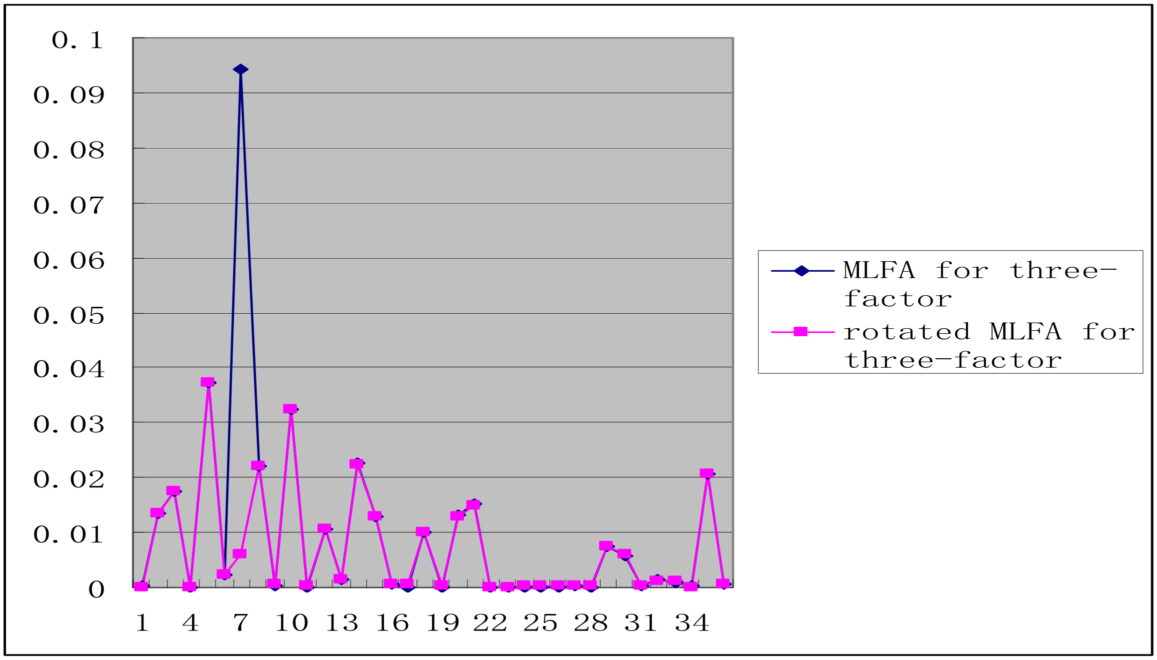

Fig. 3.5 Compare Three-factor MLFA with Three-factor Rotated MLFA

In this figure, we obviously find the values in pink line are smaller than the black line. Hence, we get the rotated MLFA is the best model in this case. All in all, in this project we used several models to extract the underlying factor and determine rotated MLFA for three-factor is the best model.

4. Summary and Conclusion

Factor analysis can be highly useful and powerful multivariate statistical technique for effectively extracting information from large data bases. Factor analysis helps the investigator make sense of large bodies of interrelated data. When it works well, it points to interesting relationships that might not have been obvious from examination of the raw data alone, or even a correlation matrix.

In this project, we propose to apply Principal Factor Model (PFA) and Maximum-likelihood Factor Model (MLFA) to extract the underlying factors. We recover three hidden factors. We also find for this case Maximum-likelihood Factor Analysis is the best model to extract the hidden factors. In some traditional application of factor models, the returns are related to some systematic factors or macro-economic variables. From this project, we can learn how to extract the underlying factors and we can know what the extract hidden factors are. But the financial market nowadays is extremely complex and dynamic, especially due to globalization and many newly introduced indices, such as IT index, it is not an easy task to decide which variables, among so many nonsystematic factor, systematic factors and macro-economic variables, should be included in the model as factors. These models serve as a data mining technique to identify the hidden factors from historical data. And we also made attempts to correlate the factors extracted to some known variables. They are possible to be applied in many aspects in finance. For example, we can perform risk analysis and construct portfolios which are less sensitive to the hidden factors.

References

1. George Casella and Roger L. Berger, *Statistical Inferenc*, Duxbury, Second Edition, ISBN 0-534-24312-6, 2002.
2. Larry Hatcher, A s*tep-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling*, SAS Institute Inc, NC, ISBN 1-55544-643, 1994
3. Richard L. Gorsuch, Factor Analysis, Lawrence Eribaum Associates, Inc., Second Edition, ISBN 0-89859-202-X, 1983.
4. Richard A. Johnson and Dean W. Wichern, Applied Multivariate Statistical Analysis, Prentice Hall, Fifth Edition, ISBN 0-13-092553-5, 2002.
5. T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley-Interscience, Third Edition, ISBN 0-471-36091-0, 2003.
6. Alvin C. Rencher, *Methods of Multivariate Analysis*, Wiley-Interscience, Second Editon, ISBN 0-471-41889-7, 2002.
7. C. Chatfield and A. J. Collins, *Introduction to Multivariate Analysis*, Chapman and Hall, ISBN 0-412-16030-7(cased) 0-412-16040-4 (paperback), 1980.
8. William R. Dillon and Mattew Goldstein, 1984, *Multivariate Analysis Methods and Applications*, ISBN 0-471-08317-8
9. *SAS online help*, http://v8doc.sas.com/sashtml/
10. Philip R. Bevington, 1992, *Data reduction and Error Analysis for the Physical Science*, McGraw-Hill, Inc. ISBN 0-07-911243-9
11. Siu-Minf CHA and Lai-Wan CHAN, *Applied Independent Component Analysis to Factor Model in Finance*, Spinger, Pages 538-544, 2000.

12. Brian S everitt and Graham Dunn, 1991, *Applied Multivariate Data Analysis*, British Library Cataloguing in publication Data, ISBN 0-340-54529-1

13. Pam Gao, Eni and Zhiwei Ren, 2005, *Introduction of Factor Model*

14. John P. Van de Geer, 1971, *Introduction to Multivariate Analysis for the social sciences*, W.H. Freeman and Company, ISBN 0-7167-0932-5

15. W. J. Krzanowski, 1988, *Principles of Multivariate Analysis*, CLARENDON PRESS. OXFORD, ISBN0-19-852211-8

16. K.V. Mardia , J. M .Bibby and J. T. Kent, 1979, *Multivariate Analysis*, ACADEMIC PRESS, INC., ISBN 0-12-471250-9

17. Donald F. Morrison,1967, *Multivariate Statistical Methods*, McGRAW-HILL BOOK COMPANY, ISBN 0-07-043186-8

18. Joseph F. Hair, Jr., Rolph E. Anderson, Ronald L. Tatham and Bernie J. Grablowsky, 1979, *Multivariate Data Analysis*, Petroleum Publishing Company, ISBN 0-87814-077-9

19. *Heywood Cases and Other Abinalies*, http://www.id.unizh.ch/software/unix/statmath/sas/sasdoc/stat/chap26/sect21.htm

Appendix. SAS Code and Log File

1. SAS Code

```
data stockprice1;
        input x1 x2 x3 x4 x5 x6 x7 x8;
        cards;

0.008800655 0.009341317 0.006092798 0.026613966 -0.01459854 -
0.010736764 -0.004520796      -0.006372133
0.015800416 0.007723872 -0.015079243      0.008235919 0.013313609
      0.001854599 -0.013820776      0.033816425
-0.004861901      -0.015680684      -0.002455871      0.042370534
      0.064566929 -0.004431315      0.027485112 0.009308511
-0.004223321      0.010564226 0.039406509 -0.017682263      0.004746835
      -0.00550863 -0.059862188      -0.056461731
0.024915541 0.026367669 -0.028969791      0.021962747 0.044628099
      0.023684211 0.077994429 0.030159414
0.002858655 0.015261446 0.170019938 -0.004703929      -0.081942337
      -0.009679821      0.035079289 -0.039321192
-0.009127151      -0.011622156      -0.005945946      0.003888889 -
0.078321678 -0.103770437      -0.009990485      0.019839595
0.02186964  0.07868765  0.014254386 -0.005250069      0.042274052
      0.022517912 0.22995904  0.014126712
0.039331476 0.069309755 0.044274809 0.038450502 -0.045897079
      0.047908473 0.016052319 0.049415993
0.027005378 0.001428163 0.105252057 0.019602106 0.048104956 0.008291276
      0.080976864 0.045070423
0.02993518  -0.010178117      0.041978022 0.003523194 0.014792899 -
0.008931761 0.111428571 0.033478894
-0.021450813      -0.017772841      0.164874552 -0.028799544      -
0.043847242 0      0.037037037 0.002919708
-0.000115314      0      0.010085337 0.009499136 0.032116788 -0.00920354
      0.005212211 -0.014388489
0.027123061 0.027389444 0.037007241 0.025686448 0.055469954 0.035177721
      0.081320451 0.035767511
-0.01527875 -0.013787282      0.004038772 -0.023074704      -
0.070200573 -0.007997092      -0.022047244      -0.022340942
-0.01175657 -0.01305193 0.035405631 -0.017568716      0.087227414
      0.000727537 0.04013104  0.00097229
0.028083896 0.021560284 0.018166336 0.008862207 0.10880829  0.014016968
      0.120183486 0.025935162
-0.006475159      0.001420455 0.025618632 0.010982659 -0.060064935
      -0.012745812      -0.088628763      -0.089464124
-0.003636364      0.009753299 0.11525974  0.009334889 -0.015974441
      0.009558824 -0.000835422      0.018501388
0.015969491 0.010141988 -0.001297017      0.050566963 0.004815409
      0.006661732 0.082278481 0.002782931
0.005150934 0.010837727 0.035594359 -0.0245142  -0.056060606      -
0.004421518 -0.017761989      -0.053140097
```

-0.041231193	-0.037496476	-0.079158936	-0.034632035
	-0.035087719	-0.047385047	-0.098478783	-
0.066037736
0.026163819 0.009391007 0.053420195 -0.028050491	0.00736377
	0.016410988 0.030528053 0.075430084
0.00676317	-0.007905138	-0.046583851	-0.018717314	-
0.043661972 0.020014556 -0.041139241	-0.002639683
0.004648945 -0.014468559	0.072261072 -0.016779432	0.024531025
	-0.013639627	-0.14652262 -0.144523899
-0.036190257	0.023057216 -0.033783784	-0.038761707	-
0.110397946 -0.024851243	-0.030124427	0.009115078
-0.028029034	-0.024708495	-0.077744807	0.010780962 -
0.113765643 0	-0.01483871 -0.052195824
-0.005662891	0.034760126 0.024004862 0.034829932 -0.002270148
	0.007760141 0.041666667 0.005065123
-0.004421844	-0.008262108	0.070592062 -0.013687601
	0.084975369 0.059021292 -0.055238095	-0.034916201
0.033120146 0.002857143 0.068102849 0.039620536 -0.039053254
	0.031599229 0.041666667 0.017768301
-0.011626594	-0.006810443	0.025659301 -0.027935991	-
0.059020045 -0.0106748	-0.027652733	-0.014360771
0.016756571 0.017908723 0.03504242	0.053428571 -0.102	0.013132484
	0.091994382 0.036297641
0.008332369 -0.002880184	0.001847746 0.007774259 -0.024390244
	0.001160093 0.02667628	0.018860947
-0.020183694	-0.029082774	0.014623172 -0.058807588
	0.031187123 0.003103181 -0.069751844	0.021533812
0.000226835 0.028177113 0.034522886 -0.0390625	0.048523207 -
0.013394566 0.048523207 0.028760202
-0.03407099 -0.032814238	-0.069314079	-0.106769016	-
0.161061947 -0.051198257	-0.130806846	-0.065383218
-0.010836584	0.016968326 -0.050719671	0.058866995 0.01618705
	0.094594595 0.019314642 0.040831758
-0.009020619	-0.007577884	0.059934617 -0.032872797	-
0.074104913 -0.01255887 -0.055882353	-0.033613445
-0.011464968	0.027986151 0.001090909 -0.015247478	-
0.029102668 -0.014313346	-0.025787966	-0.026671408
0.015414466 0.037724551 0.017011834 0.025252525 0.107430618 0.032760687
	0.125080593 0.027027027
0.012551845 0.014580802 0.045630317 0.053992395 0.126008065 0.016240357
	0.01704918	0.033597584
-0.018006431	-0.003933434	-0.061683599	-0.01375	-
0.02745098	-0.029550827	0.016666667 -0.043336945
-0.032659409	0.024488531 0.030665669 -0.047619048	-
0.073569482 -0.036812144	-0.0234375	-0.043523316
-0.000518135	-0.01194487 0.117892977 0.025891549 0.081532417 -
0.006784772 0.024 -0.008561644
-0.008323913	-0.039988239	0.067857143 -0.028476507	-
0.045923149 -0.001505457	0.023890785 -0.026666667
-0.024069802	-0.015629522	-0.026086957	0.023809524 -
0.156521739 -0.000752162	-0.019410977	-0.00464499
0.007782494 0.048558422 0.012769705 0.016296296 -0.054559043	-
0.018094535 0	-0.023963731
-0.002922503	-0.01465311 0.006648936 -0.003690037	0.055205047
	-0.020614828	0.00538358	0.011795544
0.013585291 -0.028753994	0	-0.015500121	0.018473896 -
0.029143258 -0.03002611 -0.038740157

0.027066723 -0.016566695 -0.007042254 0.043994943 0.02807597
0.024092053 -0.085918854 -0.034660991
0.045060849 -0.015743604 -0.012173913 0.072978839 0.075488455
0.005422993 0.078507079 -0.031792758
-0.003713818 0.037328667 0.080827068 0.053744997 0.060263653
0.007650273 0.045760431 0.056281095
-0.014531755 0.015699052 0.024061598 0.019825073 0.006635071
0.008820287 0.010884354 0.025510204
-0.002576766 0.01290129 0.054822335 0.007342144 0.104712042 -
0.005482456 0.03085554 0.025842329
0.004638119 -0.013905325 -0.056965055 0.023444545 0.013800425
0.026641651 -0.066142764 -0.009397278
0.022837599 -0.020289855 0.001918465 -0.009526645 0.022801303
0.025789069 -0.049782203 -0.038629283
0.001104484 -0.002024877 -0.00286944 -0.022125182 -
0.042619543 0.01050175 -0.105731775 -0.04293381
0.021550265 0.001448436 0.031065089 0.013872491 0.135773318 0.023894863
0.069642857 0.037747525
-0.017950139 -0.020431328 -0.054986021 -0.033380884
-0.082340195 -0.015294118 -0.059350504 -
0.014634146
0.02254702 -0.021111111 -0.046222222 -0.030696903
0.230666667 -0.022988506 0.049970606 -0.031591379
-0.013634332 0 -0.017038008 0.027272727 0.277683135 -
0.001530222 0.119078947 0.027921093
0.011988238 0 0.012831858 0.026538349 0.051971326 -0.017662533
0.118469463 0.06016731
-0.009077664 -0.011260643 -0.006593407 -0.08315508 -
0.071547421 -0.080193571 -0.028591851 -0.018319646
-0.037120967 0.026501269 -0.039273649 0.022137196 -0.00166113
0.000691802 0.056646526 0.040420638
0.022396293 0.016041249 0.091747349 0.001094391 0.034364261 -
0.005845942 0.113540791 0.027693347
0.017855138 0.022554189 0.048332528 0.033946252 0.127906977 0.031571479
0.093836247 0.085808581
-0.045347341 -0.020654045 -0.083702391 -0.054560043
-0.091549296 -0.059078772 -0.142744479 -
0.065135413
0.051634724 0.031360947 -0.022510823 0.010813733 0.055762082
0.05716302 0.042763158 0.029287227
0.020126509 -0.008506894 0.026666667 0.096 0.048732943 -
0.001409443 0.004958678 -0.012887496
0.060236557 0.045064378 -0.004865104 0.024279211 0.019880716
0.070135747 0.071744907 0.004197272
-0.011570447 0.015250545 0.082854406 0.016661524 0.151029748
0.011441648 0.13925328 0.04381161
0.014427192 -0.006493506 0.059360731 0.035463259 0.197260274
0.026624902 0.270512821 0.093413174
0.011251236 0.045248869 0.003564155 0.053162853 0.03988604 -
0.001563722 0.074380165 0.062340967
-0.004798819 -0.068913632 -0.052580801 -0.033181522
-0.019553073 -0.022544899 0.009735744 -0.057553957
-0.027289048 -0.004195385 -0.037604457 -0.037871674
0 -0.026775753 0.004189944 0.004415897
-0.002030578 0.002704327 0.032598274 -0.007147296 -0.04787234
0 0.005617978 0.010137875

-0.013782542 -0.008343266 0.050881612 0.004996877 0.041551247
-0.015378982 -0.011111111 0.05655527
0.011197141 0.047113885 0.037637219 0.014896989 0.040345821 0.030566038
0.073025335 0.074585635
0.006353392 0.01136005 0.021356113 0.127188282 0.048338369 0.033944596
0.051724138 0.05590666
-0.017663683 -0.008758211 -0.024479167 -0.050542741
-0.020710059 -0.026585644 -0.060382916 -
0.004837929
0.026223565 0.022385673 0.102181401 0.016551724 -0.055865922
0.068154158 0.041411043 -0.032303371
0.033728919 -0.002551834 0.01574344 -0.016949153 -
0.050397878 0.041402619 -0.059163059 -0.018382353
-0.090754203 -0.000637552 -0.044568245 0.079005121
0.128742515 -0.038195855 -0.048076923 0.045148895
0.032606146 0.063750424 -0.020196507 0.051134179 0.077419355
0.016102395 0.053545586 0.106858054
-0.041914822 -0.054504649 -0.025531915 0.035429936
0.02310231 -0.052796246 -0.090789474 -0.035384615
0.016447744 0.003216468 0.027322404 0.024052181 0.059440559 -
0.029969651 0.034013605 -0.004085802
-0.000228389 0.051403449 0.266435986 -0.004868154 -
0.027210884 0.009961686 -0.006756757 0.028901734
0.043991416 0.01580213 0.082397004 0.030518395 0.109433962 0.034892942
-0.018567639 0.041028446
-0.00450985 -0.013220339 0.017530488 -0.274931798 0.039215686
-0.010980392 -0.061021171 -0.020364416
0.069968254 0.042771297 -0.006060606 0.041022405 0.0625
0.053719008 0.123076923 0.114695341
-0.02525065 -0.020090059 -0.083969466 -0.089106065 -
0.008264463 -0.0354723 0.007042254 -0.018181818
-0.000742115 0.042238267 -0.010981469 -0.05359086 0.075555556
0.017024726 0.1414791 -0.009872242
-0.04770318 -0.03752606 -0.028666667 -0.036940005 -
0.042553191 -0.071509221 -0.13006993 -0.089852008
0.074683544 0.086037736 0.01488498 0.055586283 0.035242291 0.068785197
0.239168111 0.101921957
0.014120668 -0.008604564 0.017205781 -0.007138935 0.031818182
0.055178268 0.099047619 0.072454716
-0.000641437 -0.008531157 -0.031978681 -0.040821701
-0.153846154 -0.005907173 -0.043715847 -
0.071884058
-0.025015635 0.004470939 0.000666667 -0.028154594 -
0.209726444 -0.037758831 0.018552876 0.02739726
0.03227889 0.041521149 0.022494888 -0.002552974 0.071661238 -
0.490062112 0.056862745 0.039628483
0.004539559 0.104113111 0.036749117 0.015292898 0.069686411 0.036925719
0.011904762 0.073089701
-0.014066496 -0.021793797 -0.014623955 -0.034292866
-0.114197531 -0.01854193 -0.038167939 -0.038952746
-0.010001266 -0.021328958 0.04057971 -0.043800862 -
0.018181818 -0.047943831 -0.034990792 -0.011987382
-0.028413284 -0.018123238 -0.021276596 -0.017865538
-0.057142857 -0.031286436 -0.099502488 -0.02998776
-0.072764599 -0.085451197 -0.042119565 -0.007929104
-0.090909091 -0.079756795 -0.186234818 -
0.061997704

0.073851806 -0.023029867 -0.012080537 0.017560513 0.087570621
0.039598438 0.067723343 0.053204353
0.055455016 0.028497409 0.059743954 0.012980769 0.089230769 0.015480461
0.092913386 0.008536585
-0.030454944 -0.013868613 -0.005657709 0.004830918 -
0.055232558 -0.001319759 -0.040785498 -0.035861258
-0.002625 0.028914758 -0.043948614 -0.005046864 0.124183007
0.010285714 -0.095628415 -0.032423208
-0.028182702 -0.070506108 -0.010702341 -0.0414651 -
0.120689655 -0.053542455 -0.097410604 -0.06039551
-0.052922227 0.001748252 -0.035483871 -0.021195039 -
0.025210084 -0.03831484 -0.098888889 -0.103927203
0.029491887 -0.000349528 -0.031855091 0.027333797 0.056213018
-0.009275163 0.311953353 0.041396509
0.055243095 -0.040576794 0.003761755 0.000927429 0.027355623
0.026988887 0.058641975 0.066489362
0.031189586 0.001007049 0.006944444 0.010070258 0.166666667 0.028856624
0.062295082 0.035812672
-0.034950249 0.027950311 -0.031784841 -0.039370079
0.040590406 0.039622642 -0.062980031 -0.008196721
0.07832618 -0.002066116 0.060959792 0.014840183 0.092741935
0.006074412 0.068965517 0.103074141
0.004175084 0.019305019 0.075313808 -0.001595623 0.012244898 -
0.008842897 0.384090909 0.147302905
0.16160826 0.078758046 -0.011716058 0.038343195 0.155660377
0.087579292 0.170212766 -0.049934297
0.129328622 0.066639742 0.034212402 0.025485437 0.024154589 0.116518163
0.071225071 0.110138585
-0.062292909 0.018930041 -0.046875 -0.026465028 0.014705882
-0.032707182 -0.388501742 -0.062243502
-0.055694618 -0.021345147 -0.010087424 0.007619048 -
0.051162791 -0.046565529 -0.103125 -0.018791946
-0.118344828 -0.073507463 0.049400141 -0.01777362 -
0.096638655 -0.009392611 -0.112343967 -0.070492826
-0.009562842 0.026819923 -0.014603616 0.022721837 -
0.081081081 0.001882058 -0.103233831 -0.011713933
-0.028920138 -0.019534185 -0.025084746 -0.039071478
-0.003846154 -0.025672372 -0.091525424 -
0.026994601
-0.062437811 -0.044165171 -0.062301335 -0.043105344
-0.061371841 -0.060130218 -0.084798345 -
0.071826281
0.013232514 0.011623683 -0.005060089 0.033409091 0.601156069
0.0444 -0.026183283 -0.042133333
0.104691633 0.075390625 0.054 -0.00990099 -0.049450549 0.039068994
0.124575311 0.049832027
0.058190925 0.060920017 0.038062284 0.040505736 -0.066666667
0.083539743 0.207934337 0.068821065
0.022289157 0.057869356 0.007670851 -0.044519016 -0.101382488
-0.020727674 -0.065217391 -0.061763054
-0.077777778 -0.090510367 -0.04144385 -0.034557235 -
0.088235294 -0.084947538 -0.078916372 -0.045040214
0.040312094 0.001997603 -0.145631068 -0.090373281 -0.0703125
-0.044350174 -0.067032967 0.036687048
-0.058367347 -0.036937284 -0.065635005 -0.042693248
-0.120274914 -0.054512306 -0.025695931 -
0.079324463

0.020833333 -0.005738332 0.057562077 0.001318267 -0.093457944
0.00274223 -0.03909465 0.069512863
0.047272727 0.090075063 0.051632047 0.091918569 -0.027272727
0.046289212 0.208955224 -0.02455953
-0.09741368 -0.067651633 -0.161691542 -0.072831268
0.054313099 -0.053755656 -0.1625 -0.119830827
-0.027203065 -0.02130898 -0.060747664 -0.054954955 -
0.036923077 0.062908811 -0.021406728 -0.032727273
-0.026724674 -0.02122905 -0.081545064 -0.044914817 -
0.052478134 0.021017482 -0.141732283 -0.203475742
-0.03188929 -0.013592946 -0.035196687 -0.021717172 -
0.075471698 -0.044123169 -0.01888412 -0.036287509
-0.03022523 -0.026118068 -0.034386246 0.010032307 -
0.043814433 -0.049437801 -0.045081967 -0.081410256
0.075426707 0.170435511 0.072469983 0.074155251 0.193846154 0.119480519
0.105072464 0.15512773
-0.02567865 -0.018092105 -0.008081667 0.013513514 -
0.092178771 0.009887006 -0.000904977 0.016942771
-0.034588596 -0.060278207 0.021729683 -0.026666667 -
0.002785515 -0.037669903 -0.013392857 -0.055476529
-0.048202247 -0.053401609 -0.07886309 0.042253521 -
0.084183673 -0.09965035 -0.078947368 -0.065780731
0.039719626 0.046707504 -0.003192338 0.046991742 -0.02244389
0.022706955 -0.075285171 0.060232476
-0.119794344 -0.011355034 0.012934519 -0.012427184
0.089673913 0.001073922 -0.025925926 -0.055241265
-0.064903846 0.011872846 0.045204901 -0.003868472 -
0.053984576 -0.073619632 -0.082256968 -0.011838211
-0.015151515 -0.016942771 -0.01743462 -0.011472275
0.04851752 -0.002315964 -0.003387534 -0.005884276
-0.011143365 0.032659409 -0.034468938 -0.053393665 -
0.0725 -0.032645223 0.010958904 -0.036231884
0.016176611 -0.110342442 0.011759935 -0.023679095 -
0.036144578 -0.022830336 -0.099321407 -0.071115013
0.020093186 0.093419062 0.051599147 0.183891213 0.066838046 0.042040085
0.069966997 0.102969658
0.046419502 0.112794613 0.031222515 0.023335474 0 0.058285911
0.178988327 0.049102608
-0.043152882 -0.071875 -0.048535565 0.041471572 -
0.017676768 -0.037190769 -0.115013774 -0.085475379
-0.020001905 -0.024761905 -0.005409904 0.057783019 -
0.081206497 -0.006924979 -0.047868852 -0.007072571
-0.02787037 -0.020522388 -0.01556739 -0.052725648 -
0.175908222 -0.032077881 -0.034810127 -0.062013268
-0.011713031 -0.05766526 0.049892473 -0.035344828 0.011605416
-0.017868339 -0.030674847 0.029394299
-0.043501094 0.012099644 0.048714479 0.03803132 0.015717092 -
0.034795764 -0.030339084 0.005973716
-0.050369878 -0.031034483 0.053206651 -0.059936909 -
0.02303263 -0.036583588 -0.122193211 -0.030969609
-0.042117834 -0.019939169 -0.111439426 -0.040750454
0.072016461 -0.004208999 -0.0425 -0.034646549
0.021969081 0.06707537 0.056174766 0.071783784 0.042918455 0.015176072
0.221001221 0.110111663
0.007377049 0.005438724 0.068095238 0.042605951 0.002150538 0.004885993
0.032136106 -0.005245295

0.007431874 -0.019203414 0.029916626 -0.03982684 0.037946429
0.0014828 -0.091585575 -0.025849113
0.005813953 -0.028334485 -0.09538598 -0.04742268 -0.028199566
-0.005749668 -0.021288515 0.000902527
0.041612596 0.036161833 0.058215962 0.016771488 0.082159624 0.056377511
0.316371681 0.017758726
0.002080624 0.059559939 0.07358871 0.024704619 0.067669173 -
0.007726781 0.01497006 0.0515132
0.007423581 -0.009022556 0.045861887 0.007575758 -0.0025 -
0.01581749 -0.021245421 0.014038524
0.003681627 0.033009709 0.013896312 0.16080402 0.017811705 0.008280939
0.101694915 0.098637016
0.041826484 0.033306581 0.007539041 0.002014099 0.034210526 0.062052117
0.011428571 0.060076046
-0.014933429 -0.022745098 -0.005356186 -0.015125217
-0.054726368 -0.012861736 0.184719536 0.017014695
0.082903069 0.06029106 0.020218579 -0.016341463 0.05511811
0.074265976 0.125136017 0.070807453
0.017949226 -0.003315375 0.016102165 0.024743814 -0.01038961
0.026959915 -0.116346154 -0.034772182
0.028769639 0.069592199 0.115861214 0.135035461 0.010498688 -
0.023215523 0.209302326 0.139344262
0.068687309 0.217485159 0.040618956 0.061746988 0.040983607 0.12800469
0.055214724 0.074363992
0.013480663 0.114251353 -0.013986014 -0.096598639 0.180645161
0.029370348 -0.123655914 0.059067358
-0.063050005 -0.228306265 -0.089699074 -0.114030858
-0.209183673 -0.102707581 -0.191304348 -
0.254538432
-0.033616808 0.007951356 -0.068463612 -0.076169265 -
0.138461538 -0.028921998 -0.151291513 -0.074034335
-0.065800542 -0.075259516 -0.001077006 -0.062826132
-0.040084388 -0.080580177 -0.096666667 -
0.038514443
0.022946744 0.005217391 0.027670172 0.000208768 0.102325581 0.002747253
0.042390549 0.035981475
-0.003430205 -0.145299145 -0.049947424 -0.03815261 -
0.113402062 -0.055555556 -0.112276373 -0.062771285
-0.029857645 -0.041325258 -0.024615385 -0.022954679
-0.03960396 -0.020481387 -0.157922078 -0.054608586
0.033237822 0.02221413 0.028481013 -0.017919075 -0.005905512
0.021689323 0.088800905 0.084188912
-0.009460738 -0.015417712 -0.051051051 -0.115541922
0.007936508 -0.053628216 0.04 -0.023722018
-0.026075739 -0.002146691 -0.195975855 -0.010455312
0.003984064 -0.030277544 -0.193548387 -0.008612123
0.019061033 0.090518923 0.128007263 0.056099733 0.001996008 0.079927339
0.013461538 0.061906437
-0.061674009 -0.019885277 -0.052473118 -0.146656535
-0.119507909 -0.095068493 -0.280276817 -
0.028034188
0.005581643 0.117043998 0.044474394 -0.051599885 0.168377823
0.060584048 0.171463316 0.063249727
-0.006426056 -0.037417763 0.089041096 -0.022954513
0.02742616 0.011908262 -0.104862119 -0.006141618
-0.021533161 -0.053328143 -0.041275797 -0.060590025
0.017167382 -0.070638065 -0.113826367 -0.097489403

0.02843476 0.0438846 0.020584011 -0.045460285 -0.016877637
0.040517486 0.050675676 0.067153793
-0.041680815 -0.088518519 -0.082161687 -0.042558336
0.002114165 -0.008038358 -0.073262367 -0.012371134
0.003065395 0.08914885 -0.032724182 0.01922366 -0.042510121
0.041415773 -0.016928286 0.011821975
0.050353278 0.012663399 0.0297593 0.006823821 -0.057251908 -
0.018876081 0.173771676 0.029348604
-0.034955981 -0.052631579 -0.112621359 0.007877954 -
0.02962963 -0.019081272 -0.106808648 -0.095207254
-0.00292599 -0.006153846 -0.017175573 0.046043165 0.139240506
0.054082241 0.033 0.023194168
0.01193068 -0.136786189 0.046325879 0.021239647 0.004237288 -
0.027246377 0.034482759 -0.069380204
0.193659044 0.078796562 0.116859946 0.043635857 0.102803738 0.10968157
0.167002012 0.153271693
-0.017866258 0.12535268 0.088878096 0.013135593 0.086294416
0.106602598 0.239401496 0.190516511
0.018402994 -0.034254574 -0.067059357 -0.020069204 -
0.212 0.027427318 -0.244536549 -0.102242493
0.028553096 -0.06377551 -0.040434783 -0.008916324 -
0.075785582 -0.033062235 -0.098505435 -0.086775425
0.037846837 0.158294639 0.172273191 0.076809453 0.06916996 0.036656891
0.258119658 0.033728023
-0.092557156 0.013692769 -0.031111111 -0.041755131 -
0.165016502 -0.037572764 0.004291845 -0.053328804
-0.029423265 0.0593835 0.051948052 0.026591107 0.037671233 0
0.07175713 0.004435346
-0.016346154 -0.051182796 0.023391813 -0.043502432 -
0.074484945 -0.001057269 -0.016289593 -0.021042084
-0.095652174 -0.010638298 -0.01 0.00418702 -0.136798906
-0.009771419 -0.105625253 -0.129144852
0.026785714 0 -0.006276151 -0.006379143 -0.056774194
-0.030779638 0.01395158 0.026268657
0.015688764 -0.067090115 -0.072744908 0.023417542 -
0.067388688 -0.027627035 0.020092089 -0.061361726
-0.03432875 -0.049433962 0.054192229 -0.042142469 -
0.015402844 -0.04984375 0.117922321 -0.005295429
0.026426966 0.034348165 0.003076923 -0.040813665 0.080665813
0.049180328 -0.028636364 0.069129917
0.185907686 0.158227848 0.134380454 0.012409241 0.033068783 0.140186916
0.235261089 0.044506692
-0.002021277 0.164210526 0.050091631 0.004908464 0.423728814
0.08894769 0.113125 0.002183406
0.105882353 0.089449541 0.100134409 0.084604317 0.017241379 0.132549562
0.158580739 0.066533599
-0.04494382 -0.050626021 -0.008 0 -0.125628141 -
0.065891473 -0.034940601 -0.087431694
0.013551987 -0.075490689 0.06685633 -0.061571699 -
0.081538462 -0.055905672 -0.116121062 0.01541307
-0.094742268 0.070581897 -0.066401062 -0.058000509 -
0.004594181 -0.096436444 0.008094645 -0.045882353
0.014432127 0.006507592 -0.117233294 0.05445279 0.153710247 -
0.038672082 0.04015544 -0.003808966
-0.043225936 -0.243331966 -0.116519938 -0.015709571
-0.085621971 -0.190305977 -0.258049015 -0.22325899

-0.019619384 -0.022854852 0.043783784 0.001719122 -
0.201290323 0.012742543 0.012159533 0.058795181
0.096129032 0.084347826 -0.029380902 -0.063529412 -
0.057177616 0.02508535 0.024925224 0.121621622
-0.071114662 -0.293611794 -0.143370787 -0.064418955
-0.070135747 -0.012893773 -0.155368421 -
0.194601654
0.06863059 0.165354331 0.198814655 -0.008614748 0.036342321
0.008273009 0.161937378 -0.009486848
-0.011187335 -0.017580872 -0.048205128 -0.102474227
0.030193237 0.03834944 -0.075949367 0.077101719
-0.131212177 0.041376785 -0.116047144 0.019764508 -
0.007194245 0.212837209 0.011431184 0.06636949
-0.059827586 0.07902015 -0.005858495 -0.0275023 -
0.070234114 -0.032577394 -0.005909091 0.011016525
0.030012431 -0.17851347 -0.138252427 -0.03100852 -0.097585513
-0.078759741 -0.083333333 -0.038979788
-0.090821022 -0.142737896 -0.506610462 -0.09875
0.028985507 -0.046482213 -0.094339623 -0.13308302
-0.00904 0.005314685 -0.055042549 0.061007958 -0.025227043
-0.014644026 -0.045045045 -0.166550765
-0.034749035 -0.080504115 -0.061990489 -0.034306102
-0.078139535 -0.073871014 -0.12927518 -0.120085666
-0.030906234 -0.09707385 -0.071878941 -0.048485376 -
0.118129614 -0.012537398 -0.150133333 -0.115904788
0.035891473 0.114677712 0.116704805 0.096706444 0.102169982 -
0.006088927 0.086956522 0.013709899
0.071072733 -0.003096774 0.1362 0.070187985 0.119433198 -
0.005352113 -0.508896797 0.033730159
-0.00149229 0.028124171 0.048437828 0.030315789 0.07508161 -0.01987852
0.216450216 0.105782793
0.040994218 -0.089173514 0.006542845 0.025696394 0.003275109
0.048032407 -0.079681275 0.019980818
0.0363116 -0.052872511 -0.019250673 0.024104379 0.014396456
-0.008179079 -0.116197183 -0.515526988
-0.025620915 -0.165903016 -0.098020911 -0.071267201
-0.124151309 -0.036232886 -0.113829256 -
0.065561908
0.104002309 -0.01281568 -0.07158953 -0.04172407 0.054192229 -
0.083987839 -0.109777778 -0.057945327
-0.010660575 0.046754784 0.059698751 -0.009744689 0.060737527
-0.037317073 0.097560976 0.052975379
-0.041073384 0.02798621 0.039327988 0.088113668 -0.003243243
0.025 0.061488673 0.04203755
-0.020648968 0.039418212 0.01334881 0.002657878 0.027777778
0.029733556 -0.11965812 -0.005134693
-0.01218543 -0.001263158 -0.433161531 0.042216066 0.003344482
0.07070011 0.063636364 0.066000317
-0.053805665 0.054150022 -0.047624021 -0.114153907 0
0.054497893 -0.070422535 -0.007870297
0.09999081 0.040406373 0.034464131 0.113442623 0.044237485 0.037701704
-0.015201953 -0.053133617
0.017486441 0.021944313 0.071668403 0.008931525 0.077791719 0.079264323
0.217837838 0.14145968
0.004697482 -0.097145292 -0.081170213 -0.545276775 -
0.120309051 -0.055640947 -0.117050471 -0.002630016

0.019149751 -0.05894146 -0.126556402 -0.089315068 -
0.102081269 -0.054497893 -0.02193955 0.024956522
-0.031797534 0 -0.048620934 -0.011554432 -
0.082727273 -0.032480315 -0.068586957 -0.067844695
-0.032556054 -0.004987034 -0.088183137 -0.018038381
-0.439633214 0.019641577 0.051428571 -0.027127198
0.072115385 0.003804565 0.043661142 -0.045961945 -0.03962818 -
0.116417532 0.118210863 0.099159227
-0.00952381 0.048498845 0.062572629 0.012847966 0.069037657 0.065029682
0.185606061 0.044072398
-0.147173489 -0.136981337 -0.150891841 -0.141544118
-0.211546392 -0.167752077 -0.125827815 -
0.192310504
0.040128411 0.023173897 -0.029894706 -0.028987577 -
0.152986378 -0.161788235 0.279661017 0.036910717
-0.018653623 -0.044294826 -0.020765737 0.158430172 -
0.128993003 -0.04870624 0.092592593 -0.05120092
0.096545455 0 0.10952 -0.025195517 0.114237288 0.123981081
0.08 0.070680628
0.045130641 0.101268293 -0.005964215 0.017682051 -0.059611093
-0.016138614 -0.038461538 0.080622348
-0.025462963 0.108108108 -0.017578125 -0.185980497 -
0.023349938 0.050769871 0.253012048 0.0078826
0 0.121212121 0.159735435 0.085399449 0.001871491 0.052677691
0.024691358 0.052980132
-0.04 0.029705442 -0.007910112 0.031789962 -0.060375147 -
0.039448769 -0.058139535 0.074784094
-0.024876484 0.086225597 0.022988506 0.023973652 -0.122202213
-0.048829298 -0.054945055 -0.004722773
-0.002162256 -0.042077922 0.006944444 -0.03041105 0.079722222
-0.062124625 0.179979253 0.010692124
0.036392972 0.033557047 0.062782917 0.107998025 0.19244783 0.084580153
0.093900709 0.114361702
-0.0818107 -0.148571429 -0.087054173 0.034297872 -
0.158349596 -0.053012048 -0.069429778 -0.040228711
0.015716435 -0.005681818 0.108223815 0.035242291 0.064707628 -
0.075723831 -0.061911838 -0.049679798
0.053920705 -0.04741286 0.009447236 -0.031487328 -0.084759576
0.007268485 0.242461538 0.256829268
0.052192454 -0.094313725 -0.032195312 -0.176892011 -
0.078367551 -0.045481799 0.123013131 -0.00376625
-0.006904806 -0.032074397 -0.006666667 0.204016913
0.164092101 -0.005875341 0.011180992 -0.009744947
-0.012545455 0.156496927 0.035 0.170096972 0.393582453 0.019002169
-0.006594932 0.012917378
0.009174312 0.096246391 -0.029126214 0.079066788 0.17574021
0.227761798 -0.015042735 0.135778547
-0.025654778 -0.082763187 -0.104347826 0.022881171 -
0.002857143 -0.02340824 -0.071428571 -0.081840132
0.065428571 0.055193293 0.209762255 -0.015166183 0 0.054872695
0.1375948 -0.019439252
0.010198191 0.039457758 0.028342709 0.052768656 -0.101796407
0.059534884 0.021017699 0.004757731
0.084176489 -0.010538922 0.020083867 0.000679887 0.25093633 -
0.035766342 0.021084337 0.048300302
0.062271468 0.024539877 0.026157853 0.071710486 -0.077037037 -
0.025884666 0.249882353 -0.075027316

-0.081424936 0.015449788 0.102221667 0.030923318 0.009471585 -
0.010803803 0.07269056 0.06366219
0.045880349 0.006268806 0.083581282 0.074382944 0.193337299 -
0.001402525 0.059925094 0.054466231
-0.129136924 -0.068441953 -0.008315451 -0.018159007
-0.072807501 0.052577788 0.064350797 0.036263581
-0.049603524 -0.059120879 0.137278829 -0.033135415 -
0.123731271 -0.072466821 -0.05387931 -0.063680803
-0.036093418 0.069833059 0.062216461 0.007266881 0.107009096
0.055123361 0.075942029 0.010008006
-0.058 -0.026550698 -0.049584232 0.045659337 -
0.090953307 -0.010556411 -0.010894495 -0.009647152
-0.002951264 -0.106178396 -0.155965688 0.011908002 -
0.073456512 -0.057030278 -0.1413097 -0.105766958
-0.071333333 -0.013720743 -0.006456612 0.112406328 -
0.077722361 0.015157895 -0.06362379 -0.031475335
0.047648611 0.002427184 0.053605442 0.003341688 0.005012531 -
0.009074789 0.032857143 -0.021722092
0.039193548 0.050573736 0.135135135 0.025706941 0.027027027 0.028096515
0.066531234 0.076024096
0.018480493 0.088092486 0.093934786 -0.056380476 -0.092643052
0.118507857 0.07890411 0.038278709
-0.013131231 0.043677606 -0.014485514 0.052615289 0.017425743
-0.015586256 -0.036430834 0.002382445
-0.001052632 0.040944486 0.109551081 0.078617927 0.049459684 -
0.005051692 0.090385723 0.114449413
-0.017423821 -0.026174168 -0.02801221 -0.04144 -
0.065631068 -0.008041021 0 0.037101449
0.007050717 -0.002926829 0.044644532 0.070022256 -0.037383178
-0.049196676 0.077543424 0.07493379
-0.083963303 -0.052023121 0.004711647 -0.068198134 -
0.12838058 -0.092417538 -0.082004556 -0.043653159
-0.00822536 0.010277972 -0.046198095 0.097810858 0.03612424
0.066380697 0.032941176 0.013132075
0.03879017 0.15297603 0.201252429 0.014389767 0.057857143 0.013586957
-0.011052938 0.043307087
0.074155296 0.005143476 0.097653472 0.08218783 0.12 0.083117495 -
0.003478261 0.148074489
0.019627329 0.00653951 -0.104816465 0.007261813 0.01010101 -
0.000705882 -0.08 0.00673462
0.056338028 0.073413279 0.014857881 0.049486841 0.067270375 0.088069636
0.115407496 0.009184423
-0.014568966 0.009149941 -0.035113235 0.095879733 -
0.005148005 -0.018099548 -0.062987737 0.023500658
0 -0.016260163 0.092374035 -0.037410226 -0.008085106
-0.014004214 -0.03027027 -0.01609323
-0.49648407 -0.076923077 0.002730997 0.018338609 -0.04549147
0.040355854 -0.097560976 -0.051578947
-0.03707419 -0.094197621 -0.009914376 -0.004022614
0.012751954 0.008976194 0.093333333 -0.017241379
0.101265823 0.024881811 -0.032693984 -0.012136183 0.034468085
-0.027700481 0.064131669 -0.048244175
0.038529566 -0.024277737 -0.002608696 0.008120398 0.056179775
-0.027671873 0.067878788 -0.004085635
0.047259074 -0.042093023 0.173768819 -0.037114262 -
0.077911314 -0.054534884 -0.036214953 -0.009068826

0.172379387 0.124183007 0.105812641 -0.036270471 0.090867993 -0.0071577 0.045815516 0.07860262
-0.085502657 -0.121900826 -0.035646259 -0.031620938 -0.110932476 -0.080954907 0.04734485 -0.562342329
0.05259887 0.057538237 0.019134775 0.073644626 -0.019314151 0.016830295 0 0.082237114
0.026801253 0.087381204 0.084511278 0.015918497 0 -0.479649694 -0.023125 0.035821407
0.022662553 -0.058881988 -0.007462687 -0.038959714 -0.03130966 0.040539751 -0.00373599 -0.019411765
-0.053033708 -0.04598246 0.009340163 0.016167478 0.085370908 0.068668456 -0.026666667 0.006342495
-0.002130284 -0.509760632 0 0.177856445 0.199304175 0.033884084 -0.105206074 0.031669866
0.05083947 0.074138792 -0.046538351 0.100483611 0.03871967 0.0320389 0.031897034 -0.044355511
-0.010953796 0 -0.063995698 -0.007466667 0.057891862 0.016108291 0.014188422 -0.063407778
-0.006483357 -0.108490041 -0.013266118 0.04427736 -0.026581606 -0.063391442 0.003416856 0.012332016
0.040662651 -0.105236957 0.038006059 -0.009515929 0.041528239 -0.0140625 0.0446163 -0.008309815
-0.094133697 0.0044 -0.118475358 -0.004257072 -0.113837095 -0.085714286 -0.267218832 -0.094934014
0.019471488 0.204819277 0.062967742 0.036288601 0.062011464 0.12 0.064007421 0.093575419
-0.028063156 0.050632911 -0.061970467 0.02092111 -0.04336989 0.043405676 -0.036209209 -0.048013
-0.013968863 0.015293664 -0.082 0.004524227 0.028717949 -0.000867361 -0.201071429 0.043873853
0.01724699 0.063123378 0.099169516 -0.045150502 0.075565361 0.080683539 -0.034482759 0.093876518
-0.018942216 -0.018374464 0.043057325 -0.006094183 0.054682955 -0.021587302 0.035714286 -0.05152
0.095476801 0.098570797 0.115373686 -0.013661202 -0.003478261 0.028590088 -0.002138275 0.041666667
0.021190476 0.010120554 0.042666667 0.002602383 -0.03144301 0.028432836 -0.070553163 0.030573686
0.02283105 0.018030303 0.030534351 -0.02 -0.007246377 0.052053074 -0.037615556 0.001117703
-0.033823529 0.034320639 -0.065887051 -0.019736842 0 -0.005388099 0.074683111 0.057363636
0.06163742 -0.008699705 -0.007080147 0.042810099 0.067221892 0.127090301 0.06377551 -0.018383009
0.017085874 0.006725055 -0.01064724 0.095445664 0.174703005 0.032909091 0.025794393 0.080096386
0.050020008 -0.02664028 -0.0622701 -0.0556423 -0.02120383 0.006312323 0.094517185 0.087070411
0.00969697 0.002900763 0.025047132 0.072951569 -0.016812374 0.03249268 0.083333333 0.070075121
0.049025148 0.124463519 0.045915493 -0.111261505 0.066714491 -0.004700573 0.293577982 0.022234957
0.041341768 0.027518081 -0.032433906 0.074305657 0.109872611 0.012469065 0.014543339 0.041791045
0.067787291 0.073470934 0.04440649 0.029799341 0.122430742 0.084546299 0.131665569 -0.000715905

0.020030446 -0.157600893 0.001996577 0.059327411 -0.06281407 -
0.069631195 -0.189866667 -0.000715393
-0.065093633 -0.051014229 -0.095225806 -0.102888857
-0.030056864 -0.079073059 -0.082681018 -
0.050277432
0.070655225 0.141326883 0.054421769 -0.044075636 -0.010450161
0.072980924 0.1393534 0.063975904
-0.014775601 0.007660167 -0.023385597 -0.022213355
0.070567986 0.010743405 0.04 -0.022839652
0.060232889 -0.47543379 0.071164247 0.03496696 0.14033366 0.078969158
0.210526316 0.085495208
-0.026028068 -0.077894737 0.027493419 -0.006700397 -
0.052093023 -0.081995249 -0.061882818 0.016233766
-0.04152655 0.010638298 -0.204883721 -0.08838195 -0.022727273
-0.048544567 -0.137911464 -0.093477749
0.019940975 0.112478697 0.061728395 0.018149511 0.005484461 0.061103118
-0.021111111 -0.014502843
-0.029117943 -0.015198135 0.109589041 -0.024634811
0.012025902 -0.015301785 0 -0.005767678
-0.025433962 -0.012339994 0.053999422 -0.044023669 -
0.050087873 -0.037020193 0.043478261 0.026646139
0.06639839 0.03547249 -0.001729605 -0.081122227 -
0.075548335 -0.034004042 0.061538462 0.016614496
0.033779849 -0.107944879 -0.059381779 -0.043279234 -
0.061737805 -0.035017806 -0.018719807 -0.000721848
0.014261603 0.166154151 0.150343107 0.10967444 0.044585987 0.041964838
0.055449331 0.042257053
0.028735133 0.085495854 0.105517241 0.018339995 0.009646302 0.055384615
-0.063283582 0.088737201
0.019380531 0.005413013 0.028733593 0.043424926 -0.0048 0.0269054
-0.056338028 -0.040853738
0.064832265 0.089911504 0.041759054 -0.010559534 -0.019607843
0.102967578 0.014285714 0.091
-0.087139785 0.022686135 -0.03769559 0.044100874 0.0625
0.10425656 0.037344398 0.022793688
-0.022040885 -0.019985809 0.046520283 -0.03343949 0.079136691
-0.005797101 -0.078645549 -0.019624696
0.011659574 0.026089067 0.009391435 -0.020163226 0.059047619
0.01697913 -0.06390593 -0.022816349
-0.03641135 -0.03749124 -0.044851094 -0.017568683 -
0.040219378 -0.008765778 -0.024925224 -0.038621989
-0.024011526 -0.048666667 -0.057171854 0.006646884
0.041904762 -0.043381038 -0.098426966 -0.074710497
0.041166667 -0.019073569 -0.02890933 0.059348673 0.05 0.04303207
-0.142251349 -0.044610992
0.026781894 0.091482275 0.087142857 -0.049820789 0.006036217 -
0.043182325 -0.058780842 0.023623965
-0.004260032 0.101559429 0.002147459 -0.032258065 0.016359918
-0.027138515 -0.053571429 0.000731172
0.089280742 0.130518519 0 0.006516932 -0.073863636 0
0.049369369 0.091368533
0.012973583 -0.002806914 0.090554254 -0.005785028 0.068825911
0.03505618 0.042449286 0.019525424
0.016047378 0.004749889 -0.053215078 0.022840237 -0.036097561
-0.043010753 -0.067927171 -0.039087948
0.003066015 -0.001925926 0.004454343 -0.014577259 -
0.005819593 0.059104886 0.156743621 -0.026879675

```
0.022633745 0.081037796 0.021615472 0.022293753 0.024850895 0.073340667
      0.103710326 0.041188119
0.024493074 -0.010929827      -0.027654867      -0.009447331      -
0.024248303 -0.006798592      0.019598906 -0.011354738
0.015287403 -0.544187726      0.109656301 -0.042081448      -
0.005785921 -0.004592145      0.005960569 -0.019201229
-0.060513213      -0.009865599      0.034716342 -0.014602608      -
0.014258555 -0.02359882 -0.071914894      -0.128999889
0.017735334 0.107434091 0.181 0.038069891 0.284493284 -0.453684007
      0.189873418 -0.023091167
0.002442122 0.139262199 0.025641026 -0.023833729      -0.060779817
      -0.015047619      0.032949791 0.095454003
0.044591837 0.018370534 0.054054054 -0.038970908      0.142857143 -
0.003984064 0.033513514 -0.042718447
-0.007594937      0.094831054 0.010376843 -0.010738832      0.080736544
      0.059923587 -0.083704804      0.080864198
-0.004435931      0.048834511 -0.061025641      -0.022670025      -
0.004231312 0.079132007 0.114238411 0.049358725
-0.055333333      0.022210243 0.036682616 0.014480409 -0.037991859
      0.022181146 0.017405952 0.031813929
0.049370378 0.114917658 0.034084662 0.058730696 0.004087193 0.06496063
      -0.024109589      0.040907194
-0.052641545      -0.029854227      0.119384615 -0.006940557      -
0.096059113 -0.031495463      -0.054893837      -0.010327733
;
run;
proc iml;
      use stockprice1;
      read all var {x1 x2 x3 x4 x5 x6 x7 x8} into x;
      n=nrow(x);
      mean=t(x)*J(n,1,1/n);
      S=t(x)*(I(n)-J(n,n,1/n))*x/(n-1);
      R=inv(sqrt(diag(S)))*S*inv(sqrt(diag(S)));
      print mean;
      print S;
      print R;

/* we also can use the following code to get correlation matrix */
proc corr data=stockprice1 cov noprob;
      var x1-x8;
run;
/* method1-principal factor method */

            simple
                  method=prin
                  priors=smc
                  nfact=4
            scree
                  rotate=promax
                  round
                  flag=0.40;
      var x1 x2 x3 x4 x5 x6 x7 x8;
      run;
Proc factor corr data=stockprice1 method=principal nfactors=1 scree
rotate=varimax msa;
        var x1 x2 x3 x4 x5 x6 x7 x8;
run;
```

```
proc factor corr data=stockprice1 method=principal nfactors=2
rotate=promax msa;
        var x1 x2 x3 x4 x5 x6 x7 x8;
run;
proc factor corr data=stockprice1 method=principal nfactors=3
rotate=promax msa;
        var x1 x2 x3 x4 x5 x6 x7 x8;
   ;
proc factor corr data=stockprice1 method=principal nfactors=4
rotate=varimax msa;
        var x1 x2 x3 x4 x5 x6 x7 x8;
run;
/* method2-maximum-likehood method */

heywood;
        var x1 x2 x3 x4 x5 x6 x7 x8;
   ;
proc factor corr data=stockprice1 method=ml nfactors=2 rotate=varimax
heywood;
        var x1 x2 x3 x4 x5 x6 x7 x8;

proc factor corr data=stockprice1 method=ml nfactors=3 rotate=promax
heywood score;
        var x1 x2 x3 x4 x5 x6 x7 x8;
run;
proc factor corr data=stockprice1 method=ml nfactors=4 rotate=varimax
heywood;
        var x1 x2 x3 x4 x5 x6 x7 x8;
run;
proc factor data=stockprice method=ml heywood n=3;
      title3 'Maximum-Likelihood Factor Analysis with Three Factors';
   run;
/* score for analysis */
proc factor data=stockprice1 outstat=FactOut
               method=prin nfactors=3 rotate=varimax score;
      var x1 x2 x3 x4 x5 x6 x7 x8;
      title 'FACTOR SCORING EXAMPLE';
      run;

   proc print data=FactOut;
      title2 'Data Set from PROC FACTOR';
   run;

   proc score data=stockprice1 score=FactOut out=FScore;
      var x1 x2 x3 x4 x5 x6 x7 x8;
      run;

   proc print data=FScore;
      title2 'Data Set from PROC SCORE';
   run;
proc factor corr data=stockprice1 outstat=Factout method=ml nfactors=3
rotate=promax heywood score;

      var x1 x2 x3 x4 x5 x6 x7 x8;
      title 'FACTOR SCORING EXAMPLE';
      run;
```

```
   proc print data=FactOut;
      title2 'Data Set from PROC FACTOR';
   run;

   proc score data=stockprice1 score=FactOut out=FScore;
      var x1 x2 x3 x4 x5 x6 x7 x8;
      run;

   proc print data=FScore;
      title2 'Data Set from PROC SCORE';
   run;
proc insight data=stockprice1;
run;
```

2. SAS Log

```
        1164  data stockprice1;
1165          input x1 x2 x3 x4 x5 x6 x7 x8;
1166          cards;

NOTE: SAS went to a new line when INPUT statement reached past the end of a line.
NOTE: The data set WORK.STOCKPRICE1 has 364 observations and 8 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


1532  ;
1533  run;

1534  proc corr data=stockprice1 cov noprob;
1535      var x1-x8;
1536  run;

NOTE: PROCEDURE CORR used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


1537  proc iml;
NOTE: IML Ready
1538      use stockprice1;
1539      read all var {x1 x2 x3 x4 x5 x6 x7 x8} into x;
1540      n=nrow(x);
1541      mean=t(x)*J(n,1,1/n);
1542      S=t(x)*(I(n)-J(n,n,1/n))*x/(n-1);
1543      R=inv(sqrt(diag(S)))*S*inv(sqrt(diag(S)));
1544      print mean;
1545      print S;
1546      print R;
1547  run;
NOTE: Module MAIN is undefined in IML; cannot be RUN.
NOTE: Exiting IML.
NOTE: PROCEDURE IML used (Total process time):
      real time           47.56 seconds
      cpu time            1.39 seconds


1548  Proc factor data=stockprice1
1549              simple
1550              method=prin
1551              priors=smc
1552              nfact=2
1553              scree
1554              rotate=promax
1555              round
1556              flag=0.40;
1557      var x1 x2 x3 x4 x5 x6 x7 x8;
1558      run;

NOTE: 2 factors will be retained by the NFACTOR criterion.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.11 seconds
      cpu time            0.11 seconds


1559  Proc factor data=stockprice1
1560              simple
1561              method=prin
1562              priors=smc
1563              nfact=3
1564              scree
1565              rotate=promax
1566              round
1567              flag=0.40;
1568      var x1 x2 x3 x4 x5 x6 x7 x8;
1569      run;

NOTE: 3 factors will be retained by the NFACTOR criterion.
```

```
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.20 seconds
      cpu time            0.06 seconds


1570  Proc factor data=stockprice1
1571              simple
1572              method=prin
1573              priors=smc
1574              nfact=4
1575              scree
1576              rotate=promax
1577              round
1578              flag=0.40;
1579      var x1 x2 x3 x4 x5 x6 x7 x8;
1580      run;

NOTE: 3 factors will be retained by the MINEIGEN criterion.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.21 seconds
      cpu time            0.07 seconds


1581  proc factor corr data=stockprice1 method=principal nfactors=1 rotate=promax msa;
1582          var x1 x2 x3 x4 x5 x6 x7 x8;
1583  run;

NOTE: 1 factor will be retained by the NFACTOR criterion.
NOTE: Rotation not possible with 1 factor.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.06 seconds
      cpu time            0.04 seconds


1584  proc factor corr data=stockprice1 method=principal nfactors=2 rotate=promax msa;
1585          var x1 x2 x3 x4 x5 x6 x7 x8;
1586  run;

NOTE: 2 factors will be retained by the NFACTOR criterion.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.26 seconds
      cpu time            0.09 seconds


1587  proc factor corr data=stockprice1 method=principal nfactors=3 rotate=promax msa;
1588          var x1 x2 x3 x4 x5 x6 x7 x8;
1589  run;

NOTE: 3 factors will be retained by the NFACTOR criterion.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.20 seconds
      cpu time            0.07 seconds


1590  proc factor corr data=stockprice1 method=principal nfactors=4 rotate=varimax msa;
1591          var x1 x2 x3 x4 x5 x6 x7 x8;
1592  run;

NOTE: 4 factors will be retained by the NFACTOR criterion.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.12 seconds
      cpu time            0.04 seconds


1593  proc factor corr data=stockprice1 method=ml nfactors=1 rotate=varmax heywood;
                                                                     ------
                                                                     1
WARNING 1-322: Assuming the symbol VARIMAX was misspelled as varmax.
1594          var x1 x2 x3 x4 x5 x6 x7 x8;
1595  run;

NOTE: 1 factor will be retained by the NFACTOR criterion.
NOTE: Convergence criterion satisfied.
NOTE: Rotation not possible with 1 factor.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.20 seconds
      cpu time            0.04 seconds


1596  proc factor corr data=stockprice1 method=ml nfactors=2 rotate=varimax heywood;
1597          var x1 x2 x3 x4 x5 x6 x7 x8;
```

```
run;

NOTE: 2 factors will be retained by the NFACTOR criterion.
NOTE: Convergence criterion satisfied.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.18 seconds
      cpu time            0.09 seconds


proc factor corr data=stockprice1 method=ml nfactors=3 rotate=promax heywood score;
        var x1 x2 x3 x4 x5 x6 x7 x8;
run;

NOTE: 3 factors will be retained by the NFACTOR criterion.
NOTE: Convergence criterion satisfied.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.28 seconds
      cpu time            0.04 seconds


proc factor corr data=stockprice1 method=ml nfactors=4 rotate=varimax heywood;
        var x1 x2 x3 x4 x5 x6 x7 x8;
run;

NOTE: 3 factors will be retained by the MINEIGEN criterion.
NOTE: Convergence criterion satisfied.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.20 seconds
      cpu time            0.07 seconds


proc factor data=stockprice1 outstat=FactOut
               method=prin nfactors=3 rotate=varimax score;
      var x1 x2 x3 x4 x5 x6 x7 x8;
      title 'FACTOR SCORING EXAMPLE';
      run;

NOTE: 3 factors will be retained by the NFACTOR criterion.
NOTE: The data set WORK.FACTOUT has 26 observations and 10 variables.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.17 seconds
      cpu time            0.06 seconds


   proc print data=FactOut;
      title2 'Data Set from PROC FACTOR';
   run;

NOTE: There were 26 observations read from the data set WORK.FACTOUT.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.04 seconds
      cpu time            0.01 seconds


   proc score data=stockprice1 score=FactOut out=FScore;
      var x1 x2 x3 x4 x5 x6 x7 x8;
      run;

NOTE: There were 364 observations read from the data set WORK.STOCKPRICE1.
NOTE: There were 26 observations read from the data set WORK.FACTOUT.
NOTE: The data set WORK.FSCORE has 364 observations and 11 variables.
NOTE: PROCEDURE SCORE used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


   proc print data=FScore;
      title2 'Data Set from PROC SCORE';
   run;

NOTE: There were 364 observations read from the data set WORK.FSCORE.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.04 seconds
      cpu time            0.04 seconds
```

```
1624  /* score for maximum-likelihood factor analysis*/
1625  proc factor corr data=stockprice1 outstat=Factout method=ml nfactors=3
rotate=promax heywood
1625!  score;
1626
1627        var x1 x2 x3 x4 x5 x6 x7 x8;
1628        title 'FACTOR SCORING EXAMPLE';
1629        run;

NOTE: 3 factors will be retained by the NFACTOR criterion.
NOTE: Convergence criterion satisfied.
NOTE: The data set WORK.FACTOUT has 48 observations and 10 variables.
NOTE: PROCEDURE FACTOR used (Total process time):
      real time           0.34 seconds
      cpu time            0.03 seconds


1630
1631     proc print data=FactOut;
1632        title2 'Data Set from PROC FACTOR';
1633     run;

NOTE: There were 48 observations read from the data set WORK.FACTOUT.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


1634
1635     proc score data=stockprice1 score=FactOut out=FScore;
1636        var x1 x2 x3 x4 x5 x6 x7 x8;
1637        run;

NOTE: There were 364 observations read from the data set WORK.STOCKPRICE1.
NOTE: There were 48 observations read from the data set WORK.FACTOUT.
NOTE: The data set WORK.FSCORE has 364 observations and 11 variables.
NOTE: PROCEDURE SCORE used (Total process time):
      real time           0.01 seconds
      cpu time            0.00 seconds


1638
1639     proc print data=FScore;
1640        title2 'Data Set from PROC SCORE';
1641     run;

NOTE: There were 364 observations read from the data set WORK.FSCORE.
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.12 seconds
      cpu time            0.03 seconds


1642  proc insight data=stockprice1;
1643  run;
```