A CONTRIBUTION TO MODERN DATA REDUCTION TECHNIQUES AND THEIR APPLICATIONS BY APPLIED MATHEMATICS AND STATISTICAL LEARNING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

HATİCE SAKARYA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
SCIENTIFIC COMPUTING

DECEMBER 2010

Approval of the thesis:

A CONTRIBUTION TO MODERN DATA REDUCTION TECHNIQUES AND THEIR APPLICATIONS BY APPLIED MATHEMATICS AND STATISTICAL LEARNING

Submitted by **HATİCE SAKARYA** in partial fulfilment of the requirements for the degree of **Master of Science in Department of Scientific Computing**, **Middle East Technical University** by,

Prof. Dr. Ersan Akyıldız ____________________
Director**, Institute of Applied Mathematics**

Prof. Dr. Bülent Karasözen ____________________
Head of Department, **Scientific Computing**

Prof. Dr. Gerhard-Wilhelm Weber ____________________
Supervisor, Assistant to the Director**, IAM, METU**

Assist. Prof. Dr. Hakan Öktem ____________________
Co-Supervisor, **Department of Scientific Computing, METU**

Examining Committee Members:

Assist. Prof. Dr. Özlem İlk ____________________
Department of Statistics, METU

Prof. Dr. Gerhard-Wilhelm Weber ____________________
Assistant to the Director, IAM, METU

Assist. Professor Dr. Tolga Can ____________________
Department of Computer Engineering, METU

Date: ____________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : HATİCE SAKARYA

Signature :

ABSTRACT

A CONTRIBUTION TO MODERN DATA REDUCTION TECHNIQUES AND THEIR APPLICATIONS BY APPLIED MATHEMATICS AND STATISTICAL LEARNING

Sakarya, Hatice
M.S., Department of Scientific Computing
Supervisor: Prof. Dr. Gerhard Wilhelm Weber
Co-Supervisor: Assist. Prof. Dr. Hakan Öktem

December 2010, 81 pages

High-dimensional data take place from digital image processing, gene expression micro arrays, neuronal population activities to financial time series. Dimensionality Reduction - extracting low dimensional structure from high dimension - is a key problem in many areas like information processing, machine learning, data mining, information retrieval and pattern recognition, where we find some data reduction techniques. In this thesis we will give a survey about modern data reduction techniques, representing the state-of-the-art of theory, methods and application, by introducing the language of mathematics there. This needs a special care concerning the questions of, e.g., how to understand discrete structures as manifolds, to identify their structure, preparing the dimension reduction, and to face complexity in the algorithmically methods. A special emphasis will be paid to Principal Component Analysis, Locally Linear Embedding and Isomap Algorithms. These algorithms are studied by a research group from Vilnius, Lithuania and Zeev Volkovich, from Software Engineering Department, ORT Braude College of Engineering, Karmiel, and others. The main purpose of this study is to compare the results of the three of the algorithms. While the comparison is beeing made we will focus the results and duration.

Keywords: Data Reduction Techniques, Locally Linear Embedding, Isomap, Principal Component Analysis

ÖZ

MODERN VERİ AZALTMA TEKNIKLERI VE UYGULAMALARINA
UYGULAMALI MATEMATIK VE ISTATISTIK ÖĞRENME İLE BIR KATKI

Sakarya, Hatice
Yüksek Lisans, Bilimsel Hesaplama Bölümü
Tez Yöneticisi : Prof. Dr. Gerhard Wilhelm Weber
Ortak Tez Yöneticisi: Yard. Doç. Dr. Hakan Öktem

Aralık 2010, 81 sayfa

Çok boyutlu veriler dijital görüntü işleme, mikro dizilerde gen açıklaması, nöronsal popülasyon aktiviteleri gibi birçok alanda karşımıza çıkar. Boyut indirgeme- çok boyutlu datayı, daha az boyuta indirgeme- bilgi işleme, veri madenciliği, örüntü tanıma gibi birçok alanda problem oluşturur. Bu tezde Modern Veri Boyut İndirgeme Metodlarını teorik olarak inceleyip, matematiksel olarak uygulamalarını yapacağız. Bunun içinde, parçalı yapıların büklüm olup olmadığı, büklümlerin yapıları, algoritmalarda karşılaşılan zorluklar gibi bazı konularla ilgileneceğiz. Özellikle Vilnius, Litvanya da çalışan bir araştırma grubunun ve Karmiel ORT Braude College of Engineering Üniversitesi Bilgisayar Mühendisliği bölümü öğretim üyelerinden Prof. Dr. Zeev Volkovich'in çalışma konusu olan Yerel Lineer Gömülmesi Analizi, Temel Bileşenler Analizi ve Isomap Algoritmalarına özel bir ilgi gösterip, uygulamalar yapacağız. Bu çalışmanın asıl amacı öncelikli olarak incelediğimiz üç algoritmayı kıyaslamaktır. Algoritmaları kıyaslarken her algoritmanın sonucunu ve hesaplama süresini göz önüne alarak kıyaslama yapacağız.

Anahtar Kelimeler: Veri İndirgeme Metotları, Yerel Lineer gömülmesi, Isomap, Temel Bileşenler Analizi

To my beloved family

ACKNOWLEDGMENTS

I would like to express my deep and sincere gratitude to my supervisor, Professor Gerhard Wilhelm Weber. His wide knowledge and his logical way of thinking have been of great value for me. His understanding, encouraging and personal guidance have provided a good basis for the present thesis.

I am deeply grateful to my advisor, Assist. Prof. Dr. Hakan Öktem, for his detailed and constructive comments, and for his important support throughout this work.

I am also greatly indebted to Prof. Dr. Zeev Volkovich, from Software Engineering Department, ORT Braude College of Engineering, Karmiel, and Prof. Dr. Gintautas Dzemyda from Lithuenia Vilnius Pedagogical University, for the interest and helps of applications, Assist. Prof. Dr. Özlem İlk and Assist. Prof. Dr. Tolga Can who helped my reformulation and style of this work.

Another thanks are for the Secretary of General of Northeast Anatolia Development Agency, Dr. Rıfat Altan. Without his understanding and encouragements, this work would never ended.

During this work I have collaborated with many colleagues, Müşerref, Bilgen, Süleyman, Esma, Melih and Ayşe for whom I have great regard, and I wish to extend my warmest thanks to all those who have helped me with my work.

Last but not least, owe my loving thanks to my dear husband Utku, my dear daughter Elifnaz, my mothers Serpil and Saadet, my father and dear brother. They have lost a lot due to my research. Without their encouragement and understanding it would have been impossible for me to finish this work.

Ankara, Turkey, August 2010
Hatice Sakarya

CONTENTS

LIST OF FIGURES

ABBREVIATIONS

Principal Component Analysis	: PCA
Factor Analysis	: FA
Metric Multi Dimensional Scaling	: MDS
Locally Linear Embedding	: LLE

CHAPTER 1

INTRODUCTION

Generally machine learning problems starts with the pre-processing of multidimensional data sets. We may find some data reduction techniques' applications in different areas like images of any objects, speech signals, spectral histograms, information retrieval and pattern recognition etc.. The goal of data pre-processing is to obtain more informative, descriptive and useful data representations for subsequent operations like classification, visualization, clustering, outlier detection, etc. [24].

Dimensionality reduction is one of the operations of pre-processing. Data which has high dimension include some repetitions and correlations, some hidden relationships. Therefore the main purpose of the pre-processing is to eliminate the redundancies of the data to be processed.

Dimensionality reduction can be done in two different ways;

- Feature selection,
- Feature extraction.

The first one of them, feature selection is selection of a subset of original aspects according to a criterion. It is an important and frequently used dimensionality reduction method for data mining [17, 28]. After the application of feature selection, the number of features in data is reduced and removed irrelevant, redundant, or noisy data are removed. This method is widely used for speeding up a data mining algorithm or improving mining performance.

Feature extraction methods obtain informative projections by applying certain operations to the original features. Feature extraction is used when reduced data representation and first data has the same dimensionality. Then the transformed features may provide better results in further data analysis [24].

Reduction of dimensionality of data is done with two ways which are supervised and unsupervised. We use supervised learning, when data labels are provided, and unsupervised, when no data labels are given. Generally in practice, no prior knowledge about data is available, since it is very expensive and time consuming to assign labels to the data samples. Besides, human experts may assign different labels to the same data sample which may cause the incorrectly discovered relationships between the data samples. Therefore, nowadays the prime interest of the scientists is discovering the hidden structure of the data with unsupervised methods.

1.1 Purpose of the Study and the Method

In this thesis, we will give a survey about modern data reduction techniques, representing the state-of-the-art of theory, methods and application, by introducing the language of mathematics. This needs a special care concerning the questions of, e.g., how to understand discrete structures as "manifolds", to identify their structure, preparing the dimension reduction, and to face complexity in the algorithmical methods.

To compare the results of three algorithms, we have used the results of a questionnaire which is prepared for understanding the customer's personal identification, and comfort, usage and appearance of the driver seat of a car. Generally, customers have difficulty in explaining why they like or dislike a product. The main propose of questionnaire is to demonstrate, model and analyze customer requirements.

That customer satisfaction data set was collected by [9]. While it was collecting, 80 randomly selected customers were chosen from predetermined customer segments. The segments of customers were chosen by the company's sales experts' advice.

There are five chapters in the questionnaire. In the first one demographic questions about the customers are asked. In the second chapter, questions about vehicle usage and the purpose of use are demanded. In the third chapter, customers are observed freely while they are sitting on the driver seat, adjusting and using it. In the fourth section there were questions about comfort, appearance and usage of the seat. Finally in the last section there were anthropometric information were recorded like weight, length, arm length, etc. which could effect the comfort of the seat. In this part, extra features for which the customer accepts to pay more are also demanded. These observations were recorded by video and notes [9].

Our data is 81×70 matrix which is shown in Figure 1. 2 and the data is in Appendix A.

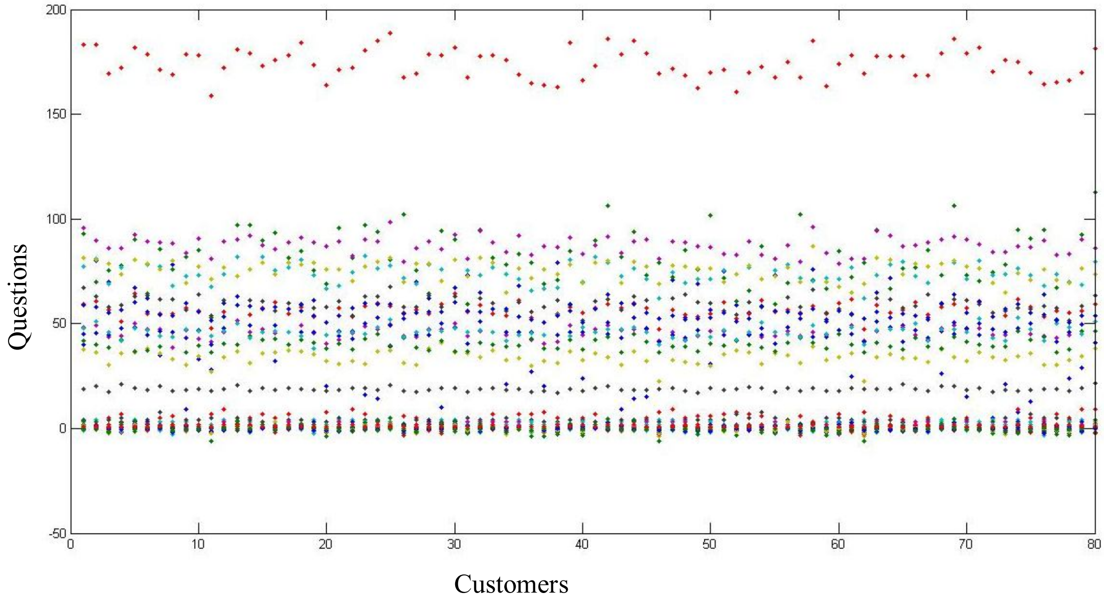


Figure 1.1 The questionnaire data before regularization

To see the better results and make a better comparison we applied the same algorithms on the Swiss Roll[1], Twin peaks[2] and Gaussian[3] data sets and draw the graph of them, measure the duration of calculation.

[1] http://people.cs.uchicago.edu/~dinoj/manifold/swissroll.html
[2] http://www.math.ucla.edu/~wittman/mani/
[3] http://en.pudn.com/downloads74/sourcecode/math/detail265562_en.html

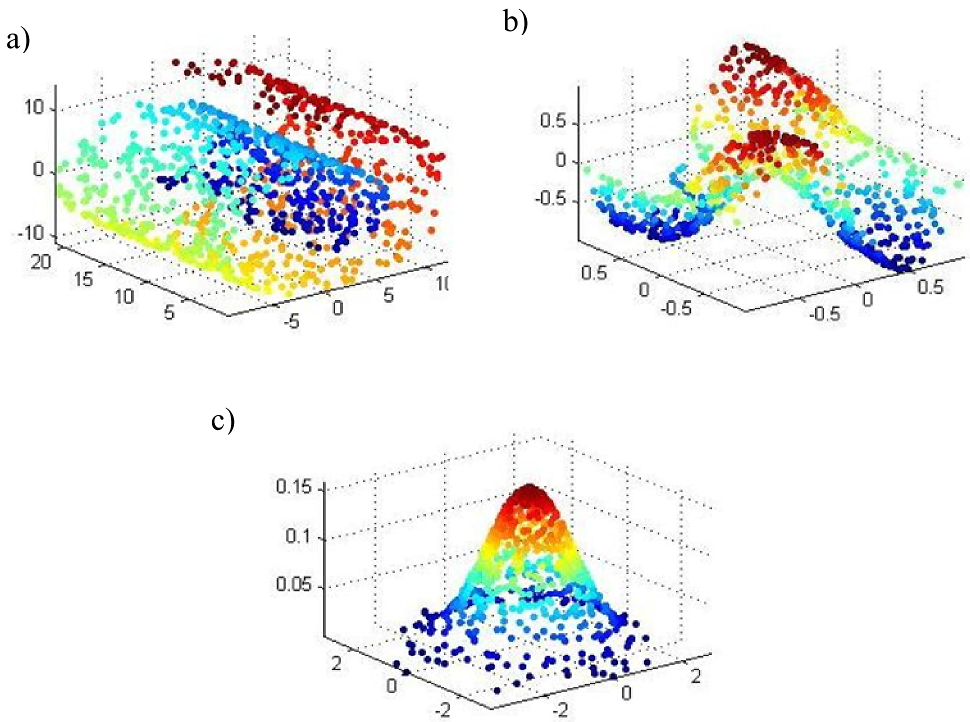


Figure 1.2 a) Swiss Roll data set,
b) Twin Peaks data set,
c) Gaussian

1.2 Scope of the Thesis

This thesis is comprised of six main chapters and an Appendix. Briefly summarizing, the contents are organized as follows:

Chapter 1: this chapter is the introduction part of the thesis. The objectives and outlines of the study is given in this chapter.

Chapter 2: The background information about well known data reduction techniques is given in this chapter.

Chapter 3: Theory and approaches of Principal Component Analysis Algorithm are included in this chapter.

Chapter 4: Theory and approaches of Locally Linear Embedding Algorithm are included in this chapter.

Chapter 5: Theory and approaches of ISOMap Algorithm are included in this chapter.

Chapter 6: Comparison of the method is made in this section.

Chapter 7: Conclusion and further studies are given in this chapter.

CHAPTER 2

A SURVEY ON DIMENSIONALITY REDUCTION TECHNIQUES

Dimensionality Reduction –extraction of low-dimensional data set from a high dimensional data set- is a big problem in most of the following areas like information processing, data mining, machine learning, and information retrieval and pattern recognition. High-dimensional data take place in digital image processing, gene expression micro arrays, neuronal population activities to financial time series.

In this chapter, we will give some information about some data reduction techniques which are some of so-called *dimensionality reduction techniques*. Our aim is not to give detailed information but to describe the simplest form of the algorithms.

2.1 Linear Methods

2.1.1 Principal Component Analysis

Principal Component Analysis, or *PCA*, which is a widely used statistical data reduction technique. It is applied some different areas like face recognition and image compression and pattern recognition in high dimensional data [35]. PCA is a method for pattern recognition with the way as to define their similarities and differences. The mathematical definition the method is an orthogonal linear transformation which transforms the data to a new coordinate system.

While the algorithm is working the firs coordinate is the greatest variance of any projection of the data comes to lie on which is called the *first principal component*, the second greatest variance on the second coordinate which is called *second principal component*, and so on. The approach of PCA is to make the optimum transformation for given data in least square terms. About this method, we will give detailed information and applications about the method in the third chapter.

2.1.2 Factor Analysis

Like PCA, another linear method is *factor analysis* (*FA*). It depends on the data summaries which are second order and includes both *component analysis* and *common factor analysis*. Factor analysis has affected from confusion concerning its very purpose more than other methods [9]. Aim of the method is to find basic patterns in the pattern of relationships between the variable. The first usage and definition of the method is done by psychologist Charles Spearman, whose hypothesis was about the largest differences of tests of mental ability - measures of mathematical skill, artistic skills, ability of logical reasoning, vocabulary, other verbal skills, etc.. What he called *g* is one underlying "factor" of general intelligence [9].

His hypothesis was if *g* might be measured and it could be selected a small group of people with the same score on *g*, in the small group, someone may find no correlations between one mental ability test. This means *g* was the only one factor accepted to all those measures [9].

The variables of some typical examples of the FA algorithm are defined as the scores of group members. The purpose of the algorithm is to acknowledge the relation and then FA algorithm can be applied which will be shown here.

The zero-mean *p*-dimensional random vector $x_{p\times 1}$ which is define as the *k*-factor model as with the covariance matrix Σ if

$$x = \Lambda f + u, \tag{2.1.1}$$

where $\Lambda_{p\times q}$ is a matrix which includes constants, and $f_{k\times 1}$ and $u_{p\times 1}$ are random common factors and specific factors [9]. All of the factors are uncorrelated and the standardisation of the variance of common factors is one:

$$E(f) = 0, \quad Var(f) = I, \tag{2.1.2}$$

$$E(u) = 0, \quad Cov(u_i, u_j) = 0 \;\; for \quad i \neq j \quad , \tag{2.1.3}$$

$$Cov(f, u) = 0. \tag{2.1.4}$$

Furthermore the diagonal covariance matrix of u can be shown as $Cov(u) = \psi = diag(\psi_{11}, ..., \psi_{pp})$.

If the data's covariance matrix could be decomposed like $\Sigma = \Lambda\Lambda^T + \Psi$, then we can show that the k-factor model holds. Since x_i can be written as

$$x_i = \sum_{j=1}^{k} \lambda_{ij} f_j + u_i i_{i=1,2,...,p} \qquad (2.1.5)$$

and its variance could be decomposed as

$$\sigma_{ii} = \sum_{j=1}^{k} \lambda_{ij}^2 + \Psi_{ii}. \qquad (2.1.6)$$

Here, at the beginning, $h^2 = \sum_{j=1}^{k} \lambda_{ij}^2$ is defined as the *communality* and defines the variance of x_i which is common to all variables. Then, Ψ_{ii} is defined as the *specific* or *unique variance* and it is the contribution in the variability of x_i. The term λ_{ij}^2 measures the magnitude of the dependence on the common factor f_j.

The difference between PCA and FA, in PCA the factor model depends on the scale of the variables. On the other hand in FA, the factor model also holds for orthogonal rotations of the factors.

2.1.3 Metric Multidimensional Scaling

The main purpose of *Metric Multidimensional Scaling*, or *MDS*, is to transform the distance matrix into a cross-product matrix and then to find its principal component analysis (PCA) [42].

MDS is based on computing the lower dimensional representation of a higher dimensional data set that the inner products between different input patterns. The outputs $\Psi_i \in \mathbb{R}^m$ are chosen to minimize

$$\varepsilon_{MDS} = \sum_{i,j=1}^{m} \left(x_i \cdot x_j - \Psi_i \cdot \Psi_j \right), \tag{2.1.9}$$

where $x_i \cdot x_j = x_i^T x_j$.

The minimum error solution is applied from the spectral decomposition of the Gram matrix of inner products,

$$G_{ij} = x_i \cdot x_j, \quad (i, j = 1, 2, ..., n). \tag{2.1.10}$$

Then, denote the m eigenvectors of Gram matrix on Equation 2.1.4 by $\left[v_\alpha \right]_{\alpha=1}^{n}$ and their respective eigenvalues by $\left[\lambda_\alpha \right]_{\alpha=1}^{n}$, the outputs of MDS are $\Psi_{i\alpha} = \sqrt{\lambda_\alpha} v_{\alpha i.}$ [42].

The motivation of classical MDS starts by preservation of pair wise distances, since its main purpose of the method is preserving the inner products of it. Let $S = (S_{ij})_{i,j=1,2,3,...n}$ where and $S_{ij} = \left\| x_i - x_j \right\|_2$, be the matrix of squared pair wise distances between input patterns. We generally assume that the inputs are centred on the origin. A Gram matrix consistent with these squared distances can be derived from the transformation $G = -\frac{1}{2}\left(I - uu^T \right) S (I - uu^T)$, where I is the $n \times n$ identity matrix and $u = \frac{1}{\sqrt{n}}(1,1,...,1)^T$ is the uniform vector of unit length [42].

2.2 Graph Based Methods

While we were analyzing high-dimensional data, graph-based methods are also used frequently. In general graph based methods begin with construction of a sparse graph. From these graphs, we define the general attitude of the data. In this sub section we will give just a small information about two famous graph based data reduction techniques, Locally Linear Embedding and ISOMap but concentrate on third and fourth chapters.

2.2.1 Locally Linear Embedding

Locally Linear Embedding, or *LLE*, is also an approved data reduction method. The algorithm starts with considering that the manifold is well sampled. The meaning of this assumption is the number of data point is enough, and every point of data and its neighbours lie on or close to a locally linear patch. Then a data point can be approximated as a weighted linear combination of its neighbours.

LLE algorithm work with some linear transformations such as translation, rotation, and scaling. While computing the lower dimensional version of data set, two constrained least squares optimization problems are solved [21]. We will give detailed information and make some applications about the method in the fourth chapter.

2.2.2 Isomap

Isomap is developed for the basic problem of classical MDS which is the way of calculation of distances in the high-dimensional Euclidean space. Since MDS algorithm fails in high dimensional data like data of planer nature, ISOMap is developed by [40].

We will give detailed information and make some applications about the method in the fifth chapter.

CHAPTER 3

PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis (PCA) is a statistical data reduction technique that alters a original multidimensional data set into essentially smaller set of uncorrelated variables. The new smaller data set involves most of the information of the first multidimensional data set.

The method was designed in 1901 by Karl Pearson and it has been a widely used method in a large area of analysis such as neuroscience to computer graphics. Because of It was a simple and non parametric method, scientists use it frequently. Another advantage of PCA is with a small job, the method afford a roadmap about how to reduce the dimension of a multidimensional data set without losing its underlying structures, hidden dynamics of it.

The aim of the algorithm is reduction of dimensionality of data. It reduces the dimensionality of data with the way as to focus on their similarities and differences. The idea was designed by [29] and developed by [18]. This chapter includes the algorithm and some real life applications.

The mathematical definition of PCA is an orthogonal linear transformation which chances the coordinate system of the data. PCA changes the data set into a new coordinate system (Figure 3.1). The projection onto the first coordinate have the greatest variance between whole of the achievable projections, and the projection onto the second coordinate have the second greatest variances, and so on.

After finding that coordinates which are called principal components, we can simulate the distribution of the original data set which is projected onto a lower dimensional space. This means that, PCA gives us a new viewing angle which can distribute the data almost attainable [36].

The goal of the algorithm is to compute the most meaningful basis to re-express a noisy, garbled data set.

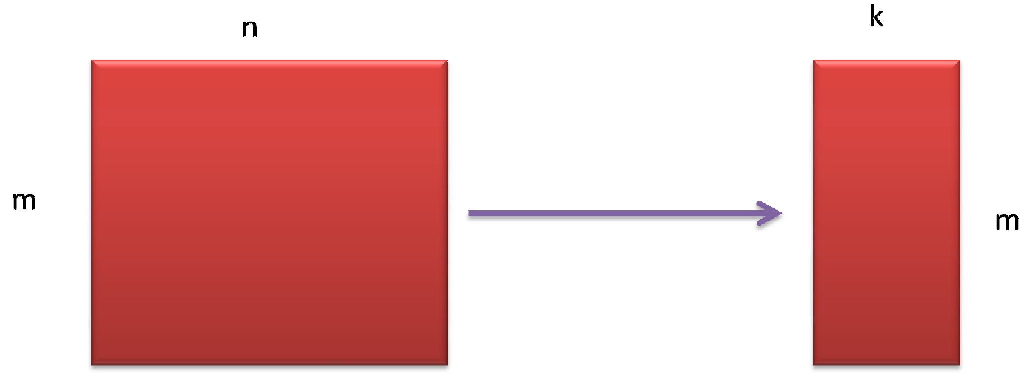


Figure 3. 1 General simulation of PCA

Generally, every data set is a vector in *m* dimensional space, where *m* is the number of measurement types and every data set is a vector that lies in an *m*-dimensional vector space which is spanned by an orthonormal basis. All of the vectors which are for measurement in that space are a linear combination of this set of unit length basis vectors. A trusted and basic choice of a basis *B* is the identity matrix *I*.

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I,$$

where each row is a basis vector b_i with b_i components.

Then PCA tries to find the answer of the question of there were another basis, that were a linear combination of the original basis, which has most of the structures of the higher dimensional data set?

And, PCA makes an assumption: linearity. The meaning of that assumption is if the system was linear, the problem will be simplified by restricting the set of potential bases, formalizing the implicit assumption of continuity in a data set.

Let X and Y be $m \times n$ matrices which are related by a linear transformation P. X is the original multidimensional data set and Y is the lower dimensional new representation of first that data set.

$$PX = Y, \tag{3.2.1}$$

Here x_i and y_i are column vectors and p_i are the rows of P, x_i are the columns of X, y_i are the columns of Y.

Equation 3.2.1 shows a alteration of basis. Hence it can have many interpretations. P is the matrix which transforms X to Y. The meaning of this in geometry is P is a rotation and a stretch that transforms X into Y repeatedly.

Then, $p_1, ..., p_m$, which are the rows of P, are a set of new basis vectors for defining the columns of X.

$$PX = \begin{bmatrix} p_1 \\ \vdots \\ p_m \end{bmatrix} \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}, \tag{3.2.2}$$

$$Y = \begin{bmatrix} p_1 \cdot x_1 & \cdots & p_1 \cdot x_n \\ \vdots & \ddots & \vdots \\ p_m \cdot x_1 & \cdots & p_m \cdot x_n \end{bmatrix}, \tag{3.2.3}$$

We can note the form of each column of Y.

$$y_i = \begin{bmatrix} p_1 \cdot x_i \\ \vdots \\ p_m \cdot x_i \end{bmatrix}. \tag{3.2.4}$$

We recognize that each coefficient of y_i is a dot product of x_i with the corresponding row in P [36]. Here, y_i is a projection on to the basis of $\{p_1, ..., p_m\}$. Therefore, the rows of P are absolutely a new set of basis vectors for representation of columns of X.

Since our assumption was linearity the problem reduces to data to the suitable change of basis. $p_1, ..., p_m$ which are the row vectors in this linear transformation will be the principal components of X.

Then we will try to find what the attitude of the data is. To find the answer we need to talk about *noise*. Noise in any data set must be low or no information about a system can be extracted [36]. An accepted scope is the signal to-noise ratio (SNR), or a ratio of variances σ^2.

$$SNR = \frac{\sigma^2_{signal}}{\sigma^2_{noise}}. \qquad (3.2.5)$$

A high SNR which means the SNR is greater then 1, signifies a higher accuracy of data, then a low SNR signifies noise contaminated data. The determination of SNR is done by calculation of variances. The basic way of computation of repetition quantity among exclusive data points is the computation of something like the variance [36].

Assume that two sets of simultaneous measurements with zero mean.

$$A = \{a_1, a_2, ..., a_n\} \quad , \quad B = \{b_1, b_2, ..., b_n\}.$$

The variance of *A* and *B* are;

$$\sigma^2_A = \langle a_i a_i \rangle_i, \quad \sigma^2_B = \langle b_i b_i \rangle_i,$$

where the expectation is the average over *n* variables. The covariance between *A* and *B* is a straightforward generalization:

$$\text{Covariance of } A \text{ and } B \ \equiv \sigma^2_{AB} = \langle a_i b_i \rangle_{i_{i=1,2,...,n}}.$$

There are two essential facts about the covariance;

- $\sigma^2_{AB} = 0$ if and only if *A* and *B* are completely uncorrelated.
- $\sigma^2_{AB} = \sigma^2_A$ if *A* = *B*.

Then, we may convert the sets of *A* and *B* into the following row vectors.

$$a = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix},$$

$$b = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix},$$

then now we may show the covariance matrix as a computation of dot product matrix,

$$\sigma_{ab}^{2} \equiv \frac{1}{n-1} ab^{T}, \tag{3.2.6}$$

where the beginning term is a constant for normalization.

Here the basic attainable normalization is $\frac{1}{n}$. But, the normalization of $\frac{1}{n}$ affords a partiality belief of variance especially for small *n*. Our aim in this chapter is to show that the correct normalization for an impartial estimator is $\frac{1}{n-1}$ [35].

Finally, if we made a generalisation from two vectors to an arbitrary number, we can rewrite the row vectors $x_1 \equiv a,\ x_2 \equiv b$ and acknowledge row vectors $x_3, ..., x_m$. Then let us define a $m \times n$ matrix *X*,

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}. \tag{3.2.7}$$

Every row of *X* are the complement of all amounts of a appropriate type of x_i. Every column of *X* is complement to a set of measurements from one exact testing.

Now we can define the covariance matrix;

$$S_X := \frac{1}{n-1} XX^{T}. \tag{3.2.8}$$

Here, the dot product between the vector of the i^{th} measurement type with the vector of the j^{th} measurement type is the ij^{th} element of the variance and. In addition to this, the ij^{th} value of XX^{T} is equivalent to replacement of x_i and x_j into Equation 3.2.6.

Furthermore, S_X is a square symmetric $m \times m$ matrix.

- The variance of very measurement types are the terms of S_X which are diagonal.
- The covariance between measurement types are the terms of S_X which are not diagonal.

Since the aim of the method is to reduce repetition, then one should co vary each variable the smallest with other variables. The meaning of this is to remove repetition between separate measurements' covariance to be zero [35].

The construction of the method is based on the eigenvector decomposition. Our data set is X which is an $m \times n$ matrix, where m is the number of measurement types and n is the number of data trials. The point what we can touch et the end is finding the orthonormal matrix P where $Y = PX$, such that $S_Y = \frac{1}{n-1} YY^T$ is diagonalized. Here the rows of P are the principal components of X.

Let's rewrite S_Y in terms of our variable of choice P.

$$\begin{aligned} S_Y &= \frac{1}{n-1} YY^T \\ &= \frac{1}{n-1} (PX)(PX)^T \\ &= \frac{1}{n-1} PXX^T P^T \\ &= \frac{1}{n-1} P(XX^T) P^T \\ S_Y &= \frac{1}{n-1} PAP^T. \end{aligned}$$

Since if A is any matrix, the matrices $A^T A$ and AA^T are both symmetric, A is symmetric.

A symmetric matrix is diagonalized by a matrix of its orthonormal eigenvectors. Therefore for a symmetric matrix A:

$$A = EDE^T, \tag{3.2.9}$$

where D is a diagonal matrix and E eigenvectors matrix of A which is classified as columns.

A has $r \leq m$ orthonormal eigenvectors and rank of the matrix A is r. The rank of A, which is r, and smaller than m when A is degenerate or all data occupy a subspace of dimension $r \leq m$. Preserving the restriction of orthogonality, we can solve this problem by selection of l orthonormal vectors which were added to spill the matrix *E*. Since the variances associated with these directions are zero, the added vectors do not affect the solution [36].

Let us select the matrix *P* to be a matrix where each row p_i is an eigenvector of XX^T. Then $P \equiv E^T$, and substitute into Equation 3.2.9, we find $A = P^T DP$. After this relation $P^{-1} = P^T$ and we can finish evaluating S_Y:

$$\begin{aligned} S_Y &= \frac{1}{n-1} PAP^T \\ &= \frac{1}{n-1} P(P^T DP)P^T \\ &= \frac{1}{n-1} (PP^T)D(PP^T) \\ &= \frac{1}{n-1} P(P^{-1}DP)P^{-1}, \\ S_Y &= \frac{1}{n-1} D. \end{aligned}$$

3.1 Singular Value Decomposition

In this subsection we will show a different algebraic solution for PCA which is related to singular value decomposition (SVD). We will apply the algorithm to our data sets in two of the ways of PCA.

Let X is an $n \times m$ matrix and $X^T X$ is a rank *r*, square, symmetric $n \times n$ matrix. Let's define the variables which will be used here.

- $\{\hat{v}_1, \hat{v}_2, ..., \hat{v}_r\}$ is the set of orthonormal $m \times 1$ eigenvectors with associated eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_r\}$ for the symmetric matrix $X^T X$:

$$(X^T X)\hat{v}_i = \lambda_i \hat{v}_i.$$

- $\sigma_i \equiv \sqrt{\lambda_i}$ are positive real and termed the singular values.

- $\{\hat{u}_1, \hat{u}_2, ..., \hat{u}_r\}$ is the set of orthonormal $n \times 1$ vectors defined by $\hat{u}_i = \frac{1}{\sigma_i} X\hat{v}_i$.

Since $\hat{u}_i \cdot \hat{u}_j = \delta_{ij}$ and $\|X\hat{v}_i\|_2 = \sigma_i$, we can construct the decomposition matrix. The definition of the singular value decomposition is will be:

$$X\hat{v}_i = \sigma_i \hat{u}_i . \qquad (3.2.10)$$

X multiplied by an eigenvector of $X^T X$ is equal to a scalar times another vector and then we started by constructing a new diagonal matrix Σ:

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & 0 & & \\ & & \sigma_r & & & \\ & 0 & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}.$$

Here $\sigma_{\tilde{1}} \geq \sigma_{\tilde{2}} \geq \cdots \geq \sigma_{\tilde{r}}$ are sets of singular values which are ordered by rank. Similarly we construct complementary orthogonal matrices V and U:

$$V = \begin{bmatrix} \hat{v}_{\tilde{1}} & \hat{v}_{\tilde{2}} & \cdots & \hat{v}_{\tilde{m}} \end{bmatrix},$$

$$U = \begin{bmatrix} \hat{u}_{\tilde{1}} & \hat{u}_{\tilde{2}} & \cdots & \hat{u}_{\tilde{n}} \end{bmatrix},$$

where we have added an additional *(m − r)* and *(n − r)* orthonormal vectors to spill the matrices for V and U respectively. Then;

$$XV = U\Sigma ,$$

where each column of V and U perform the singular value decomposition (Equation (3.2.10)) [18]. Since V is orthogonal, we can multiply both sides by $V^{-1} = V^T$ to arrive at the final form of the decomposition.

$$X = U\Sigma V^T . \qquad (3.2.11)$$

This means any matrix X can be converted to an orthogonal matrix, a diagonal matrix and another orthogonal matrix.

Here we will interpret singular value decomposition step by step. Let us start reinterpretation of Equation 3.2.10.

Here, $Xa = kb$, where a and b are column vectors and k is a scalar constant. Then, we can manage Equation 3.2.10,

$$X = U\Sigma V^T,$$
$$U^T X = \Sigma V^T,$$
$$U^T X = Z,$$

where $Z \equiv \Sigma V^T$.

There is a symmetry to SVD such that we can define a similar quantity - the row space [18].

$$XV = \Sigma U,$$
$$(XV)^T = (\Sigma U)^T,$$
$$V^T X^T = U^T \Sigma,$$
$$V^T X^T = Z,$$

where $Z \equiv U^T \Sigma$. Then let us define a new matrix Y which is $n \times m$;

$$Y \equiv \frac{1}{\sqrt{n-1}} X^T,$$

where each column of Y has zero mean.

$$Y^T Y = \left(\frac{1}{\sqrt{n-1}} X^T\right)^T \left(\frac{1}{\sqrt{n-1}} X^T\right),$$
$$= \frac{1}{n-1} X^{TT} X^T,$$
$$= \frac{1}{n-1} X X^T,$$
$$Y^T Y = S_X.$$

We can conclude that finding the principal components amounts to finding an orthonormal basis that spans the column space of X.

3.2 Application of Principal Component Analysis

In this subsection, we will take the steps we needed to perform a Principal Components Analysis on some data sets.

In our first example, we are going to use the results of a questionnaire which is prepared for to understand the customer's personal identification, and comfort, usage and appearance of the driver seat data set which is defined at the end of previous chapter with its details (Appendix A).

While we were performing the algorithm, we subtracted off the mean for each row and covariance matrix of data which is in Appendix B. .

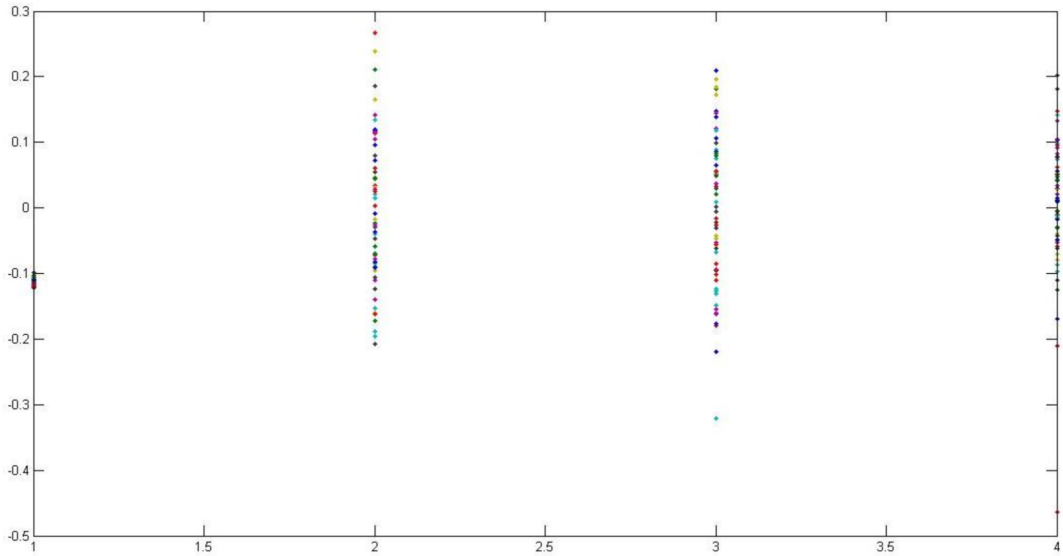


Figure 3.2 The Graph of the data of car seat after PCA (SVD)

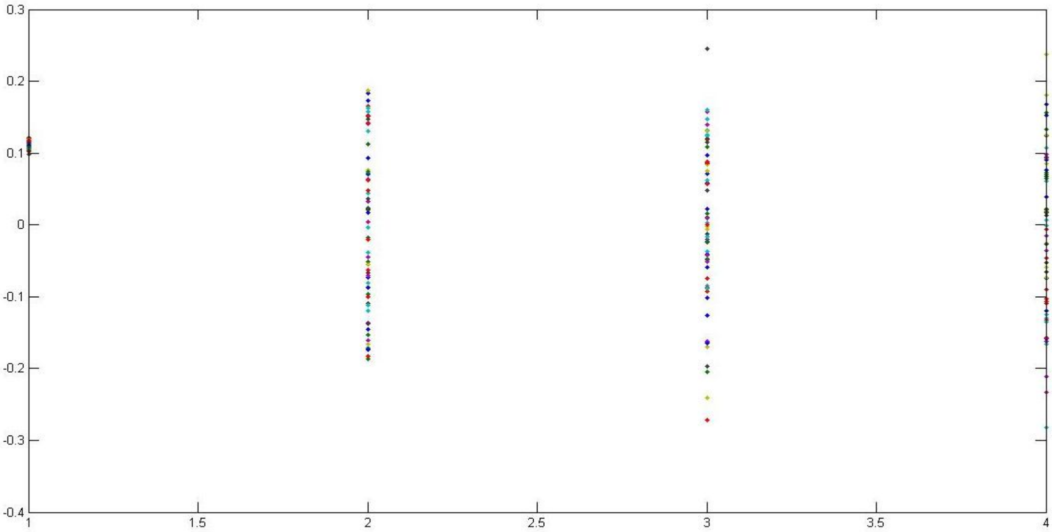


Figure 3. 3 The Graph of the data of car seat after PCA (EVD)

Then we applied PCA in two different ways, singular value decomposition and eigenvalue decomposition. The graphs of the algorithm are in Figure 3.2 and Figure 3.4.3.

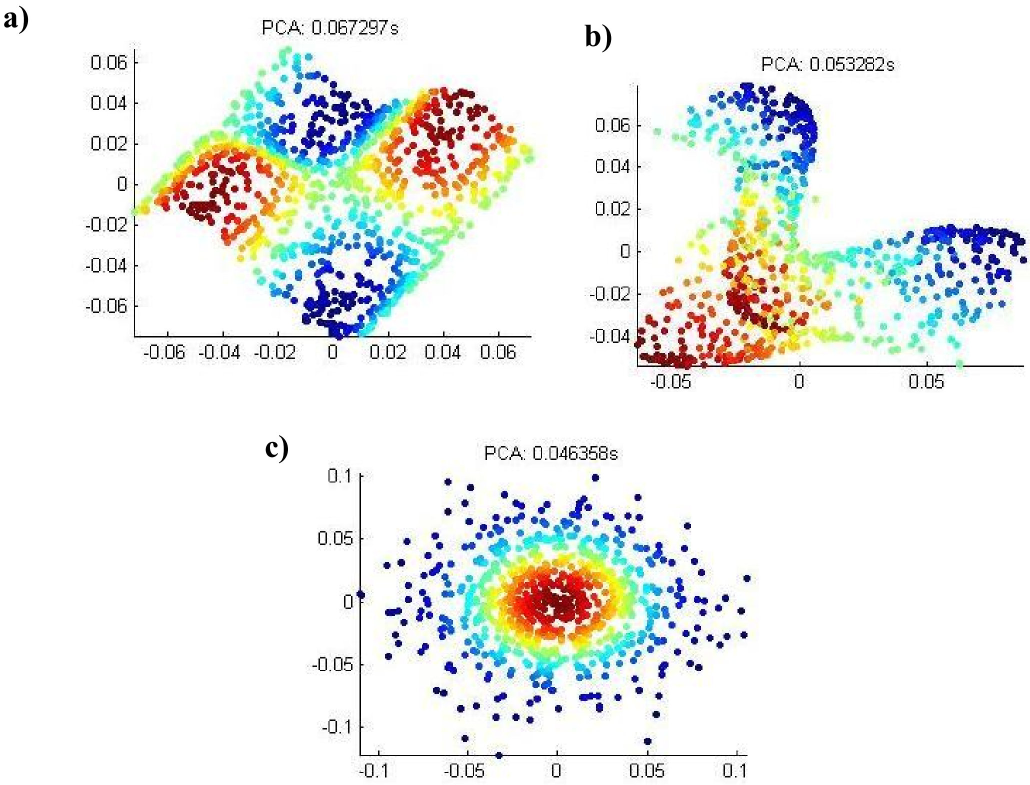


Figure 3.4 PCA Results of
a) Reduction of Swiss Roll data set,
b) Twin Peaks data set,
c) Gaussian data set.

In Figure 3.4 we may see the reduced forms of Swiss Roll, Twin Peaks and of Gaussian data sets with PCA. The duration of calculations are in the bottom of each figure.

Finally, performing PCA is quite simple in practice. At the beginning we organized a data set as an $m \times n$ matrix, here we have had four different data sets, where *m* is the number of measurement types and *n* is the number of trials. Then we subtracted off the mean for every row of our data set, x_i. And at the end we calculated the singular value decomposition or the eigenvectors of the covariance matrix (Appendix A, B,C).

CHAPTER 4

LOCALLY LINEAR EMBEDDING

In today's world in any branch of science, data is often high-dimensional. To understand the general attitude of higher dimensional data is difficult. But underlying structure of a data set can be easily generally. There are many methods for data reduction and some of them are presented in previous chapters. In this chapter we will concentrate on Locally Linear Embedding (*LLE*) algorithm which is a well known graph based*, manifold learning* method for data reduction technique.

This chapter deals with LLE Algorithm. The purpose of the algorithm is to maintain distances when mapping multi dimensional data set to a lower dimensional space. After the operation the data points which are close in the high dimensional space is also close in the lower dimensional one. In other words closer points in input space are closer in output space.

The quality of mapping strongly depends on two control parameters; the first one is the number of nearest neighbours of each data point and the second one is the regularization parameter of a local Gram matrix.

There are some different ways proposed for the selection of the number of the nearest neighbours of each data point [22].

In [22] new two ways are proposed and applied the LLE qualification to multidimensional data visualization. Furthermore also the estimation of topology maintenance, the Spearman's rho [22] is used. Here we will concentrate on the original LLE Algorithm.

There are different applications of Locally Linear Embedding Algorithm such as an image based facial animation system for describing mouth images recognition, analyzing a gait cycle and dynamic features in the gait cycle, hand gesture recognition and tracking, visualizing economic data and to recognize handwritten digits.

LLE is also a dimensionality reduction technique which is widely used in manifold learning.

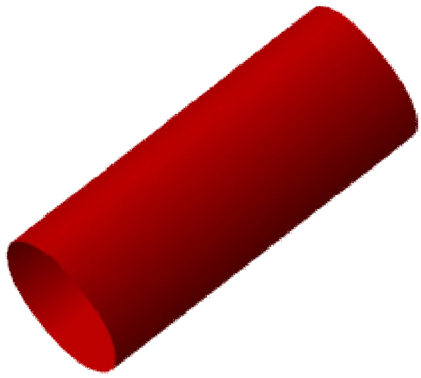


Figure 4.1 A finite cylinder is a manifold with boundary.

Here we should focus on the term of ***manifold*** which is a mathematical space which has a specific dimension that called the dimension of manifold. If we define it a real life experiment, a line and a circle are one-dimensional manifolds, a plane and sphere (the surface of a ball) are two-dimensional manifolds, and so forth.

A ***topological manifold*** is a locally Euclidean Hausdorff space which is defined as a separated space or T_2 space in which distinct points have disjointed neighbourhoods. It is common to place additional requirements on topological manifolds. In addition, topological manifolds may have requirements such as paracompact or second countable

The applications of LLE which includes manifold learning, particularly interesting, since a recent finding in neuroscience, where it is believed to be similar to the human brain's learning process. In [35] there is a new hypothesis and LLE is suggested as one possible algorithm.

The first assumption of LLE is the manifold is the dataset was well sampled. The meaning of this term is there are enough data, and each data point and its neighbours close to or lies on a locally linear patch. Because of the fact that, a data point can be approximated as a weighted linear combination of its neighbours. The main idea of LLE is the combination of such linear transformations;

- translation,
- rotation and
- scaling.

With the algorithm, two unconstrained optimization problem were solved and the configuration of lower data acquired. LLE algorithm has the input, which is the data set at the beginning, consists of *m* *n*-dimensional vectors (points) $X_i = (x_{i1}, ..., x_{in})^T{}_{i=1,2,...,n}$ and $X_i \in \mathbb{R}^n$ which accumulates a matrix of size $n \times m$. This algorithm's output, which is the lower dimensional new data set, consists *m* *d*-dimensional vectors $Y_i = (y_{i1}, ..., y_{id})^T{}_{i=1,2,...,m}$, and $Y_i \in \mathbb{R}^n$ which accumulates a matrix of $d \times m$.

The algorithm has three steps, the first step is selecting a local neighbourhood, then changing each point into a coordinate system based on its neighbours and finally finding new (*q*-dimensional) coordinates which reproduce these local relationships.

At the beginning, identification of *k* neighbours of each data point X_i is done. Here there are different applications. First one identification of a fixed number of *k* nearest neighbours of every data point in terms of *Euclidean distances. O*r the second one is choice of all points within a ball of a fixed radius. Then, we compute the weights w_{ij} which reorganize every data point X_i best from the neighbours $X_{i1}, ..., X_{ik}$ with the minimizing the following error function [33].

$$E(W) = \sum_{j=1}^{m} \left\| X_i - \sum_{j=1}^{k} w_{ij} X_{ij} \right\|^2, \qquad (4.1.1)$$

subject to the constraint $\sum_{j=1}^{k} w_{ij} = 1$. Here, $X_{ij} = (x_{i1}^j, ..., x_{in}^j)$, $(i = 1, ..., m;\ j = 1, ..., k)$ and $\|\cdot\|$ is the Euclidean norm.

In this step, we may recognize that this was a constrained least squares optimisation problem. The solution of the problem is the solution of a linear system.

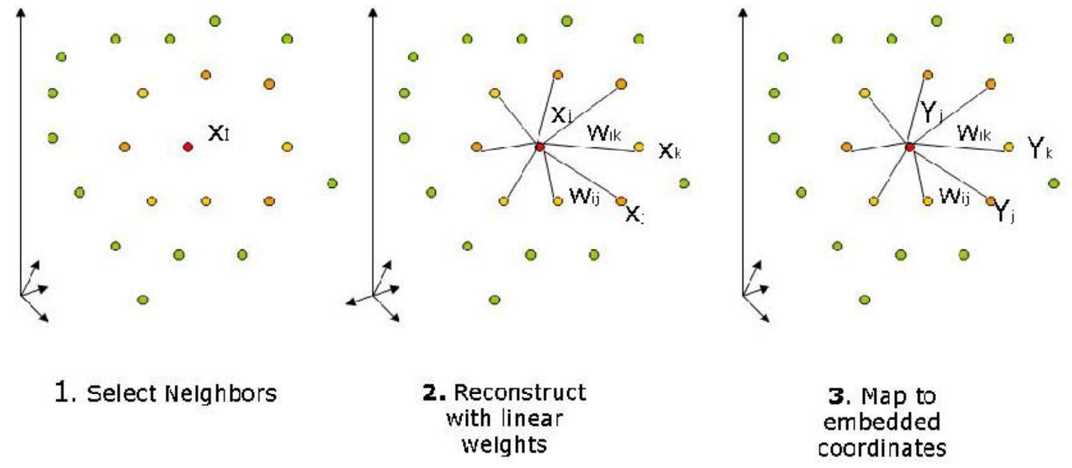


Figure 4.2 Locally Linear Embedding algorithm

To solve the constrained optimization problem which is mentioned before, let us consider a particular data point X_i with k-nearest neighbours X_{ij} and the reconstruction error will be;

$$E^i(W) = \left\| X_i - \sum_{j=1}^{k} w_{ij} X_{ij} \right\|^2 ,$$

$$= \left\| \sum_{j=1}^{k} w_{ij} (X_i - X_{ij}) \right\|^2 . \qquad (4.1.2)$$

Here $C^i = c^i_{jl}$ and $j, l = 1, \ldots, k$ is the local Gram matrix which is

$$c^i_{j,l} = (X_i - X_{ij}) \cdot (X_i - X_{il}) , \qquad (4.1.3)$$

where X_{ij} and X_{il} are the neighbours of X_i .

LLE can be applied by using other metric distances like kernel distance to find the nearest neighbours in the kernel's feature space. Kernel-based learning methods such as support vector machines, the kernel PCA and others [4] are generally used in machine learning and data mining. In [5], the use of distances based on Mercer kernels is given and kernelized form of LLE, KLLE, has been suggested [23].

Assume that $\phi(.)$ is a function of mapping from the n-dimensional space into another higher dimensional, maybe an infinite dimensional space. And assume that X_a and X_b are two points in $\mathbb{R}^n$, then the kernel function will be;

$$\kappa\left(X_a, X_b\right) = \phi\left(X_a\right) \cdot \phi\left(X_b\right), \tag{4.1.4}$$

where $\kappa\left(X_a, X_b\right)$ is the inner product of $\phi\left(X_a\right)$ and $\phi\left(X_b\right)$. In several examples, it may sometimes be useful to compute the coordinates of points in the kernels feature space explicitly. In (4.1.5), various Mercer kernels like the radial basis function kernel, the polynomial kernel, or Gaussian kernel and the linear kernel may be applied to LLE.

Here the kernel Gram matrix's elements are shown;

$$\begin{aligned} c_{j,l}^{i} &= \left(\phi\left(X_i\right) - \phi\left(X_{ij}\right)\right) \cdot \left(\phi\left(X_i\right) - \phi\left(X_{il}\right)\right), \\ &= \kappa(X_i, X_i) - \kappa(X_i, X_{ij}) - \kappa(X_i, X_{il}) + \kappa(X_{ij}, X_{il}), \end{aligned} \tag{4.1.5}$$

where X_{ij} and X_{il} are the neighbours of X_i.

There are different approaches and applications about the computation of distance which are geodesic, Euclidean or Kernel. Here we have computed kernel distance [22]. Our aim is to minimize the error in (4.1.2), then to solve the following equations which are linear; linear system of equations;

$$\sum_{i=1}^{k} c_{jl}^{i} w_{il} = 1, \tag{4.1.6}$$

Finally, we rescale the weights;

$$w_{ij} \leftarrow {w_{ij}} \Big/ {\sum_{l=1}^{k} w_{il}} = 1. \tag{4.1.7}$$

In the situation of the Gram matrix was nearly singular or singular, or when the data points were not in the position which was there were be generally, there can be conditioned by adding a small multiple of the identity matrix [33].

$$c_{jl}^{i} \leftarrow c_{jl}^{i} + \delta_{jl} Tr\left(C^{i}\right) t, \tag{4.1.8}$$

where $Tr\left(C^i\right)$ is the trace of C^i and $\delta_{jl} = 1$ if $j = 1$, otherwise 0, and t is a control parameter which is defined by user between 0 and 1. Let define the regularization parameter;

$$\varepsilon = Tr\left(C^i\right)t \, . \tag{3.1.9}$$

Finally the weights are fixed and embedded coordinates Y_i are found by minimisation of the following function:

$$\phi(Y) = \sum_{i=1}^{m} \left\| Y_i - \sum_{j=1}^{k} w_{ij} Y_{ij} \right\|_2 \, , \tag{3.1.10}$$

subject to $\frac{1}{m}\sum_{i=1}^{m} Y_i Y_i^T = 1$ and $\sum_{j=1}^{m} Y_i = 0$ where I is the $\left(d \times d\right)$ unit matrix and the solution is unique. The most useful method for solving the d-dimensional coordinates is to find the eigenvectors of sparse matrix which is;

$$M = \left(I - W\right)^T \left(I - W\right), \text{ where } W = \left(w_{1j}, w_{2j}, ..., w_{mj}\right), j = 1, 2, ..., k \, . \tag{3.1.11}$$

The eigenvectors are related to the d +1 the smallest eigenvalues of M . The eigenvector at the bottom, having eigenvalue closest to zero, is the unit vector carrying all equal components and it is discarded. The remained d eigenvectors form the d embedding coordinates found by LLE [18].

4.1 Application of Locally Linear Embedding Algorithm

In this subsection, we will take the steps we needed to perform a Locally Linear Embedding Algorithm on the set of data which is the results of questionnaire results of a driver's seat.

While we were applying the LLE algorithm firstly we computed five nearest neighbourhood for each data point. In terms of differential geometry and our data sample the geodesic distance would be the most appropriate metric to identify nearest neighbours. The geodesic distance is the shortest distance on any kind of nonlinear structure. Then we computed reconstructed weights. The main point here is to find the optimal reconstruction for each data point by its neighbours. We computed it to minimise the reconstruction error which is in Equation 4.1.1. And finally with the using of reconstruction weights we computed the embedding into two dimensional spaces. The result of the data is in Figure 4.2.

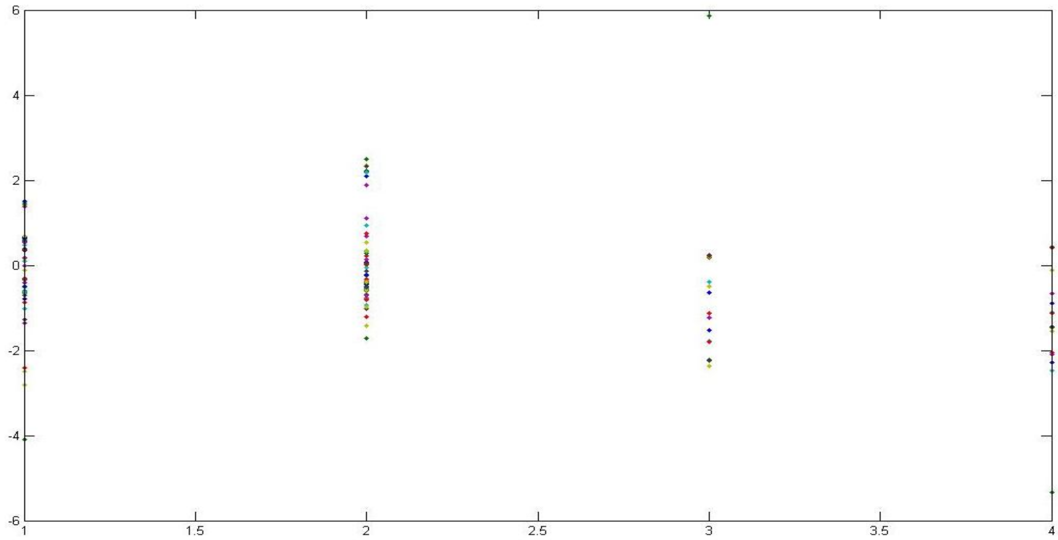


Figure 4.3 The Reduced data after LLE Algorithm

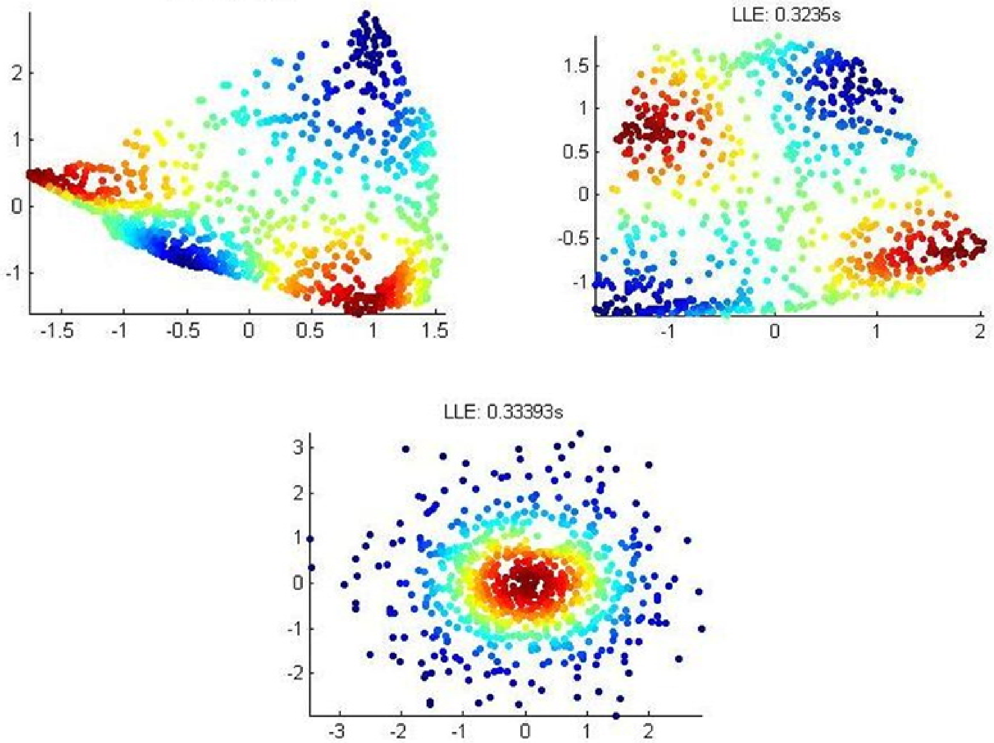


Figure 4.4 LLE Results of

a) Swiss Roll data set,
b) Twin Peaks data set and
c) Gaussian data set.

For a better comparison we applied the LLE algorithm the three well-known data sets which are Swiss Roll, Twin Peaks and of Gaussian data sets. We computed the reduction again with five nearest neighbours and with the geodesic distance. The results of the algorithm are in Figure 4..2. The duration of calculation is again in the bottom of each figure.

Finally performing LLE was simple but taking the results took more time. The result of the algorithm gave different graphs when we changed the number of neighbours while we were computing *k*-nearest neighbours. Another weakness of the algorithm was when we had distant data points the algorithm did not gave healthy results.

CHAPTER 5

ISOMAP ALGORITHM

Because of the difficulties of studying with dense and large data sets like global climate patterns, stellar spectra, or human gene distributions, people who studied science need to reduce the data. It could be possible with reducing the dimensionality of data, or selecting some features of data. ISOMAP is also data reduction method. It is developed by [37] and an extension of the classical multidimensional scaling method for dimension reduction, which finds a linear subspace in which dissimilarity between data points is preserved.

The basis of ISOMAP algorithm is to replace Euclidean distances using an estimation of the geodesic distances on the manifold. The principle aim of this usage to calculate the geodesic distances between points representatives of the shortest paths along the curved surface of the manifold (Figure 5.1) .

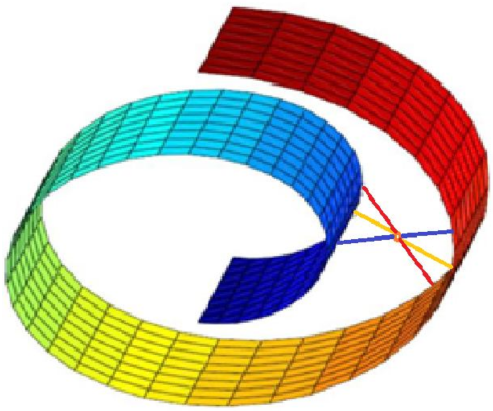


Figure 5.1 Geodesic distance of two data points in Swiss Roll Data Set

Assume that there are M sets of N-dimensional data $X_i = \{x_{i1}, x_{i2}, ..., x_{iM}\}$. The algorithm for dimensionality reduction has three steps. In the first step find the neighbourhood points of each data points $X_{i,j} = 1, 2, ..., m$ by connecting points *i* and *j*. Then if *i* is the *k* nearest neighbours of *j* in Euclidean space let us compute the distance $d_x(i, j)$ among two neighbourhood points.

In the second step, let us estimate the shortest geodesic distance which is $d_G(i, j)$, between two points X_i and X_j. If the two are neighbour points, $d_G(i, j) = d_x(i, j)$; in another way, $d_G(i, j)$ should be computed considering the shortest path iteratively, the shortest distances between all pairs of points from a distance matrix will be $D_G = [d_G(i, j)]$. And final step is the application of classical MDS which was mentioned at second chapter to D_G to obtain an embedding of the data in a *d*-dimensional Euclidean space *Y* carrying the manifolds estimated intrinsic geometry [37].

5.1 Generalization of ISOMAP

Here we will consider that there is a data set which is $X = \{x_1, x_2, ..., x_n\}$ with *n* points in some *D*-dimensional space as the training set. In addition to this we will consider that $Y = \{y_1, y_2, ..., y_n\}$ is the *d*-dimensional embedded coordinates of X (*d* << *D*) from ISOMAP. And let us suppose that $d_g(i, j)$ and $d_G(i, j)$ particularly Euclidean distances and the geodesic distances among points of pairs *(i,j)* respectively. Then Euclidean distance will be;

$$d_g^2(i, j) = \|x_i - x_j\|^2, \qquad (5.1.1)$$

$$= (x_i - x_j)^T (x_i - x_j), \qquad (5.1.2)$$

$$= x_i^T x_i - 2x_i^T x_j + x_j^T x_j, \qquad (5.1.3)$$

$$= \|x_i\|^2 - 2x_i^T x_j + \|x_j\|^2. \qquad (5.1.4)$$

For the embedding y_{new} of a new data point x_{new} and *Y*, the Equation 5.1.4 will be altered into

$$d_E^2\left(y_{new}, i\right) = \|y_i\|^2 - 2y_i^T y_{new} + \|y_{new}\|^2 . \tag{5.1.5}$$

To obtain y_{new} in accordance with *X* and the embedding of it which is Y, the embedding should absorb geodesic distances among x_{new} and every data point x_i, appropriately, x_{new} which means that $d_G(x_{new}, i) = d_E(y_{new}, i)$ for every x_i. Then the Equation (5.1.5) will be like this;

$$d_G^2(x_{new}, i) = \|y_i\|^2 - 2y_i^T y_{new} + \|y_{new}\|^2 \tag{5.1.6}$$

and it can be transformed into

$$y_i^T y_{new} = \left(\|y_{new}\|^2 + \|y_i\|^2 - d_G^2\left(x_{new}, i\right)\right)/2. \tag{5.1.7}$$

Each y_i meets the above formula. Let $f_i = \left(\|y_{new}\|^2 + \|y_i\|^2 - d_G^2\left(x_{new}, i\right)\right)/2$ and $F = (f_1, f_2, ..., f_n)^T$. And we have:

$$Y^T y_{new} = F, \tag{5.1.8}$$

$$YY^T y_{new} = YF = F', \tag{5.1.9}$$

$$Cy_{new} = F', \tag{5.1.10}$$

$$y_{new} = C^{-1} F', \tag{5.1.11}$$

$$C = YY^T. \tag{5.1.12}$$

where

$$F' = YF = (y_1, y_2, ..., y_n)(f_1, f_2, ..., f_n)^T = \sum_{i=1}^{n} y_i f_i \tag{5.1.13}$$

$$= \frac{1}{2}\sum_{i=1}^{n} y_i (\|y_{new}\|^2 + \|y_i\|^2 - d_G^2(x_{new}, i)),$$

$$= \frac{1}{2}\|y_{new}\|^2 \sum_{i=1}^{n} y_i + \frac{1}{2}\sum_{i=1}^{n} y_i (\|y_i\|^2 - d_G^2(x_{new}, i)).$$

Then, assume that $\sum_{i=1}^{n} y_i = 0$, and then let us rewrite the equation;

$$F' = \frac{1}{2}\sum_{i=1}^{n} y_i(\|y_i\|^2 - d_G^2(x_{new}, i)). \qquad (5.1.14)$$

Then let us build up p-dimensional embedding. The matrix which represents the shortest distance can be;

$$\tau(D) = -\frac{1}{2}HSH, \qquad (5.1.15)$$

where S is the matrix of square distances $S = \left[d_G^2(i,j)\right]_{i,j=1,2,\ldots,m}$, H is a centering matrix

$h = -\frac{1}{N}[1]$ Where I is identity matrix, and $[1]$ is ones matrix.

Then compute the eigenvalues λ_p of matrix $\tau(D_G)$ and its corresponding eigenvector, rank the eigenvector v_p by the decreasing order of λ_p and p-dimensional embedding is the eigen matrix involving the first k eigenvectors.

$$\phi_{p*N} = \left[v_1, v_2, \ldots, v_p\right]^T, \qquad (4.1.16)$$

Dimensionality reduction is implemented by mapping every data X_i to Y_i.

$$Y_i = \phi X_i, \ (i = 1, 2, \ldots, m),$$

and the error function will be $E = \|\tau(D_G) - \tau(D_Y)\|_{L^2}$ which is minimal.

5.2 Application of ISOMAP Algorithm

There are some applications of ISOMAP in [37] about hand written digits, hand images or face recognition. In this subsection we will draw the graph of the data, and other data set's reduced forms with ISOMAP.

While we were applying the ISOMAP Algorithm, firstly we constructed the neighborhood graph for five nearest neighbors and then we computed the shortest paths with geodesic distance as in previous section. And finally we constructed the new two dimensional embeddings. We took the results which is showed in Figure 5.2.

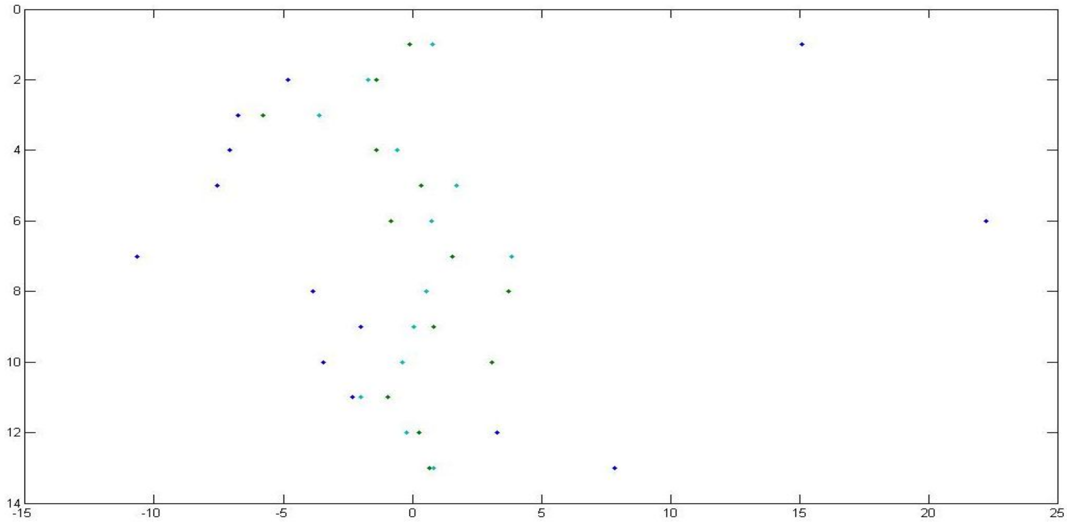
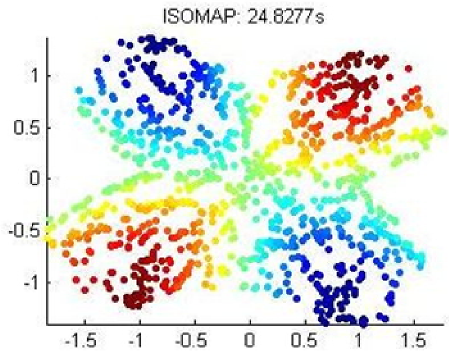


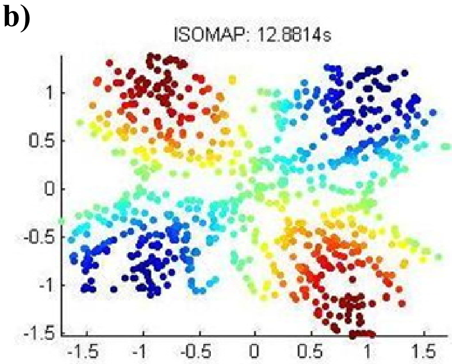
Figure 5.2 The reduced data after ISOMap Algorithm

To see the better results and make a better comparison we applied the algorithm for the three data sets with five nearest neighborhood. The graphs are in Figure 5.3, and the duration of calculation is at the bottom of each figure.

a)



b)



c)

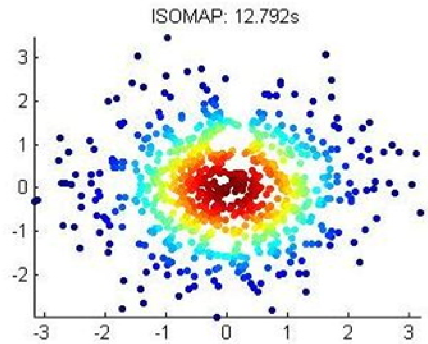


Figure 5.3 ISOMap Results of

a) Swiss Roll data set,
b) Twin Peaks data set and
c) Gaussian data set.

Finally, although ISOMAP enables to examine and manage higher dimensional entries easily based on their intrinsic nonlinear degrees, sometimes application of the algorithm can be expensive and unsuccessful when considered distant data points or data with multiple clusters.

CHAPTER 6

COMPARISON OF THE METHODS

In the previous chapters we have provided a detailed review of three dimensionality reduction techniques which are Principal Component Analysis, Locally Linear Embedding, and ISOMap. In this section, we present the results of these applications. In our experiments, we compared the algorithms in three ways; the method, the running time, and the result. We saw that each method produces different results.

Figure 6.1 shows the results on the data of queastionairre, after the reduction made by the three algorithms. We have used same values for the number of neighbourhoods.

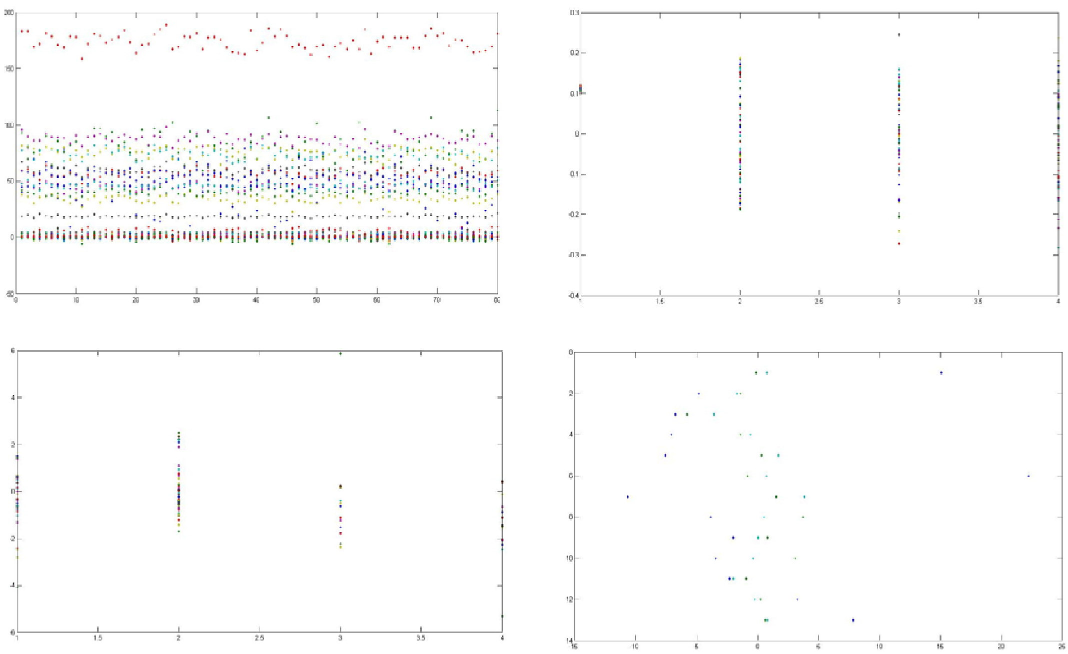


Figure 6.1 From the top left corner

a) Questionnaire data,
b) Reduced form of questionnaire data with PCA,
c) Reduced form of questionnaire data with LLE and
d) Reduced form of questionnaire data with ISOMAP.

PCA reduces the data with orthogonal linear transformation which transforms the data to the new coordinate system by finding the greatest variance. On the other hand LLE Algorithm solves the reduction problem as a graph problem, while it is reducing the data it computes the distance of each data point and prepares the new coordinate system and ISOMap Algorithm makes the reduction converting the Euclidean distance to the geodesic distance. While we apply the algorithms on the questionnaire data, we observed similar results for PCA and LLE, but ISOMap gave different results (Appendix B, C, D). Therefore we may conclude that PCA and LLE are more successful in this data set.

On the other hand, in the view of Swiss Roll data set example, PCA and ISOMap gave similar results but LLE produced different results. When we compare the duration of calculation, we see that the ISOMap algorithm took the longest time (Figure 6.2).

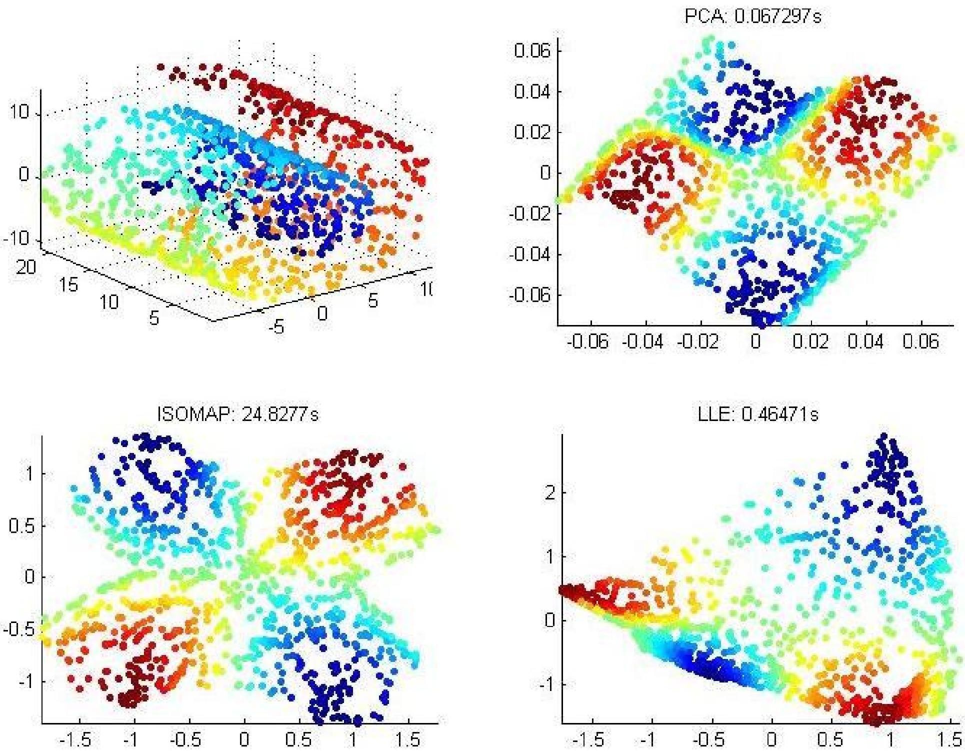


Figure 6.2 From the top left corner

e) Swiss Roll data set,
f) Reduced form of Swiss Roll data set with PCA,
g) Reduced form of Swiss Roll data set with LLE and
h) Reduced form of Swiss Roll data set with ISOMAP.

When we applied the three algorithms on the Twin Peaks data set, ISOMap and LLE algorithms gave similar results but PCA produced different results. The longest time was taken by ISOMap again (Figure 6. 3).

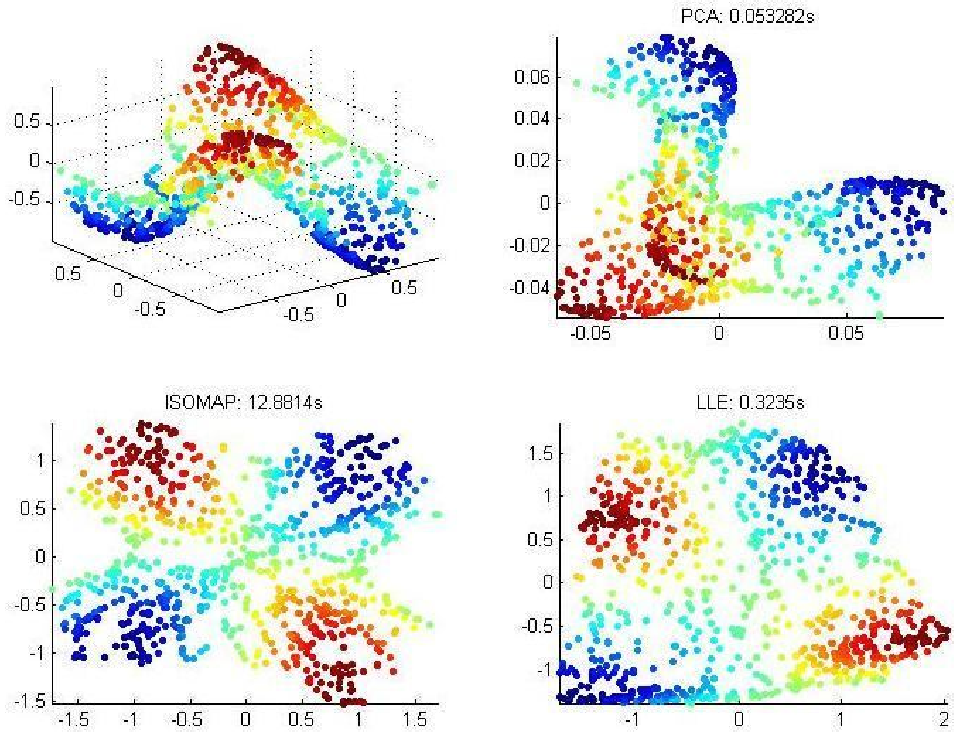


Figure 6.3 From the top left corner

a) Twin Peaks data set,
b) Reduced form of Twin Peaks data set with PCA,
c) Reduced form of Twin Peaks data set with LLE and
d) Reduced form of Twin Peaks data set with ISOMap.

And finally, when we applied the three methods on the Gaussian data set, all of the algorithms gave similar results since it is a symmetrical data set (Figure 6.4).

The biggest difference between the ISOMap and Locally Linear Embedding algorithms is in the fundamental approach to the problem of dimensionality reduction and data representation. ISOMap sees the dimensionality reduction problem as a graph problem. Here, the data is represented as connected graphs and relationship among data is described through the use of geodesic distances.

On the other hand, LLE takes a different approach, looking at this problem from a purely geometrical perspective. However both of them are concerned with preserving the neighbourhoods and their geometric and graph-based relationships.

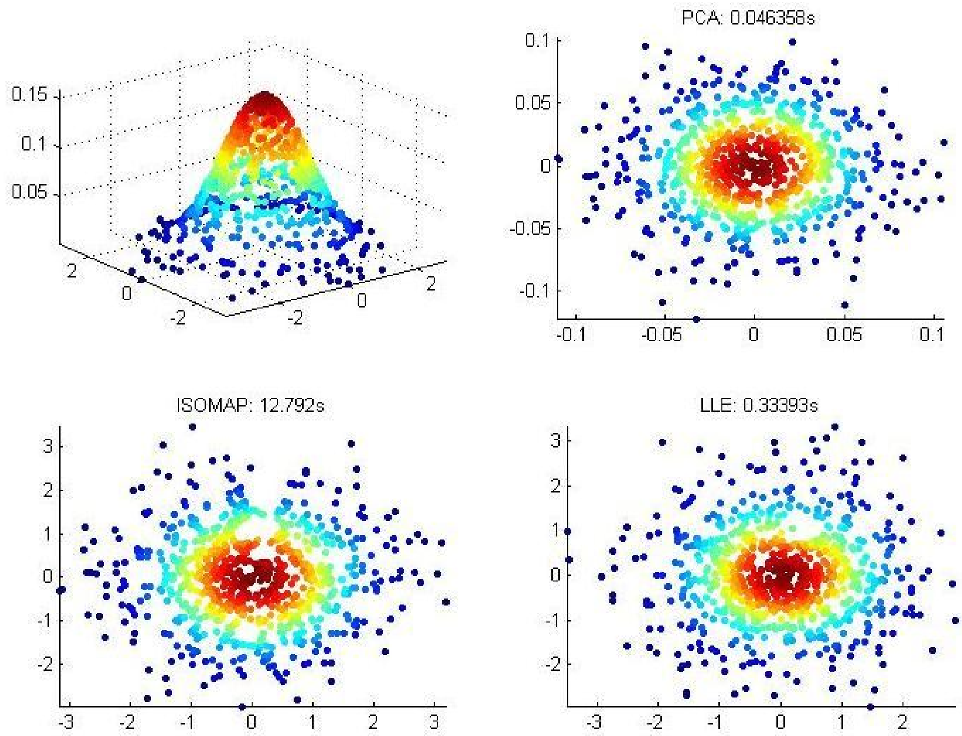


Figure 6.4 From the top left corner

a) Gaussian data set,
b) Reduced form of Gaussian data set with PCA,
c) Reduced form of Gaussian data set with LLE and
d) Reduced form of Gaussian data set with ISOMap.

However, if we look at the problem from the perspective of PCA, we may see that the strength and the weakness of the method is the *non-parametric analysis*. Parameters and coefficients do not need to be adjusted [36]. This means the principal components are not uniquely defined. Therefore the result of the algorithm does not depend on the user.

This strength may also be accepted to be a weakness. If we consider some characteristics of the dynamics in a system, it can be said that these assumptions can be incorporated into a parametric algorithm [36]

ISOMap also contains some problems first one of which is that if the data set has faraway data points, the embeddings tend to separate faraway points leading to distortion of local geometry [41]. It does not project data spread among multiple clusters successfully and it is a well conditioned algorithm which is computationally expensive for large data sets [41].

Furthermore; LLE has also some problems including the conditions that in the case of noisy data, faraway points may have sparse or weakly connected coupling; in the case of under sampled manifold, mapping close points that are faraway in original space and output strongly depends on the selection of k [41]. When we changed the value of k, we observed different results.

CHAPTER 7

CONCLUSION AND FURTHER STUDIES

Generally most of the problems of machine learning such as probability theory, pattern recognition, computational neuroscience, images of an object, spectral histograms, etc. begins with the pre-processing of raw multidimensional data sets. The aim of pre-processing of a multi dimensional data set is to get more descriptive, useful and understandable data set. Some of the operations of pre-processing of a data set are classification, visualization, clustering, outlier detection, etc.

One of the methods for pre-processing of data is dimensionality reduction. By reason the of multi dimensional data has some repetitions, correlations and unseen relationships, the purpose of dimensionality of data is trying to understand the general attitude of data.

In this thesis, we gave survey about modern, well known data reduction techniques and focused on Principal Component Analysis, Locally Linear Embedding and ISOMap Algorithms. Our purpose was to understand which method was better. While we were making the comparison, we compared the methods in two aspects; their result and duration of computations of our applications.

The applications included four different data sets. First one of them is, a real life experiment, which was the results of a questionnaire which is applied to a car company's customers and about a car seat. The data set was 81x70 matrix. The questionnaire includes questions about anthropometrical measures, demographic measures, the usage purpose of their car and comfort, appearance and usage of the car seat.

The other applications were some well known sets, which are Twin Peaks, Swiss Roll and Gaussian data sets. We choose that set because two of the algorithms which we focused on were graph based data reduction techniques. While we were applying the algorithms we have chosen same number of neighbours.

We have showed the results of Principal Component Analysis on the third chapter, Locally Linear Embedding Algorithm on the fourth chapter, ISOMap Algorithm on the fifth chapter and the comparison of the algorithms on the sixth chapter.

After the application we saw that we cannot say any of these methods is better. One method is better in one data, or other is better in any other. If we compare them in the view of time efficiency, we may say PCA is the fastest of them and LLE is the slowest one. But to make comparison with just one way is not true, since all of the algorithm did not gave the same results for all examples.

As a future study, we will compare the methods in some more dense data sets to see their efficiency and to define which kinds of criteria are needed for the success of an algorithm. On the other hand after this thesis we have seen the strength and weakness of the methods. Therefore we will try to develop a new approach to make better reductions and to develop their time efficiency.

REFERENCES

[1] Aggarwal, C. C., Hinneburg A. and Keim D. A., *On the surprising behavior of distance metrics in high dimensional space,* Lecture Notes in Computer Science. - 1973, 2001.

[2] Balasubramanian B., Schwartz E. L., T*he ISOMap Algorithm and Topological Stability,* Science, Vol. 295, No. 5552, pp. 7-7, 2002, doi:10.1126/science.295.5552.7a.

[3] Bell, A., *The independent components of natural scenes are edge filters*, Vision Research , 37(23), 3327-3338,1997.

[4] Belkin, M. and Niyogi P. *Laplacian eigenmaps and spectral techniques for embedding and clustering*, MIT Press, 2002.

[5] Bishop, C., *Neural networks for pattern recognition*, Clarendon, Oxford University Press, 1996.

[6] Cho, M., Park, H. *Nonlinear dimension reduction using ISOMap based on class information,* International Joint Conference on Neural Networks, Atlanta, USA, 2009.

[7] Cox, T., Cox, M., *Multidimensional Scaling,* Chapman & Hall, London, 1994.

[8] Cristianini, J. S., *Support Vector Machines.* Cambridge: Cambridge University Press, 2000.

[9] Çabuk V., A comprehensive analysis of customer desire and expectations for a driver seat: Modeling of voice of the customer (VOC), M.Sc. Thesis, Middle East Technical University , 2008.

[10] Darlington, R. B. *Factor Analysis*, Cornell University Web Page. - November 21 2010. - http://www.psych.cornell.edu/darlington/factor.html.

[11] Decoste, D., *Visualizing Mercer kernel feature spaces via kernelized locally-linear embeddings.* Proc. of the Eighth International Conference on Neural Information Processing, 2001.

[12] Dijkstra, E. W., *A Note on two Problems in Connection with Graphs*, Numerische Mathematik, 1, 29271,1959.

[13] Donoho, D. L., Grimes, C., On Image Manifolds which are Isometric to Euclidean Space. *Journal of Machine Imaging and Vision,* 2005.

[14] Gorban, A., Kegl, B., Wunsch, D., Zinovyev, A. (Eds.), *Principal Manifolds for Data Visualization and Dimension Reduction*, LNCSE 58, Springer, Berlin – Heidelberg – New York, 2007, ISBN 978-3-540-73749-0.

[15] Grimes C. , Donoho, D. L., *Image manifolds which are isometric to Euclidean space.* Department of Statistics, Stanford University and National Science Foundation grant DMS 00-72661, and by DARPA, 2002.

[16] Groenen, B., Modern *Multi Dimensional Scaling*. New York: Springer Verlag, 1997.

[17] Guyon, I., Elisseeff A., An *Introduction to Variable and Feature Selection,* Journal of Machine Learning Research, 2003. - 1157-1182 : Vol. 3. - ISSN 1532-4435.

[18] Hadid, A. K., Unsupervised learning using locally linear embedding: Experiments with face pose analysis. *Internat. Conf. on Pattern Recognition*, (pp. pp. 111-114), 2002.

[19] Hotelling, H., Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417-441,498-520, 1933.

[20] Ihler, A., *Nonlinear Manifold Learning Lecture Notes*, Utah University Web Page, December 01 2010, http://www.cs.utah.edu/~hal/courses/2009F_ML/out/manifold.pdf.

[21] Jollife, I. T., *Principal Component Analysis,* 2nd ed. Springer, NY, 2002, ISBN 978-0-387-95442-4.

[22] Karbauskaitė, R. and Dzemyda, G., T*opology Preservation Measures in the Visualization of Manifold*-Type Multidimensional Data, IOS Press, 2009. - 235-254 : Vol. 32.

[23] Karbauskaitė R., Kurasova O. and Dzemyda G., *Selection of the Number of Neighbours of Each Data Point for the Locally Linear Embedding Algorithm,* Lithuania : 124X Information Technology and Control, 2007.

[24] Karbauskaitė R., Kurasova O. and Dzemyda G., *Topology Preservation*

[25] *Measures in the Visualization of Manifold-Type Multidimensional Data.* Informatica , Volume 20, Number 2, 2009.

[26] Kayo, O., Locally *Linear Embedding Algorithm Extensions and Applications,* Oulu Universty Press, Finland, 2006, ISSN 1796-2226.

[27] Kuo, C., *Data Flow Design for the Backpropagation Algorithm*,International Computer Symposium, Taiwan, 2002.

[28] Lay, D.*, Linear Algebra and* Its *Applications, Addison-Wesley*, New York, 2000.

[29] Mitra, P., *Analysis Of Dynamic Brain Imaging Data*, Biophysical Journal , 76, 691-708. a. 1999.

[30] Motoda, H.and Liu, H., Feature *Selection for Knowledge Discovery and Data Mining*, Kluwer Academic Publishers, Boston, 1998. - ISBN 0-7923-8198.

[31] Mekuz, N. B., *Face recognition with weighted locally linear embedding.,* In The second Canadian conference on computer and robot vision , (pp. 290-296), (2005.

[32] Pearson, K., *On Lines and Planes of Closest Fit to Systems of Points in Space*, Philosophical Magazine 2, http://stat.smmu.edu.cn/history/pearson1901.pdf, 1901.

[33] Saul, L. K. [et al.], Spectral *methods for dimensionality reduction*, Cambridge: MIT Press, 2006.

[34] Saul, L.K. and Roweis S. T., Think *globally, fit locally: unsupervised learning of nonlinear manifolds* , Pennsylvania : University of Pennsylvania Technical Report, 2002. - MS-CIS-02-18.

[35] Saul, V. J., *Exploratory analysis and visualization of speech and music by locally linear embedding.* In Proceedings of the International Conference of Speech, Acoustics, and Signal Processing, (pp. vol. 3, pages 984-987). Montreal, Canada, 2004.

[36] Seung H. and Lee D., *The manifold ways of perception*, Science, 5500, Vol. 290 , 2000.

[37] Shlens, J., *A Tutorial on Principal Component Analysis, Derivation, Discussion and Singular Value Decomposition*, Technical Report, Version 1, 25 March 2003.

[38] Smith L. I., A *tutorial on Principal Components Analysis* ,Otago University, 26 February 2002. - 13 November 2010, http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf.

[39] Tenenbaum, J., Langford, J. and Silva V. *A Global Geometric Framework for Nonlinear Dimensionality Reduction*, Science, 2000. - 290 : Vol. 5500.

[40] Togerson, W., *Theory and Methods of Scaling*, Newyork: Wiley, 1958.

[41] Varini C., *Visualisation of Breast Tumour DCE-MRI Data using LLE,* Proc. Medical Image Understanding and Analysis (MIUA): BMVA , pp. 97-100, 2004.

[42] Veljkovic, D., Nonlinear Dimensionality Reduction, Lecture Slides, http:// cse.spsu.edu/clo/teaching/cs7123/Fall2005/NonlinearDimensionalityReduction.ppt

[43] Will, T., *Introduction to the Singular Value Decomposition.* Davidson College, 2010, www.davidson.edu/academic/math/will/svd/index.html.

[44] Zheng N. and Xue J., Advances *in Pattern Recognition*, Springer, London, 2009.

[45] Zhenyue Z., *Linear low-rank approximation and nonlinear dimensionality reduction*, Science in China Series A-Mathematics , 47, 2004.

[46] Zhou, H. Z., *Role-based collaboration and its kernel mechanisms*. IEEE Trans. Syst., Man, Cybern. C, Appl. Rev. , vol. 36

APPENDICES

Appendix A

The Covariance Matrix of Questionnaire Data

The Covariance Matrix of Questionnaire Data Columns 1 to 10

0,0481	-0,0241	0,0228	0,0000	0,0025	0,0038	0,0380	0,0177
-0,0241	4,8348	-0,0367	0,0696	0,0241	-0,2361	-0,6203	-0,2241
0,0228	-0,0367	1,1873	0,1646	-0,0481	-0,6544	0,3038	-0,0709
0,0000	0,0696	0,1646	0,4177	0,0380	0,0063	-0,0253	-0,0759
0,0025	0,0241	-0,0481	0,0380	0,0481	0,1228	-0,0253	0,0076
0,0038	-0,2361	-0,6544	0,0063	0,1228	1,9690	-0,2785	0,0873
0,0380	-0,6203	0,3038	-0,0253	-0,0253	-0,2785	0,5316	-0,0380
0,0177	-0,2241	-0,0709	-0,0759	0,0076	0,0873	-0,0380	0,2304
0,0127	0,1456	0,0253	0,0253	-0,0127	-0,0570	-0,0127	-0,0506
0,0139	-0,6842	-0,2392	-0,2342	0,0114	0,2196	-0,1013	0,5228
0,0519	-0,6082	-0,0620	-0,1203	0,0114	0,2070	-0,0506	0,5228
0,0203	-0,2633	-0,0684	-0,0633	0,0051	0,0203	-0,1519	0,3899
0,0051	-0,2557	-0,0456	-0,0886	-0,0304	-0,0709	0,0253	0,0025
0,0070	-0,0604	-0,0120	-0,0095	0,0057	0,0117	0,0380	-0,0108
0,0057	0,0478	-0,0259	0,0222	-0,0057	-0,1003	0,0127	-0,0146
0,0019	0,0560	-0,0171	0,0032	0,0108	-0,0060	0,0253	0,0120
-0,0057	-0,0288	-0,0120	-0,0095	-0,0070	-0,0453	0,0380	-0,0234
0,0082	0,1731	0,1032	0,0095	-0,0082	-0,0408	-0,0253	-0,0196
0,0006	-0,0066	0,0070	0,0728	0,0120	0,0149	-0,0127	-0,0297
-0,0120	-0,0763	0,1715	-0,0665	-0,0133	-0,0547	0,0127	0,0209
-0,0329	0,4595	0,0177	0,1139	0,0076	0,0114	-0,1139	-0,0481
0,0051	0,0734	-0,1215	0,0380	0,0329	0,1190	-0,0380	0,0278
-0,0228	-0,0329	-0,0608	-0,0759	-0,0025	-0,1114	-0,0506	0,0329
-0,0165	0,3247	-0,0165	-0,0696	-0,0089	-0,1082	-0,0506	-0,0241
-0,0070	0,1491	0,0247	0,0222	-0,0057	0,0389	-0,0127	-0,0019
-0,0190	0,2753	-0,0063	-0,0063	-0,0063	-0,0791	-0,0506	0,0063
0,0019	0,0370	-0,0044	0,0032	-0,0019	0,0003	0,0000	-0,0006

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
0,0032	0,1503	-0,0032	0,0222	0,0222	0,0174	-0,0127	0,0032
-0,0146	0,2858	-0,0462	0,0348	0,0146	0,1326	-0,1013	0,0133
-0,0203	0,2127	-0,0709	0,0000	-0,0051	-0,0203	-0,0127	0,0025
0,0051	-0,0785	0,0557	-0,0127	-0,0051	-0,1215	0,0253	-0,0101
-0,0285	-0,0554	-0,0222	-0,0475	-0,0095	-0,1503	0,0380	-0,0032
-0,0114	0,1513	0,0266	-0,0570	-0,0013	0,0171	-0,0506	-0,0215
0,0044	0,0864	0,0487	0,0158	-0,0044	-0,0541	-0,0127	-0,0310
-0,0266	0,2222	-0,0646	-0,0316	-0,0114	0,0019	-0,0506	-0,0165
-0,0152	0,0203	-0,1038	0,0000	0,0025	0,1114	0,0000	-0,0203
0,0013	0,2589	0,0266	0,0063	0,0114	-0,0399	-0,0253	-0,0089
-0,0234	0,1636	-0,0044	0,0285	-0,0019	-0,0377	-0,1013	-0,0513
0,0101	0,2608	0,0861	0,0506	-0,0101	-0,1038	0,0000	0,0051
-0,0449	0,3389	0,0500	0,1361	-0,0057	-0,0244	0,0253	-0,3943
-0,0411	0,3877	0,0918	0,1424	-0,0095	-0,0744	0,0633	-0,4209
-0,0386	0,4244	0,1196	0,1424	-0,0120	-0,1668	0,0380	-0,4386
-0,0373	0,3604	0,1335	0,1487	-0,0133	-0,2003	0,0886	-0,4475
-0,0392	0,3487	0,1633	0,1709	-0,0114	-0,1627	0,1013	-0,4342
-0,0095	0,2706	0,0222	0,0222	0,0095	-0,0142	-0,1013	0,0538
-0,0133	-0,0187	0,0563	0,0158	0,0133	-0,0275	0,0253	0,0171
-0,0196	-0,0788	-0,0386	-0,0285	0,0070	-0,0497	0,0127	0,0234
-0,0165	-0,1310	-0,0797	0,0190	0,0038	0,0057	0,0000	-0,0241
-0,0108	-0,2225	-0,0171	-0,0095	-0,0019	0,0573	0,0633	-0,0259
0,0019	-0,0579	0,0082	0,0411	-0,0146	-0,0946	0,0000	-0,0513
0,0032	-0,0649	0,0475	-0,0032	-0,0032	0,0301	0,0253	-0,0222
0,0241	-0,1259	0,1253	0,0063	0,0013	-0,0677	0,0506	0,0342
0,0089	-0,0867	0,0215	0,0190	-0,0089	-0,1652	0,0886	-0,0367
0,0013	-0,1019	0,0519	-0,0316	-0,0266	-0,1728	-0,0253	0,0418
0,0158	0,1693	0,0854	-0,0032	-0,0032	-0,0016	0,0000	-0,0348
-0,0038	-0,1437	-0,0165	0,0063	0,0038	0,0057	0,0253	-0,0114
-0,5244	2,0591	1,7009	-0,7930	-0,1832	-3,8450	0,5924	1,9267
-0,2875	-6,9740	1,9632	-0,0184	-0,2796	-2,5821	2,3608	0,4996
-0,0211	-3,4989	0,5105	-0,3070	-0,2308	-0,8164	1,1506	0,0353
-0,0935	-1,9684	0,7191	-0,2095	-0,1419	-0,8641	0,7443	0,0447
0,0283	-2,5977	0,1644	-0,0617	-0,0992	-0,2763	0,5468	0,0450

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0693	0,1106	0,0085	0,0503	-0,0180	-0,2090	0,0367	0,0180
-0,1594	-0,6159	0,2222	-0,4079	-0,0380	-0,8061	0,1734	0,4147
-0,0348	0,8383	0,0044	-0,2462	0,0892	-0,4883	0,1975	0,2449
0,0207	-2,4964	0,5998	-0,3972	-0,2346	-1,1130	0,8165	0,3108
0,1758	-0,9404	-0,2647	-0,4525	-0,1290	-0,4966	0,5215	0,3642
-0,1805	-2,5253	0,1288	-0,4547	-0,1600	-1,3160	0,3658	0,0632
-0,1027	-1,6965	-0,0159	-0,3852	-0,1644	-1,8030	0,6506	-0,2114
-0,0866	-1,9934	-0,4866	-0,2759	-0,1197	-0,8904	0,3810	0,1744
0,0240	-2,7709	0,6461	-0,2275	-0,1506	-0,9110	1,2051	0,3220

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
0,0127	0,0139	0,0519	0,0203	0,0051	0,0070	0,0057	0,0019
0,1456	-0,6842	-0,6082	-0,2633	-0,2557	-0,0604	0,0478	0,0560
0,0253	-0,2392	-0,0620	-0,0684	-0,0456	-0,0120	-0,0259	-0,0171
0,0253	-0,2342	-0,1203	-0,0633	-0,0886	-0,0095	0,0222	0,0032
-0,0127	0,0114	0,0114	0,0051	-0,0304	0,0057	-0,0057	0,0108
-0,0570	0,2196	0,2070	0,0203	-0,0709	0,0117	-0,1003	-0,0060
-0,0127	-0,1013	-0,0506	-0,1519	0,0253	0,0380	0,0127	0,0253
-0,0506	0,5228	0,5228	0,3899	0,0025	-0,0108	-0,0146	0,0120
0,1899	-0,1582	-0,0823	-0,0759	0,0380	-0,0222	0,0095	0,0158
-0,1582	1,5690	1,1259	0,7873	0,0734	-0,0035	-0,0092	0,0497
-0,0823	1,1259	1,4677	0,8759	-0,0025	-0,0541	-0,0218	0,0117
-0,0759	0,7873	0,8759	0,9013	-0,0278	-0,0241	-0,0266	0,0418
0,0380	0,0734	-0,0025	-0,0278	0,2430	-0,0329	-0,0177	-0,0101
-0,0222	-0,0035	-0,0541	-0,0241	-0,0329	0,1011	-0,0125	0,0391
0,0095	-0,0092	-0,0218	-0,0266	-0,0177	-0,0125	0,1011	-0,0264
0,0158	0,0497	0,0117	0,0418	-0,0101	0,0391	-0,0264	0,1695
0,0285	-0,0858	-0,0098	-0,0620	0,0304	-0,0223	-0,0157	-0,0210
0,0348	-0,0611	-0,0104	-0,0215	0,0278	-0,0321	0,0068	-0,0410
-0,0158	-0,0256	-0,1269	-0,0241	-0,0076	0,0172	0,0081	0,0312
-0,0158	-0,0066	0,0946	0,0646	0,0051	-0,0049	-0,0204	-0,0163
0,0506	-0,0975	-0,0975	-0,0658	-0,0228	-0,0487	0,0108	-0,0133
-0,0127	0,0354	0,0734	0,0481	-0,0354	-0,0013	0,0139	0,0215
0,0127	0,0937	0,0557	0,0810	0,0203	-0,0165	-0,0089	0,0266
0,0443	-0,0297	-0,1057	-0,0203	0,0203	0,0073	-0,0073	0,0250
0,0222	0,0035	0,0161	-0,0013	0,0203	-0,0378	0,0125	-0,0264
0,0190	0,0095	0,0095	0,0253	0,0000	-0,0364	0,0237	-0,0111
0,0158	-0,0199	0,0054	0,0038	0,0025	-0,0084	0,0084	-0,0046
-0,0032	0,0047	-0,0079	0,0190	0,0000	-0,0055	0,0055	0,0008
-0,0285	-0,0117	0,0263	0,0595	-0,0405	-0,0138	-0,0116	-0,0429
0,0000	0,0101	0,0101	0,0101	-0,0354	-0,0013	0,0392	0,0089
0,0000	0,0228	-0,0278	-0,0278	0,0278	-0,0139	0,0392	-0,0165
0,0032	-0,0111	0,0016	0,0063	0,0127	-0,0166	0,0293	0,0150
0,0443	0,0247	0,0120	-0,0228	0,0228	-0,0161	-0,0092	0,0117
0,0285	-0,0718	-0,0465	-0,0418	-0,0025	-0,0027	0,0280	-0,0065

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0190	-0,0184	0,0070	-0,0405	0,0025	-0,0250	0,0123	-0,0149
0,0380	-0,0557	-0,0430	-0,0177	0,0051	0,0038	-0,0165	0,0241
-0,0190	-0,0703	-0,0196	-0,0101	-0,0532	-0,0256	0,0003	-0,0104
0,0285	-0,0579	-0,1339	-0,0468	0,0405	-0,0274	-0,0106	0,0144
0,0380	-0,0051	0,0962	0,0203	-0,0076	-0,0089	0,0089	-0,0266
0,0601	-0,8446	-0,9585	-0,6595	0,0329	0,0508	-0,0002	-0,0138
0,0665	-0,9478	-0,9984	-0,7152	0,0000	0,0340	-0,0087	-0,0103
0,1044	-0,9744	-0,9744	-0,7354	0,0076	0,0144	0,0236	-0,0375
0,0854	-1,0130	-1,0383	-0,7709	0,0177	0,0172	0,0207	-0,0321
0,0696	-0,9677	-0,9551	-0,7241	-0,0101	0,0256	0,0250	-0,0275
0,0348	0,1377	0,1630	0,1203	-0,0253	-0,0277	0,0024	0,0166
-0,0475	0,0383	0,0003	0,0114	-0,0051	-0,0046	-0,0081	-0,0059
-0,0285	0,0415	-0,0092	0,0494	0,0203	0,0065	-0,0065	-0,0074
-0,0316	-0,0677	-0,0297	0,0304	0,0076	0,0263	0,0117	-0,0066
-0,0348	-0,0642	-0,0136	-0,0089	0,0278	0,0138	0,0116	-0,0078
0,0285	-0,1402	-0,1275	-0,0722	0,0025	-0,0305	0,0179	-0,0394
-0,0032	-0,0585	-0,0585	-0,0570	0,0127	0,0071	-0,0071	0,0008
-0,0063	0,1146	0,0892	0,1114	0,0152	-0,0092	-0,0035	0,0389
-0,0316	-0,0614	-0,1247	-0,0456	0,0582	-0,0275	0,0022	0,0155
-0,0063	0,1133	0,0500	0,1038	0,0354	-0,0035	-0,0092	-0,0263
0,0601	-0,1408	-0,0649	-0,0190	-0,0127	-0,0150	-0,0103	0,0419
-0,0063	-0,0171	-0,0044	-0,0456	-0,0051	-0,0180	0,0054	-0,0130
-1,9703	5,5261	4,1350	3,9942	-0,3334	-0,2771	-0,1938	-0,0893
-0,5057	0,9653	1,2083	1,0428	0,7104	0,0912	-0,5064	-0,1724
-0,0437	0,0517	-0,2217	0,1210	0,2916	0,1274	-0,4857	-0,2174
-0,2487	0,2240	-0,0494	0,0635	0,1162	0,1060	-0,2262	-0,1368
-0,0313	-0,1236	0,0548	0,2908	0,4365	0,1649	-0,3118	-0,1785
-0,0396	0,0720	-0,0128	0,0741	0,0266	0,0251	-0,0504	0,0270
-0,4927	1,1949	0,9367	0,7837	0,0170	0,0486	0,0109	-0,0450
-0,2842	0,5573	0,6130	0,6013	0,0456	-0,1027	-0,0353	-0,0672
-0,1339	0,5726	0,5738	0,7072	0,3643	-0,1167	-0,0947	-0,4184
-0,1032	0,5342	0,4444	0,8922	0,2433	0,1497	0,0199	-0,1429
-0,0521	0,5737	0,2762	0,0783	0,6690	0,1674	-0,2433	-0,1303
0,0680	-0,3193	-0,1104	-0,3004	0,7239	-0,0083	-0,1702	-0,1226

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,2608	0,4015	0,5116	0,3129	0,5301	0,1254	-0,1305	-0,1768
-0,1921	0,7222	0,5678	0,5454	0,2159	0,0081	-0,1764	0,0510

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0057	0,0082	0,0006	-0,0120	-0,0329	0,0051	-0,0228	-0,0165
-0,0288	0,1731	-0,0066	-0,0763	0,4595	0,0734	-0,0329	0,3247
-0,0120	0,1032	0,0070	0,1715	0,0177	-0,1215	-0,0608	-0,0165
-0,0095	0,0095	0,0728	-0,0665	0,1139	0,0380	-0,0759	-0,0696
-0,0070	-0,0082	0,0120	-0,0133	0,0076	0,0329	-0,0025	-0,0089
-0,0453	-0,0408	0,0149	-0,0547	0,0114	0,1190	-0,1114	-0,1082
0,0380	-0,0253	-0,0127	0,0127	-0,1139	-0,0380	-0,0506	-0,0506
-0,0234	-0,0196	-0,0297	0,0209	-0,0481	0,0278	0,0329	-0,0241
0,0285	0,0348	-0,0158	-0,0158	0,0506	-0,0127	0,0127	0,0443
-0,0858	-0,0611	-0,0256	-0,0066	-0,0975	0,0354	0,0937	-0,0297
-0,0098	-0,0104	-0,1269	0,0946	-0,0975	0,0734	0,0557	-0,1057
-0,0620	-0,0215	-0,0241	0,0646	-0,0658	0,0481	0,0810	-0,0203
0,0304	0,0278	-0,0076	0,0051	-0,0228	-0,0354	0,0203	0,0203
-0,0223	-0,0321	0,0172	-0,0049	-0,0487	-0,0013	-0,0165	0,0073
-0,0157	0,0068	0,0081	-0,0204	0,0108	0,0139	-0,0089	-0,0073
-0,0210	-0,0410	0,0312	-0,0163	-0,0133	0,0215	0,0266	0,0250
0,1201	0,0153	-0,0239	0,0172	0,0019	-0,0013	-0,0038	0,0009
0,0153	0,1378	-0,0305	0,0296	0,0437	-0,0165	-0,0114	0,0092
-0,0239	-0,0305	0,1960	-0,1109	0,0589	-0,0013	-0,0544	-0,0339
0,0172	0,0296	-0,1109	0,2530	-0,0551	-0,0266	0,0595	0,0104
0,0019	0,0437	0,0589	-0,0551	0,2304	-0,0101	-0,0177	0,0139
-0,0013	-0,0165	-0,0013	-0,0266	-0,0101	0,0911	-0,0051	-0,0177
-0,0038	-0,0114	-0,0544	0,0595	-0,0177	-0,0051	0,2127	0,0354
0,0009	0,0092	-0,0339	0,0104	0,0139	-0,0177	0,0354	0,1462
-0,0030	0,0448	-0,0299	0,0176	0,0234	-0,0114	-0,0089	0,0054
-0,0174	0,0301	-0,0332	0,0237	0,0570	-0,0127	0,0253	0,0285
-0,0052	0,0065	-0,0100	-0,0068	-0,0006	-0,0038	0,0013	0,0187
-0,0087	0,0024	0,0087	-0,0071	0,0158	0,0190	-0,0190	0,0269
-0,0296	0,0283	-0,0059	0,0163	0,0766	0,0165	-0,0013	0,0003
-0,0139	-0,0038	-0,0266	-0,0139	0,0278	0,0278	0,0076	0,0203
-0,0139	0,0089	-0,0013	0,0114	-0,0101	-0,0101	0,0329	-0,0051
0,0119	-0,0182	-0,0245	0,0419	-0,0032	0,0063	0,0316	0,0047
0,0155	0,0149	-0,0383	-0,0066	0,0165	-0,0025	0,0304	0,0589

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
0,0005	-0,0017	0,0147	-0,0074	0,0196	-0,0089	-0,0013	-0,0028
0,0446	0,0136	-0,0472	0,0478	0,0468	-0,0101	0,0203	0,0234
0,0038	-0,0646	-0,0089	-0,0089	0,0177	0,0177	0,0025	0,0278
-0,0066	-0,0263	-0,0161	0,0155	0,0165	0,0228	0,0177	0,0019
0,0138	0,0634	0,0090	0,0122	0,0627	-0,0165	0,0519	0,0440
0,0291	0,0494	-0,0089	0,0038	0,0557	0,0051	-0,0228	-0,0101
0,0223	0,0068	0,0714	-0,0204	0,0614	-0,0494	-0,0468	0,0434
0,0372	0,0324	0,0388	-0,0340	0,0854	-0,0443	-0,0570	0,0554
0,0555	0,0495	0,0255	-0,0093	0,0930	-0,0494	-0,0595	0,0592
0,0521	0,0454	0,0441	-0,0350	0,0842	-0,0646	-0,0544	0,0421
0,0573	0,0389	0,0541	-0,0155	0,0848	-0,0481	-0,0684	0,0234
0,0071	0,0245	-0,0198	0,0403	0,0411	0,0063	0,0190	0,0332
-0,0014	-0,0201	0,0191	0,0350	-0,0082	0,0139	0,0418	-0,0421
0,0033	0,0005	-0,0109	-0,0267	-0,0272	0,0139	0,0418	0,0180
0,0073	-0,0098	0,0231	0,0168	-0,0241	0,0203	0,0101	-0,0310
0,0296	0,0097	-0,0068	0,0343	-0,0259	0,0089	0,0139	-0,0130
0,0486	0,0160	-0,0005	-0,0100	0,0247	-0,0165	-0,0114	-0,0003
-0,0087	-0,0103	0,0087	0,0055	0,0032	-0,0063	-0,0190	-0,0111
-0,0282	0,0104	-0,0377	0,0826	-0,0544	0,0152	0,0456	0,0044
0,0168	0,0187	0,0263	-0,0180	-0,0241	-0,0177	0,0101	-0,0247
0,0282	0,0275	-0,0256	0,0313	-0,0089	-0,0278	0,0430	-0,0044
-0,0055	0,0372	-0,0071	0,0150	0,0158	-0,0190	0,0316	0,0206
0,0389	0,0092	-0,0085	-0,0275	-0,0114	-0,0051	0,0101	-0,0057
0,7083	-0,9845	-0,1509	0,1612	-0,2999	-0,3790	0,1529	0,2703
0,3159	-0,2733	-0,3997	0,9124	-1,0244	-0,5390	0,2185	-0,4015
0,2066	-0,2936	-0,2387	0,2088	-0,5166	-0,3197	-0,1618	0,0164
0,0952	-0,1694	-0,0823	0,2791	-0,4059	-0,1990	0,0701	-0,0764
0,1171	-0,0666	-0,1321	-0,0115	-0,3018	-0,1920	-0,1153	-0,0041
0,0045	-0,0494	-0,0172	0,0711	-0,0174	-0,0361	0,0323	0,0527
0,1419	-0,2157	0,0639	0,0927	-0,1599	-0,0368	0,0224	0,0662
0,2359	-0,0498	-0,0612	0,0217	0,1209	-0,0076	-0,0848	0,1402
0,1507	-0,0331	-0,2479	0,1132	-0,3854	-0,3059	-0,1659	-0,1281
-0,1737	-0,0357	-0,0493	-0,0069	-0,2890	-0,0706	-0,0220	-0,0413
0,1431	0,0966	-0,6264	0,6219	-0,6431	-0,3226	0,1688	0,2671

0,3584	0,1389	-0,6251	0,4882	-0,4810	-0,2908	-0,0243	0,2962
0,2495	-0,0439	-0,4265	0,2368	-0,5041	-0,0889	-0,0565	0,1473
0,0856	-0,2814	-0,0041	0,1608	-0,4780	-0,2733	-0,0186	-0,2795

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0070	-0,0190	0,0019	0,0032	-0,0146	-0,0203	0,0051	-0,0285
0,1491	0,2753	0,0370	0,1503	0,2858	0,2127	-0,0785	-0,0554
0,0247	-0,0063	-0,0044	-0,0032	-0,0462	-0,0709	0,0557	-0,0222
0,0222	-0,0063	0,0032	0,0222	0,0348	0,0000	-0,0127	-0,0475
-0,0057	-0,0063	-0,0019	0,0222	0,0146	-0,0051	-0,0051	-0,0095
0,0389	-0,0791	0,0003	0,0174	0,1326	-0,0203	-0,1215	-0,1503
-0,0127	-0,0506	0,0000	-0,0127	-0,1013	-0,0127	0,0253	0,0380
-0,0019	0,0063	-0,0006	0,0032	0,0133	0,0025	-0,0101	-0,0032
0,0222	0,0190	0,0158	-0,0032	-0,0285	0,0000	0,0000	0,0032
0,0035	0,0095	-0,0199	0,0047	-0,0117	0,0101	0,0228	-0,0111
0,0161	0,0095	0,0054	-0,0079	0,0263	0,0101	-0,0278	0,0016
-0,0013	0,0253	0,0038	0,0190	0,0595	0,0101	-0,0278	0,0063
0,0203	0,0000	0,0025	0,0000	-0,0405	-0,0354	0,0278	0,0127
-0,0378	-0,0364	-0,0084	-0,0055	-0,0138	-0,0013	-0,0139	-0,0166
0,0125	0,0237	0,0084	0,0055	-0,0116	0,0392	0,0392	0,0293
-0,0264	-0,0111	-0,0046	0,0008	-0,0429	0,0089	-0,0165	0,0150
-0,0030	-0,0174	-0,0052	-0,0087	-0,0296	-0,0139	-0,0139	0,0119
0,0448	0,0301	0,0065	0,0024	0,0283	-0,0038	0,0089	-0,0182
-0,0299	-0,0332	-0,0100	0,0087	-0,0059	-0,0266	-0,0013	-0,0245
0,0176	0,0237	-0,0068	-0,0071	0,0163	-0,0139	0,0114	0,0419
0,0234	0,0570	-0,0006	0,0158	0,0766	0,0278	-0,0101	-0,0032
-0,0114	-0,0127	-0,0038	0,0190	0,0165	0,0278	-0,0101	0,0063
-0,0089	0,0253	0,0013	-0,0190	-0,0013	0,0076	0,0329	0,0316
0,0054	0,0285	0,0187	0,0269	0,0003	0,0203	-0,0051	0,0047
0,1011	0,0617	0,0337	0,0055	0,0264	0,0139	0,0013	0,0040
0,0617	0,1108	0,0332	0,0047	0,0617	0,0506	0,0127	0,0269
0,0337	0,0332	0,0366	0,0103	0,0172	0,0089	0,0089	0,0055
0,0055	0,0047	0,0103	0,0593	0,0245	0,0063	0,0063	-0,0119
0,0264	0,0617	0,0172	0,0245	0,1695	0,0418	-0,0089	-0,0150
0,0139	0,0506	0,0089	0,0063	0,0418	0,0911	0,0025	0,0316
0,0013	0,0127	0,0089	0,0063	-0,0089	0,0025	0,0911	0,0063
0,0040	0,0269	0,0055	-0,0119	-0,0150	0,0316	0,0063	0,1543
0,0288	0,0222	0,0180	0,0301	-0,0117	0,0101	0,0228	-0,0237

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0100	0,0016	-0,0033	-0,0055	-0,0188	-0,0089	0,0038	0,0087
0,0250	0,0348	-0,0085	-0,0142	0,0149	0,0278	0,0025	0,0332
-0,0038	-0,0127	-0,0013	0,0190	0,0013	0,0304	-0,0456	-0,0063
-0,0123	0,0032	-0,0041	0,0142	0,0358	0,0101	-0,0278	0,0174
0,0400	0,0332	0,0049	-0,0087	0,0109	-0,0165	-0,0418	-0,0008
0,0215	0,0000	-0,0013	0,0190	0,0013	0,0051	-0,0203	0,0063
-0,0002	-0,0269	-0,0043	-0,0071	-0,0242	-0,0241	0,0013	0,0166
0,0166	-0,0111	0,0055	0,0008	-0,0024	0,0063	-0,0063	0,0150
-0,0017	-0,0174	0,0036	-0,0024	-0,0131	-0,0114	0,0266	0,0182
0,0081	-0,0079	0,0027	-0,0040	-0,0312	-0,0139	0,0241	0,0134
-0,0003	-0,0158	-0,0085	-0,0142	-0,0231	-0,0101	0,0152	0,0079
0,0530	0,0617	0,0134	0,0055	0,0214	0,0063	-0,0190	-0,0087
-0,0334	-0,0047	-0,0153	0,0040	0,0312	-0,0114	0,0139	0,0372
-0,0191	-0,0016	0,0147	0,0245	0,0328	0,0013	0,0139	0,0356
-0,0136	-0,0095	-0,0003	0,0079	0,0320	0,0076	-0,0177	0,0237
-0,0138	-0,0111	0,0081	0,0134	0,0204	0,0089	0,0215	0,0024
0,0179	0,0142	0,0271	-0,0055	0,0141	-0,0165	0,0342	-0,0166
-0,0071	-0,0079	-0,0024	-0,0040	0,0245	-0,0063	-0,0063	0,0008
0,0092	0,0032	-0,0054	-0,0047	-0,0009	-0,0101	0,0152	0,0237
-0,0104	-0,0032	-0,0035	-0,0016	-0,0028	-0,0430	0,0203	0,0079
-0,0345	0,0095	-0,0199	-0,0332	0,0136	-0,0152	0,0354	0,0016
0,0150	0,0111	0,0134	0,0182	-0,0166	-0,0063	0,0316	-0,0214
0,0054	-0,0095	0,0060	-0,0111	-0,0123	-0,0177	0,0076	-0,0206
-0,1102	0,8560	-0,3127	-0,3397	0,6931	0,6147	-0,5727	0,6074
0,0278	-0,0750	-0,1528	-0,3432	-0,6340	-0,2567	-0,1681	0,5653
-0,1616	-0,3832	-0,1522	-0,2524	-0,4269	-0,4008	-0,3514	-0,0103
-0,0794	-0,1509	-0,1235	-0,1552	-0,2366	-0,1446	-0,0787	0,0179
-0,0421	-0,2378	-0,0238	-0,0733	-0,2443	-0,2490	-0,1642	-0,1276
0,0040	-0,0090	-0,0480	-0,0248	-0,0105	-0,0019	-0,0728	0,0371
-0,2081	0,0777	-0,1137	-0,0776	0,0488	0,1151	0,0011	0,1811
0,0255	0,1737	0,0043	0,1210	0,1090	0,0544	0,0089	0,1049
0,0269	0,0584	-0,0636	-0,1537	-0,2183	-0,0287	-0,0528	-0,0282
-0,1016	0,0661	-0,0740	0,0160	-0,0153	0,1167	-0,0744	-0,0495

0,1060	0,0157	-0,0558	-0,0833	-0,2976	-0,0226	-0,0846	0,3451
0,0058	-0,1394	-0,0432	0,1664	-0,4901	-0,0845	-0,1351	0,3611
-0,1862	-0,1228	-0,0490	-0,0108	-0,2586	0,0808	-0,1167	0,3437
-0,0207	-0,0366	-0,0765	-0,1904	-0,3092	-0,1606	-0,1239	0,1819

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0114	0,0044	-0,0266	-0,0152	0,0013	-0,0234	0,0101	-0,0449
0,1513	0,0864	0,2222	0,0203	0,2589	0,1636	0,2608	0,3389
0,0266	0,0487	-0,0646	-0,1038	0,0266	-0,0044	0,0861	0,0500
-0,0570	0,0158	-0,0316	0,0000	0,0063	0,0285	0,0506	0,1361
-0,0013	-0,0044	-0,0114	0,0025	0,0114	-0,0019	-0,0101	-0,0057
0,0171	-0,0541	0,0019	0,1114	-0,0399	-0,0377	-0,1038	-0,0244
-0,0506	-0,0127	-0,0506	0,0000	-0,0253	-0,1013	0,0000	0,0253
-0,0215	-0,0310	-0,0165	-0,0203	-0,0089	-0,0513	0,0051	-0,3943
0,0443	0,0285	-0,0190	0,0380	-0,0190	0,0285	0,0380	0,0601
0,0247	-0,0718	-0,0184	-0,0557	-0,0703	-0,0579	-0,0051	-0,8446
0,0120	-0,0465	0,0070	-0,0430	-0,0196	-0,1339	0,0962	-0,9585
-0,0228	-0,0418	-0,0405	-0,0177	-0,0101	-0,0468	0,0203	-0,6595
0,0228	-0,0025	0,0025	0,0051	-0,0532	0,0405	-0,0076	0,0329
-0,0161	-0,0027	-0,0250	0,0038	-0,0256	-0,0274	-0,0089	0,0508
-0,0092	0,0280	0,0123	-0,0165	0,0003	-0,0106	0,0089	-0,0002
0,0117	-0,0065	-0,0149	0,0241	-0,0104	0,0144	-0,0266	-0,0138
0,0155	0,0005	0,0446	0,0038	-0,0066	0,0138	0,0291	0,0223
0,0149	-0,0017	0,0136	-0,0646	-0,0263	0,0634	0,0494	0,0068
-0,0383	0,0147	-0,0472	-0,0089	-0,0161	0,0090	-0,0089	0,0714
-0,0066	-0,0074	0,0478	-0,0089	0,0155	0,0122	0,0038	-0,0204
0,0165	0,0196	0,0468	0,0177	0,0165	0,0627	0,0557	0,0614
-0,0025	-0,0089	-0,0101	0,0177	0,0228	-0,0165	0,0051	-0,0494
0,0304	-0,0013	0,0203	0,0025	0,0177	0,0519	-0,0228	-0,0468
0,0589	-0,0028	0,0234	0,0278	0,0019	0,0440	-0,0101	0,0434
0,0288	-0,0100	0,0250	-0,0038	-0,0123	0,0400	0,0215	-0,0002
0,0222	0,0016	0,0348	-0,0127	0,0032	0,0332	0,0000	-0,0269
0,0180	-0,0033	-0,0085	-0,0013	-0,0041	0,0049	-0,0013	-0,0043
0,0301	-0,0055	-0,0142	0,0190	0,0142	-0,0087	0,0190	-0,0071
-0,0117	-0,0188	0,0149	0,0013	0,0358	0,0109	0,0013	-0,0242
0,0101	-0,0089	0,0278	0,0304	0,0101	-0,0165	0,0051	-0,0241
0,0228	0,0038	0,0025	-0,0456	-0,0278	-0,0418	-0,0203	0,0013
-0,0237	0,0087	0,0332	-0,0063	0,0174	-0,0008	0,0063	0,0166
0,2525	0,0168	0,0449	0,0582	0,0057	0,0307	0,0076	0,0035
0,0168	0,0809	0,0180	0,0013	0,0199	0,0283	0,0266	0,0533

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
0,0449	0,0180	0,1766	-0,0203	0,0006	0,0421	0,0051	-0,0003
0,0582	0,0013	-0,0203	0,2127	0,0709	-0,0266	0,0228	0,0722
0,0057	0,0199	0,0006	0,0709	0,1766	0,0085	0,0456	0,0256
0,0307	0,0283	0,0421	-0,0266	0,0085	0,2517	0,0114	0,1160
0,0076	0,0266	0,0051	0,0228	0,0456	0,0114	0,2127	-0,0291
0,0035	0,0533	-0,0003	0,0722	0,0256	0,1160	-0,0291	0,8100
0,0396	0,0214	0,0459	0,0570	0,0174	0,1131	-0,0063	0,7381
0,0509	0,0549	0,0345	0,0468	0,0288	0,0986	-0,0038	0,7704
0,0250	0,0400	0,0161	0,0291	0,0092	0,0976	-0,0215	0,7802
0,0449	0,0687	0,0500	0,0177	0,0006	0,1180	-0,0076	0,7339
0,0744	0,0356	0,0301	0,0316	0,0206	0,0451	0,0443	-0,0989
-0,0883	0,0106	-0,0288	-0,0418	0,0288	-0,0217	-0,0038	0,0046
0,0161	-0,0036	-0,0383	-0,0418	-0,0123	0,0337	-0,0291	-0,0318
-0,0044	0,0161	-0,0399	0,0278	0,0019	-0,0003	0,0278	0,0623
0,0497	0,0062	-0,0022	0,0114	-0,0231	-0,0236	-0,0139	0,0369
-0,0009	-0,0128	-0,0149	-0,0013	-0,0231	-0,0172	-0,0139	0,0685
-0,0206	-0,0055	-0,0142	0,0190	0,0142	-0,0087	-0,0190	0,0435
-0,0373	-0,0041	-0,0323	-0,0329	-0,0057	0,0326	0,0051	-0,0415
-0,0741	0,0003	-0,0082	-0,0734	-0,0044	0,0472	0,0025	0,0528
-0,0006	-0,0085	0,0323	-0,0684	-0,0070	0,0180	-0,0177	-0,0472
0,0364	-0,0150	-0,0079	-0,0316	-0,0427	-0,0055	0,0316	0,0150
-0,0044	-0,0155	-0,0019	-0,0481	-0,0234	0,0187	-0,0101	-0,0073
-0,2916	0,1560	1,4137	-0,7213	0,4128	0,6417	-0,0605	-3,3647
-0,0917	-0,1438	0,0658	0,0448	-0,4987	-0,2477	-0,0843	-0,8697
-0,0534	-0,1197	-0,3473	0,4719	-0,0084	-0,3838	0,0061	-0,0781
0,0569	-0,1047	-0,1246	0,1020	-0,0994	-0,1324	-0,0575	-0,0781
0,1194	-0,2305	-0,2893	0,1989	-0,1803	-0,1788	0,0951	-0,1092
-0,0508	-0,0445	0,0536	0,0437	-0,0169	-0,0335	-0,0348	0,0129
-0,0228	0,1394	0,3459	-0,3363	-0,0763	0,0034	0,0396	-0,8068
0,1585	0,0378	0,3193	-0,1405	-0,0218	-0,1084	0,1038	-0,5859
0,0321	-0,1046	-0,0982	0,1115	-0,1955	-0,1111	-0,0834	-0,5833
-0,0126	-0,0472	-0,2374	0,1119	-0,2461	-0,1676	-0,1742	-0,7130
0,2518	-0,1074	0,3982	0,0868	-0,2475	0,3383	0,1767	0,0525
0,3088	-0,0086	0,3951	0,2837	-0,0470	0,4120	0,4318	0,4086

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0909	-0,0814	0,2466	0,1628	-0,0301	-0,0224	0,1970	-0,1343
-0,2550	0,0311	-0,0883	0,1376	-0,0813	-0,2278	-0,0573	-0,4562

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0411	-0,0386	-0,0373	-0,0392	-0,0095	-0,0133	-0,0196	-0,0165
0,3877	0,4244	0,3604	0,3487	0,2706	-0,0187	-0,0788	-0,1310
0,0918	0,1196	0,1335	0,1633	0,0222	0,0563	-0,0386	-0,0797
0,1424	0,1424	0,1487	0,1709	0,0222	0,0158	-0,0285	0,0190
-0,0095	-0,0120	-0,0133	-0,0114	0,0095	0,0133	0,0070	0,0038
-0,0744	-0,1668	-0,2003	-0,1627	-0,0142	-0,0275	-0,0497	0,0057
0,0633	0,0380	0,0886	0,1013	-0,1013	0,0253	0,0127	0,0000
-0,4209	-0,4386	-0,4475	-0,4342	0,0538	0,0171	0,0234	-0,0241
0,0665	0,1044	0,0854	0,0696	0,0348	-0,0475	-0,0285	-0,0316
-0,9478	-0,9744	-1,0130	-0,9677	0,1377	0,0383	0,0415	-0,0677
-0,9984	-0,9744	-1,0383	-0,9551	0,1630	0,0003	-0,0092	-0,0297
-0,7152	-0,7354	-0,7709	-0,7241	0,1203	0,0114	0,0494	0,0304
0,0000	0,0076	0,0177	-0,0101	-0,0253	-0,0051	0,0203	0,0076
0,0340	0,0144	0,0172	0,0256	-0,0277	-0,0046	0,0065	0,0263
-0,0087	0,0236	0,0207	0,0250	0,0024	-0,0081	-0,0065	0,0117
-0,0103	-0,0375	-0,0321	-0,0275	0,0166	-0,0059	-0,0074	-0,0066
0,0372	0,0555	0,0521	0,0573	0,0071	-0,0014	0,0033	0,0073
0,0324	0,0495	0,0454	0,0389	0,0245	-0,0201	0,0005	-0,0098
0,0388	0,0255	0,0441	0,0541	-0,0198	0,0191	-0,0109	0,0231
-0,0340	-0,0093	-0,0350	-0,0155	0,0403	0,0350	-0,0267	0,0168
0,0854	0,0930	0,0842	0,0848	0,0411	-0,0082	-0,0272	-0,0241
-0,0443	-0,0494	-0,0646	-0,0481	0,0063	0,0139	0,0139	0,0203
-0,0570	-0,0595	-0,0544	-0,0684	0,0190	0,0418	0,0418	0,0101
0,0554	0,0592	0,0421	0,0234	0,0332	-0,0421	0,0180	-0,0310
0,0166	-0,0017	0,0081	-0,0003	0,0530	-0,0334	-0,0191	-0,0136
-0,0111	-0,0174	-0,0079	-0,0158	0,0617	-0,0047	-0,0016	-0,0095
0,0055	0,0036	0,0027	-0,0085	0,0134	-0,0153	0,0147	-0,0003
0,0008	-0,0024	-0,0040	-0,0142	0,0055	0,0040	0,0245	0,0079
-0,0024	-0,0131	-0,0312	-0,0231	0,0214	0,0312	0,0328	0,0320
0,0063	-0,0114	-0,0139	-0,0101	0,0063	-0,0114	0,0013	0,0076
-0,0063	0,0266	0,0241	0,0152	-0,0190	0,0139	0,0139	-0,0177
0,0150	0,0182	0,0134	0,0079	-0,0087	0,0372	0,0356	0,0237
0,0396	0,0509	0,0250	0,0449	0,0744	-0,0883	0,0161	-0,0044
0,0214	0,0549	0,0400	0,0687	0,0356	0,0106	-0,0036	0,0161

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
0,0459	0,0345	0,0161	0,0500	0,0301	-0,0288	-0,0383	-0,0399
0,0570	0,0468	0,0291	0,0177	0,0316	-0,0418	-0,0418	0,0278
0,0174	0,0288	0,0092	0,0006	0,0206	0,0288	-0,0123	0,0019
0,1131	0,0986	0,0976	0,1180	0,0451	-0,0217	0,0337	-0,0003
-0,0063	-0,0038	-0,0215	-0,0076	0,0443	-0,0038	-0,0291	0,0278
0,7381	0,7704	0,7802	0,7339	-0,0989	0,0046	-0,0318	0,0623
0,8631	0,7903	0,8362	0,7801	-0,1353	-0,0514	-0,0150	0,0364
0,7903	0,8922	0,8483	0,8193	-0,1005	-0,0381	-0,0460	0,0402
0,8362	0,8483	0,9049	0,8389	-0,1210	-0,0188	-0,0236	0,0484
0,7801	0,8193	0,8389	0,8854	-0,0839	0,0092	-0,0383	0,0994
-0,1353	-0,1005	-0,1210	-0,0839	0,2176	0,0071	-0,0419	0,0396
-0,0514	-0,0381	-0,0188	0,0092	0,0071	0,1960	0,0489	0,0782
-0,0150	-0,0460	-0,0236	-0,0383	-0,0419	0,0489	0,2403	0,0370
0,0364	0,0402	0,0484	0,0994	0,0396	0,0782	0,0370	0,2222
0,0403	0,0511	0,0565	0,0991	-0,0087	0,0321	0,0559	0,1326
0,0720	0,1207	0,1134	0,0864	-0,0150	-0,0122	0,0242	-0,0193
0,0388	0,0483	0,0467	0,0364	-0,0198	0,0166	0,0119	-0,0047
-0,0649	-0,0383	-0,0630	-0,0449	0,0142	0,0630	0,0092	0,0171
0,0459	0,0623	0,0896	0,1057	-0,0459	0,0750	0,0655	0,0259
-0,0744	-0,0756	-0,0636	-0,0690	-0,0142	0,0130	0,0415	-0,0171
0,0672	0,0514	0,0815	0,0427	-0,0229	-0,0182	-0,0166	-0,0491
-0,0206	0,0339	0,0294	0,0361	-0,0047	0,0085	0,0180	0,0070
-3,7191	-3,7922	-3,9483	-2,6116	0,6698	0,4711	0,1455	-0,3778
-0,6574	-0,9966	-0,6794	-0,4469	-0,0489	-0,0877	0,2152	0,3770
0,0555	0,0412	-0,0210	0,1325	-0,2707	-0,3816	0,0118	0,2797
-0,0062	-0,1000	-0,0253	0,1045	-0,1824	-0,0671	0,0763	0,1261
0,2154	-0,1774	0,0717	0,0841	-0,3756	-0,4186	0,2572	0,2966
0,0004	-0,0227	-0,0096	0,0233	-0,0010	0,0159	-0,0042	0,0445
-0,8265	-0,9126	-0,8829	-0,4503	-0,0311	0,1146	0,2673	0,1580
-0,4761	-0,5122	-0,4903	-0,2756	0,1116	-0,1679	0,0622	0,0212
-0,5561	-0,6413	-0,5428	-0,4109	-0,1901	-0,2470	-0,0067	-0,0275
-0,6255	-0,8532	-0,6531	-0,5475	-0,1328	-0,0950	0,2313	0,1409
0,2080	-0,0681	0,1230	0,1218	0,0456	-0,3563	0,0926	0,1067
0,6778	0,4904	0,6901	0,6245	-0,1897	-0,3569	0,1192	0,1640

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,1323	-0,1697	-0,1594	-0,1344	-0,3171	-0,2166	0,0935	0,1030
-0,5510	-0,6927	-0,5319	-0,5364	-0,0407	-0,0010	-0,0758	0,0376

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0108	0,0019	0,0032	0,0241	0,0089	0,0013	0,0158	-0,0038
-0,2225	-0,0579	-0,0649	-0,1259	-0,0867	-0,1019	0,1693	-0,1437
-0,0171	0,0082	0,0475	0,1253	0,0215	0,0519	0,0854	-0,0165
-0,0095	0,0411	-0,0032	0,0063	0,0190	-0,0316	-0,0032	0,0063
-0,0019	-0,0146	-0,0032	0,0013	-0,0089	-0,0266	-0,0032	0,0038
0,0573	-0,0946	0,0301	-0,0677	-0,1652	-0,1728	-0,0016	0,0057
0,0633	0,0000	0,0253	0,0506	0,0886	-0,0253	0,0000	0,0253
-0,0259	-0,0513	-0,0222	0,0342	-0,0367	0,0418	-0,0348	-0,0114
-0,0348	0,0285	-0,0032	-0,0063	-0,0316	-0,0063	0,0601	-0,0063
-0,0642	-0,1402	-0,0585	0,1146	-0,0614	0,1133	-0,1408	-0,0171
-0,0136	-0,1275	-0,0585	0,0892	-0,1247	0,0500	-0,0649	-0,0044
-0,0089	-0,0722	-0,0570	0,1114	-0,0456	0,1038	-0,0190	-0,0456
0,0278	0,0025	0,0127	0,0152	0,0582	0,0354	-0,0127	-0,0051
0,0138	-0,0305	0,0071	-0,0092	-0,0275	-0,0035	-0,0150	-0,0180
0,0116	0,0179	-0,0071	-0,0035	0,0022	-0,0092	-0,0103	0,0054
-0,0078	-0,0394	0,0008	0,0389	0,0155	-0,0263	0,0419	-0,0130
0,0296	0,0486	-0,0087	-0,0282	0,0168	0,0282	-0,0055	0,0389
0,0097	0,0160	-0,0103	0,0104	0,0187	0,0275	0,0372	0,0092
-0,0068	-0,0005	0,0087	-0,0377	0,0263	-0,0256	-0,0071	-0,0085
0,0343	-0,0100	0,0055	0,0826	-0,0180	0,0313	0,0150	-0,0275
-0,0259	0,0247	0,0032	-0,0544	-0,0241	-0,0089	0,0158	-0,0114
0,0089	-0,0165	-0,0063	0,0152	-0,0177	-0,0278	-0,0190	-0,0051
0,0139	-0,0114	-0,0190	0,0456	0,0101	0,0430	0,0316	0,0101
-0,0130	-0,0003	-0,0111	0,0044	-0,0247	-0,0044	0,0206	-0,0057
-0,0138	0,0179	-0,0071	0,0092	-0,0104	-0,0345	0,0150	0,0054
-0,0111	0,0142	-0,0079	0,0032	-0,0032	0,0095	0,0111	-0,0095
0,0081	0,0271	-0,0024	-0,0054	-0,0035	-0,0199	0,0134	0,0060
0,0134	-0,0055	-0,0040	-0,0047	-0,0016	-0,0332	0,0182	-0,0111
0,0204	0,0141	0,0245	-0,0009	-0,0028	0,0136	-0,0166	-0,0123
0,0089	-0,0165	-0,0063	-0,0101	-0,0430	-0,0152	-0,0063	-0,0177
0,0215	0,0342	-0,0063	0,0152	0,0203	0,0354	0,0316	0,0076
0,0024	-0,0166	0,0008	0,0237	0,0079	0,0016	-0,0214	-0,0206
0,0497	-0,0009	-0,0206	-0,0373	-0,0741	-0,0006	0,0364	-0,0044
0,0062	-0,0128	-0,0055	-0,0041	0,0003	-0,0085	-0,0150	-0,0155

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,0022	-0,0149	-0,0142	-0,0323	-0,0082	0,0323	-0,0079	-0,0019
0,0114	-0,0013	0,0190	-0,0329	-0,0734	-0,0684	-0,0316	-0,0481
-0,0231	-0,0231	0,0142	-0,0057	-0,0044	-0,0070	-0,0427	-0,0234
-0,0236	-0,0172	-0,0087	0,0326	0,0472	0,0180	-0,0055	0,0187
-0,0139	-0,0139	-0,0190	0,0051	0,0025	-0,0177	0,0316	-0,0101
0,0369	0,0685	0,0435	-0,0415	0,0528	-0,0472	0,0150	-0,0073
0,0403	0,0720	0,0388	-0,0649	0,0459	-0,0744	0,0672	-0,0206
0,0511	0,1207	0,0483	-0,0383	0,0623	-0,0756	0,0514	0,0339
0,0565	0,1134	0,0467	-0,0630	0,0896	-0,0636	0,0815	0,0294
0,0991	0,0864	0,0364	-0,0449	0,1057	-0,0690	0,0427	0,0361
-0,0087	-0,0150	-0,0198	0,0142	-0,0459	-0,0142	-0,0229	-0,0047
0,0321	-0,0122	0,0166	0,0630	0,0750	0,0130	-0,0182	0,0085
0,0559	0,0242	0,0119	0,0092	0,0655	0,0415	-0,0166	0,0180
0,1326	-0,0193	-0,0047	0,0171	0,0259	-0,0171	-0,0491	0,0070
0,1695	0,0112	0,0008	0,0009	0,0282	-0,0009	-0,0087	0,0250
0,0112	0,2074	0,0324	-0,0117	0,0408	0,0244	0,0356	0,0630
0,0008	0,0324	0,0593	0,0206	0,0111	-0,0079	0,0055	0,0142
0,0009	-0,0117	0,0206	0,2525	0,0361	0,0006	0,0396	-0,0082
0,0282	0,0408	0,0111	0,0361	0,2475	0,0652	0,0174	0,0639
-0,0009	0,0244	-0,0079	0,0006	0,0652	0,2525	-0,0396	0,0082
-0,0087	0,0356	0,0055	0,0396	0,0174	-0,0396	0,2176	0,0206
0,0250	0,0630	0,0142	-0,0082	0,0639	0,0082	0,0206	0,1462
-0,6272	-0,4019	-0,2587	0,4220	1,2596	1,6274	-1,4201	0,4121
0,3985	-0,4496	-0,1824	0,0753	0,4207	0,2032	-0,4970	0,0327
0,1712	0,1029	0,0603	-0,0251	0,3031	0,1416	-0,3796	0,1126
0,1607	-0,0419	-0,0160	-0,0202	0,1647	0,0961	-0,1533	0,0970
0,3177	0,0019	-0,0278	-0,1688	0,0747	0,1333	-0,0667	-0,0421
0,0131	-0,0597	-0,0197	-0,0378	0,0676	-0,0065	-0,0593	-0,0043
0,2297	-0,2292	-0,1131	0,0026	0,3317	0,3709	-0,2172	0,0827
0,1884	-0,0179	-0,1853	-0,2092	0,2244	0,1345	-0,1150	0,0237
0,0702	-0,0139	-0,0828	-0,1308	0,1481	0,3726	-0,1357	0,0579
0,2330	-0,2885	-0,0207	0,0860	-0,0103	0,0077	0,0824	-0,1970
0,1027	-0,2674	-0,0402	-0,0970	0,1411	0,3362	-0,0354	-0,0874
0,2015	-0,1752	-0,0703	-0,1622	0,4075	0,1167	0,2609	-0,0582

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
0,2067	-0,2123	-0,0829	-0,2066	0,1473	0,2889	-0,2399	-0,0400
-0,0604	-0,1243	0,0324	0,2424	0,2917	0,1576	-0,1407	-0,0314

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
-0,5244	-0,2875	-0,0211	-0,0935	0,0283	-0,0693	-0,1594	-0,0348
2,0591	-6,9740	-3,4989	-1,9684	-2,5977	0,1106	-0,6159	0,8383
1,7009	1,9632	0,5105	0,7191	0,1644	0,0085	0,2222	0,0044
-0,7930	-0,0184	-0,3070	-0,2095	-0,0617	0,0503	-0,4079	-0,2462
-0,1832	-0,2796	-0,2308	-0,1419	-0,0992	-0,0180	-0,0380	0,0892
-3,8450	-2,5821	-0,8164	-0,8641	-0,2763	-0,2090	-0,8061	-0,4883
0,5924	2,3608	1,1506	0,7443	0,5468	0,0367	0,1734	0,1975
1,9267	0,4996	0,0353	0,0447	0,0450	0,0180	0,4147	0,2449
-1,9703	-0,5057	-0,0437	-0,2487	-0,0313	-0,0396	-0,4927	-0,2842
5,5261	0,9653	0,0517	0,2240	-0,1236	0,0720	1,1949	0,5573
4,1350	1,2083	-0,2217	-0,0494	0,0548	-0,0128	0,9367	0,6130
3,9942	1,0428	0,1210	0,0635	0,2908	0,0741	0,7837	0,6013
-0,3334	0,7104	0,2916	0,1162	0,4365	0,0266	0,0170	0,0456
-0,2771	0,0912	0,1274	0,1060	0,1649	0,0251	0,0486	-0,1027
-0,1938	-0,5064	-0,4857	-0,2262	-0,3118	-0,0504	0,0109	-0,0353
-0,0893	-0,1724	-0,2174	-0,1368	-0,1785	0,0270	-0,0450	-0,0672
0,7083	0,3159	0,2066	0,0952	0,1171	0,0045	0,1419	0,2359
-0,9845	-0,2733	-0,2936	-0,1694	-0,0666	-0,0494	-0,2157	-0,0498
-0,1509	-0,3997	-0,2387	-0,0823	-0,1321	-0,0172	0,0639	-0,0612
0,1612	0,9124	0,2088	0,2791	-0,0115	0,0711	0,0927	0,0217
-0,2999	-1,0244	-0,5166	-0,4059	-0,3018	-0,0174	-0,1599	0,1209
-0,3790	-0,5390	-0,3197	-0,1990	-0,1920	-0,0361	-0,0368	-0,0076
0,1529	0,2185	-0,1618	0,0701	-0,1153	0,0323	0,0224	-0,0848
0,2703	-0,4015	0,0164	-0,0764	-0,0041	0,0527	0,0662	0,1402
-0,1102	0,0278	-0,1616	-0,0794	-0,0421	0,0040	-0,2081	0,0255
0,8560	-0,0750	-0,3832	-0,1509	-0,2378	-0,0090	0,0777	0,1737
-0,3127	-0,1528	-0,1522	-0,1235	-0,0238	-0,0480	-0,1137	0,0043
-0,3397	-0,3432	-0,2524	-0,1552	-0,0733	-0,0248	-0,0776	0,1210
0,6931	-0,6340	-0,4269	-0,2366	-0,2443	-0,0105	0,0488	0,1090
0,6147	-0,2567	-0,4008	-0,1446	-0,2490	-0,0019	0,1151	0,0544
-0,5727	-0,1681	-0,3514	-0,0787	-0,1642	-0,0728	0,0011	0,0089
0,6074	0,5653	-0,0103	0,0179	-0,1276	0,0371	0,1811	0,1049
-0,2916	-0,0917	-0,0534	0,0569	0,1194	-0,0508	-0,0228	0,1585
0,1560	-0,1438	-0,1197	-0,1047	-0,2305	-0,0445	0,1394	0,0378

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
1,4137	0,0658	-0,3473	-0,1246	-0,2893	0,0536	0,3459	0,3193
-0,7213	0,0448	0,4719	0,1020	0,1989	0,0437	-0,3363	-0,1405
0,4128	-0,4987	-0,0084	-0,0994	-0,1803	-0,0169	-0,0763	-0,0218
0,6417	-0,2477	-0,3838	-0,1324	-0,1788	-0,0335	0,0034	-0,1084
-0,0605	-0,0843	0,0061	-0,0575	0,0951	-0,0348	0,0396	0,1038
-3,3647	-0,8697	-0,0781	-0,0781	-0,1092	0,0129	-0,8068	-0,5859
-3,7191	-0,6574	0,0555	-0,0062	0,2154	0,0004	-0,8265	-0,4761
-3,7922	-0,9966	0,0412	-0,1000	-0,1774	-0,0227	-0,9126	-0,5122
-3,9483	-0,6794	-0,0210	-0,0253	0,0717	-0,0096	-0,8829	-0,4903
-2,6116	-0,4469	0,1325	0,1045	0,0841	0,0233	-0,4503	-0,2756
0,6698	-0,0489	-0,2707	-0,1824	-0,3756	-0,0010	-0,0311	0,1116
0,4711	-0,0877	-0,3816	-0,0671	-0,4186	0,0159	0,1146	-0,1679
0,1455	0,2152	0,0118	0,0763	0,2572	-0,0042	0,2673	0,0622
-0,3778	0,3770	0,2797	0,1261	0,2966	0,0445	0,1580	0,0212
-0,6272	0,3985	0,1712	0,1607	0,3177	0,0131	0,2297	0,1884
-0,4019	-0,4496	0,1029	-0,0419	0,0019	-0,0597	-0,2292	-0,0179
-0,2587	-0,1824	0,0603	-0,0160	-0,0278	-0,0197	-0,1131	-0,1853
0,4220	0,0753	-0,0251	-0,0202	-0,1688	-0,0378	0,0026	-0,2092
1,2596	0,4207	0,3031	0,1647	0,0747	0,0676	0,3317	0,2244
1,6274	0,2032	0,1416	0,0961	0,1333	-0,0065	0,3709	0,1345
-1,4201	-0,4970	-0,3796	-0,1533	-0,0667	-0,0593	-0,2172	-0,1150
0,4121	0,0327	0,1126	0,0970	-0,0421	-0,0043	0,0827	0,0237
155,3030	33,8749	11,9367	12,1035	1,6550	4,1151	29,4642	20,6655
33,8749	47,8071	21,4454	14,2412	11,4541	3,3911	7,6623	4,9616
11,9367	21,4454	21,9197	9,8790	10,8103	2,3362	3,6427	2,5226
12,1035	14,2412	9,8790	6,7200	4,6975	1,4344	3,2462	1,5547
1,6550	11,4541	10,8103	4,6975	10,5346	1,0745	2,4587	2,5328
4,1151	3,3911	2,3362	1,4344	1,0745	0,8885	0,8212	0,5818
29,4642	7,6623	3,6427	3,2462	2,4587	0,8212	10,2088	5,7651
20,6655	4,9616	2,5226	1,5547	2,5328	0,5818	5,7651	7,2933
16,1223	13,7742	8,9774	5,7786	5,7595	0,8139	3,7518	2,8711
2,8403	7,6644	4,7547	3,3077	3,7354	0,4665	1,0456	1,2135
13,4066	17,0411	7,3243	4,7381	5,8902	1,4700	4,5545	2,2360
7,5251	13,3584	6,4642	3,0750	5,7390	0,9952	3,3331	2,4372

The Covariance Matrix of Questionnaire Data Columns 1 to 10							
13,3927	13,2901	6,4841	3,4052	4,8843	1,0492	3,8623	2,4301
17,4901	16,9410	9,1987	5,2403	3,5070	1,0359	3,3011	2,0583

The Covariance Matrix of Questionnaire Data Columns 1 to 10					
0,0207	0,1758	-0,1805	-0,1027	-0,0866	0,0240
-2,4964	-0,9404	-2,5253	-1,6965	-1,9934	-2,7709
0,5998	-0,2647	0,1288	-0,0159	-0,4866	0,6461
-0,3972	-0,4525	-0,4547	-0,3852	-0,2759	-0,2275
-0,2346	-0,1290	-0,1600	-0,1644	-0,1197	-0,1506
-1,1130	-0,4966	-1,3160	-1,8030	-0,8904	-0,9110
0,8165	0,5215	0,3658	0,6506	0,3810	1,2051
0,3108	0,3642	0,0632	-0,2114	0,1744	0,3220
-0,1339	-0,1032	-0,0521	0,0680	-0,2608	-0,1921
0,5726	0,5342	0,5737	-0,3193	0,4015	0,7222
0,5738	0,4444	0,2762	-0,1104	0,5116	0,5678
0,7072	0,8922	0,0783	-0,3004	0,3129	0,5454
0,3643	0,2433	0,6690	0,7239	0,5301	0,2159
-0,1167	0,1497	0,1674	-0,0083	0,1254	0,0081
-0,0947	0,0199	-0,2433	-0,1702	-0,1305	-0,1764
-0,4184	-0,1429	-0,1303	-0,1226	-0,1768	0,0510
0,1507	-0,1737	0,1431	0,3584	0,2495	0,0856
-0,0331	-0,0357	0,0966	0,1389	-0,0439	-0,2814
-0,2479	-0,0493	-0,6264	-0,6251	-0,4265	-0,0041
0,1132	-0,0069	0,6219	0,4882	0,2368	0,1608
-0,3854	-0,2890	-0,6431	-0,4810	-0,5041	-0,4780
-0,3059	-0,0706	-0,3226	-0,2908	-0,0889	-0,2733
-0,1659	-0,0220	0,1688	-0,0243	-0,0565	-0,0186
-0,1281	-0,0413	0,2671	0,2962	0,1473	-0,2795
0,0269	-0,1016	0,1060	0,0058	-0,1862	-0,0207
0,0584	0,0661	0,0157	-0,1394	-0,1228	-0,0366
-0,0636	-0,0740	-0,0558	-0,0432	-0,0490	-0,0765
-0,1537	0,0160	-0,0833	0,1664	-0,0108	-0,1904
-0,2183	-0,0153	-0,2976	-0,4901	-0,2586	-0,3092
-0,0287	0,1167	-0,0226	-0,0845	0,0808	-0,1606
-0,0528	-0,0744	-0,0846	-0,1351	-0,1167	-0,1239
-0,0282	-0,0495	0,3451	0,3611	0,3437	0,1819
0,0321	-0,0126	0,2518	0,3088	-0,0909	-0,2550

-0,1046	-0,0472	-0,1074	-0,0086	-0,0814	0,0311
-0,0982	-0,2374	0,3982	0,3951	0,2466	-0,0883
0,1115	0,1119	0,0868	0,2837	0,1628	0,1376
-0,1955	-0,2461	-0,2475	-0,0470	-0,0301	-0,0813
-0,1111	-0,1676	0,3383	0,4120	-0,0224	-0,2278
-0,0834	-0,1742	0,1767	0,4318	0,1970	-0,0573
-0,5833	-0,7130	0,0525	0,4086	-0,1343	-0,4562
-0,5561	-0,6255	0,2080	0,6778	-0,1323	-0,5510
-0,6413	-0,8532	-0,0681	0,4904	-0,1697	-0,6927
-0,5428	-0,6531	0,1230	0,6901	-0,1594	-0,5319
-0,4109	-0,5475	0,1218	0,6245	-0,1344	-0,5364
-0,1901	-0,1328	0,0456	-0,1897	-0,3171	-0,0407
-0,2470	-0,0950	-0,3563	-0,3569	-0,2166	-0,0010
-0,0067	0,2313	0,0926	0,1192	0,0935	-0,0758
-0,0275	0,1409	0,1067	0,1640	0,1030	0,0376
0,0702	0,2330	0,1027	0,2015	0,2067	-0,0604
-0,0139	-0,2885	-0,2674	-0,1752	-0,2123	-0,1243
-0,0828	-0,0207	-0,0402	-0,0703	-0,0829	0,0324
-0,1308	0,0860	-0,0970	-0,1622	-0,2066	0,2424
0,1481	-0,0103	0,1411	0,4075	0,1473	0,2917
0,3726	0,0077	0,3362	0,1167	0,2889	0,1576
-0,1357	0,0824	-0,0354	0,2609	-0,2399	-0,1407
0,0579	-0,1970	-0,0874	-0,0582	-0,0400	-0,0314
16,1223	2,8403	13,4066	7,5251	13,3927	17,4901
13,7742	7,6644	17,0411	13,3584	13,2901	16,9410
8,9774	4,7547	7,3243	6,4642	6,4841	9,1987
5,7786	3,3077	4,7381	3,0750	3,4052	5,2403
5,7595	3,7354	5,8902	5,7390	4,8843	3,5070
0,8139	0,4665	1,4700	0,9952	1,0492	1,0359
3,7518	1,0456	4,5545	3,3331	3,8623	3,3011
2,8711	1,2135	2,2360	2,4372	2,4301	2,0583
11,1831	6,8974	5,1602	4,4684	4,5916	6,1368
6,8974	10,2547	1,6683	1,4877	2,4891	3,3254
5,1602	1,6683	17,1349	15,9045	11,4809	4,7254
4,4684	1,4877	15,9045	20,8577	12,3998	4,0001

4,5916	2,4891	11,4809	12,3998	12,0855	3,8319
6,1368	3,3254	4,7254	4,0001	3,8319	9,2385

Appendix B

Results of Questionnaire Data After Principal Component Analysis with Singular Value Decomposition

-0,11944	-0,11589	-0,10863	-0,11004	-0,11876	-0,11227	-0,11222	-0,10904
0,046102	-0,16181	-0,04745	-0,02897	0,025556	-0,19629	0,028745	-0,03655
-0,08598	0,055911	-0,09855	-0,00214	0,016587	-0,0088	-0,03689	-0,20998
0,043578	0,06234	-0,06155	0,029776	0,095312	0,049713	0,132384	-0,01724

-0,11171	-0,11425	-0,09915	-0,11247	-0,11804	-0,11614	-0,11416	-0,11395
0,002656	0,014582	-0,20727	0,072188	0,114597	0,13465	0,054653	0,113409
0,179765	-0,088	-0,04887	-0,08542	0,161279	-0,05542	-0,00144	0,094492
-0,20981	-0,04957	-0,11031	0,055657	0,092196	-0,09693	0,201554	0,051243

-0,11375	-0,11707	-0,11043	-0,10466	-0,1121	-0,10894	-0,11811	-0,11904
-0,02323	-0,01771	-0,05815	-0,08558	0,141445	0,020937	0,117192	0,080441
-0,02026	0,056167	-0,1068	-0,17295	-0,12039	0,131176	0,02666	0,021745
0,009249	-0,01667	-0,03149	-0,03947	-0,05226	-0,08639	0,014954	-0,01043

-0,12035	-0,1104	-0,10863	-0,11227	-0,11467	-0,11876	-0,10542	-0,11651
-0,08174	0,238993	-0,04745	-0,19629	0,113538	0,025556	-0,18866	0,096249
-0,06559	0,042431	-0,09855	-0,0088	0,149075	0,016587	0,122977	-0,14711
-0,0477	0,042244	-0,06155	0,049713	0,09893	0,095312	0,078352	0,015423

-0,11375	-0,10956	-0,11075	-0,10574	-0,10466	-0,10601	-0,11707	-0,10502
-0,02323	-0,00873	0,015152	0,031429	-0,08558	-0,07116	-0,01771	-0,0713
-0,02026	0,177279	-0,11829	-0,04974	-0,17295	-0,08	0,056167	0,085828
0,009249	-0,16974	0,07346	-0,07111	-0,03947	0,041734	-0,01667	-0,00503

-0,11416	-0,12161	-0,11171	-0,11904	-0,11285	-0,10404	-0,10997	-0,1094
0,054653	0,185146	0,002656	0,080441	-0,08967	0,034109	-0,09182	0,044831
-0,00144	0,030636	0,179765	0,021745	-0,05104	-0,03196	-0,10616	-0,18156
0,201554	-0,04199	-0,20981	-0,01043	-0,12488	-0,46347	0,010783	-0,0293

-0,1066	-0,11301	-0,10865	-0,10284	-0,10905	-0,11096	-0,10542	-0,1116
-0,02634	0,21105	-0,07753	-0,15987	-0,15365	0,061097	-0,18866	-0,12397
-0,14437	-0,08846	0,154812	-0,03026	-0,07569	0,110938	0,122977	0,005489
0,082872	-0,00446	0,076172	0,18196	0,141268	-0,05939	0,078352	0,078749

-0,1104	-0,11679	-0,10427	-0,11057	-0,11024	-0,10404	-0,11651	-0,1125
0,238993	-0,16113	-0,09595	0,044674	-0,08363	0,034109	0,096249	-0,17175
0,042431	-0,19659	0,04696	0,219121	0,320792	-0,03196	-0,14711	-0,08277
0,042244	-0,0798	0,053674	0,011815	-0,01414	-0,46347	0,015423	0,050797

-0,11204	-0,11036	-0,1094	-0,114	-0,12161	-0,11285	-0,11461	-0,10764
-0,06835	-0,03601	0,044831	-0,11077	0,185146	-0,08967	-0,13954	-0,03904
0,126725	-0,05687	-0,18156	0,047099	0,030636	-0,05104	0,160707	0,067877
-0,01144	0,140998	-0,0293	0,033773	-0,04199	-0,12488	0,020775	-0,03036

-0,10956	-0,11502	-0,11119	-0,10981	-0,10535	-0,10502	-0,11192	-0,11957
-0,00873	0,11934	0,104599	0,165725	-0,1054	-0,0713	0,118039	0,266147
0,177279	0,095701	0,053232	-0,18419	0,061743	0,085828	-0,13927	0,101739
-0,16974	0,103589	0,104936	0,027847	0,046996	-0,00503	0,029395	0,147132

Appendix C

Results of Questionnaire Data After Principal Component Analysis with Eigen Value Decomposition

0,119439	0,11589	0,108632	0,11004	0,118763	0,112266	0,112222	0,109037
-0,0461	0,161813	0,047447	0,028972	-0,02556	0,196294	-0,02874	0,036551
-0,08598	0,055911	-0,09855	-0,00214	0,016587	-0,0088	-0,03689	-0,20998
-0,04358	-0,06234	0,061549	-0,02978	-0,09531	-0,04971	-0,13238	0,017242

0,111714	0,114245	0,099151	0,11247	0,118043	0,116144	0,114163	0,113945
-0,00266	-0,01458	0,207271	-0,07219	-0,1146	-0,13465	-0,05465	-0,11341
0,179765	-0,088	-0,04887	-0,08542	0,161279	-0,05542	-0,00144	0,094492
0,209813	0,049569	0,110305	-0,05566	-0,0922	0,096929	-0,20155	-0,05124

0,113753	0,117068	0,11043	0,104656	0,1121	0,108945	0,118114	0,119039
0,023226	0,017713	0,058154	0,085579	-0,14144	-0,02094	-0,11719	-0,08044
-0,02026	0,056167	-0,1068	-0,17295	-0,12039	0,131176	0,02666	0,021745
-0,00925	0,01667	0,031485	0,039474	0,05226	0,086392	-0,01495	0,010427

0,120351	0,110402	0,108632	0,112266	0,114666	0,118763	0,105424	0,116505
0,081735	-0,23899	0,047447	0,196294	-0,11354	-0,02556	0,188662	-0,09625
-0,06559	0,042431	-0,09855	-0,0088	0,149075	0,016587	0,122977	-0,14711
0,047698	-0,04224	0,061549	-0,04971	-0,09893	-0,09531	-0,07835	-0,01542

0,113753	0,109564	0,110749	0,105739	0,104656	0,106005	0,117068	0,105019
0,023226	0,008729	-0,01515	-0,03143	0,085579	0,07116	0,017713	0,071295
-0,02026	0,177279	-0,11829	-0,04974	-0,17295	-0,08	0,056167	0,085828
-0,00925	0,169735	-0,07346	0,071106	0,039474	-0,04173	0,01667	0,005027

0,114163	0,121612	0,111714	0,119039	0,112852	0,104044	0,109969	0,109405
-0,05465	-0,18515	-0,00266	-0,08044	0,089665	-0,03411	0,091816	-0,04483
-0,00144	0,030636	0,179765	0,021745	-0,05104	-0,03196	-0,10616	-0,18156
-0,20155	0,041987	0,209813	0,010427	0,124884	0,463473	-0,01078	0,0293

0,106601	0,113008	0,108646	0,102838	0,109051	0,110964	0,105424	0,111603
0,026338	-0,21105	0,077534	0,159867	0,153646	-0,0611	0,188662	0,123967
-0,14437	-0,08846	0,154812	-0,03026	-0,07569	0,110938	0,122977	0,005489
-0,08287	0,00446	-0,07617	-0,18196	-0,14127	0,059392	-0,07835	-0,07875

0,110402	0,116786	0,104268	0,110565	0,110244	0,104044	0,116505	0,112504
-0,23899	0,161131	0,095952	-0,04467	0,083635	-0,03411	-0,09625	0,171746
0,042431	-0,19659	0,04696	0,219121	0,320792	-0,03196	-0,14711	-0,08277
-0,04224	0,079796	-0,05367	-0,01181	0,014142	0,463473	-0,01542	-0,0508

0,112041	0,110361	0,109405	0,114	0,121612	0,112852	0,11461	0,107638
0,068351	0,036007	-0,04483	0,110774	-0,18515	0,089665	0,13954	0,039037
0,126725	-0,05687	-0,18156	0,047099	0,030636	-0,05104	0,160707	0,067877
0,011442	-0,141	0,0293	-0,03377	0,041987	0,124884	-0,02077	0,030357

0,109564	0,115023	0,111188	0,10981	0,105352	0,105019	0,111918	0,119572
0,008729	-0,11934	-0,1046	-0,16572	0,105404	0,071295	-0,11804	-0,26615
0,177279	0,095701	0,053232	-0,18419	0,061743	0,085828	-0,13927	0,101739
0,169735	-0,10359	-0,10494	-0,02785	-0,047	0,005027	-0,0294	-0,14713

Appendix D

Results of Questionnaire Data After Locally Linear Embedding

-0,49751	-1,76007	-1,17372	-0,93574	-0,96826	-1,26904	-0,89362	-0,43913
0,235134	0,175018	0,194179	0,215191	0,210198	0,189003	0,209583	0,234813
0,42403	0,402542	0,411246	0,416736	0,415662	0,409476	0,416316	0,424573
-1	-1	-1	-1	-1	-1	-1	-1

-0,51388	-0,85791	-0,67225	-0,64722	1,075196	2,416308	-0,48562	2,4208
0,235953	0,219968	0,219153	0,23011	0,241591	0,244387	0,235598	0,243029
0,423966	0,418295	0,419997	0,421868	0,425908	0,428018	0,424225	0,428385
-1	-1	-1	-1	-1	-1	-1	-1

-0,41313	-0,5122	-0,48528	0,493353	-0,46072	-0,34552	-0,49366	-0,50866
0,234142	0,235703	0,235469	0,239572	0,235001	0,23306	0,235774	0,23593
0,424646	0,423984	0,424203	0,425207	0,424371	0,425215	0,424167	0,424016
-1	-1	-1	-1	-1	-1	-1	-1

-0,47094	-0,48192	-0,48259	-0,41036	-0,46048	-0,36985	-0,47066	-0,34788
0,234794	0,235247	0,234955	0,234175	0,235057	0,233561	0,235127	0,23292
0,42425	0,424211	0,424145	0,424731	0,424391	0,425056	0,424294	0,425174
-1	-1	-1	-1	-1	-1	-1	-1

1,042129	-0,35239	-0,35578	1,992359	2,233835	0,211777	2,265335	-0,52162
0,238723	0,232836	0,233086	0,244249	0,244584	0,236008	0,245718	0,234584
0,426578	0,42511	0,425072	0,427149	0,427577	0,425543	0,427345	0,425755
-1	-1	-1	-1	-1	-1	-1	-1

-0,61445	-0,60241	-0,68326	-0,67464	-0,77433	2,326705	1,869461	2,435363
0,236138	0,23594	0,237342	0,237296	0,214203	0,244093	0,242308	0,245102
0,425309	0,425379	0,424928	0,42484	0,418179	0,427907	0,427399	0,427872
-1	-1	-1	-1	-1	-1	-1	-1

2,404909	-0,47609	-0,39556	0,566698	0,744614	0,627527	-0,36177	-0,42551
0,245247	0,235246	0,233484	0,23874	0,239185	0,236888	0,233208	0,23439
0,427769	0,424216	0,424774	0,4256	0,425815	0,426223	0,425063	0,424545
-1	-1	-1	-1	-1	-1	-1	-1

0,081072	-0,15573	0,169557	-0,08772	-0,59492	-1,02731	-0,12343	-0,36682
-0,63923	5,858259	-1,12437	-1,78496	-1,22257	-0,50118	-1,80213	-1,52403
-2,28125	-5,32885	-2,04412	-1,10278	-0,67172	-0,11615	-1,11849	-0,90345
-1	-1	-1	-1	-1	-1	-1	-1

0,144476	-0,20526	0,189854	0,312409	0,366406	0,236665
-2,24766	-1,80385	-0,38469	-1,23455	-2,36232	-2,2187
-1,46508	-1,1198	-2,48365	-2,08991	-1,55319	-1,44082
-1	-1	-1	-1	-1	-1

Appendix E

Results of Questionnaire Data After Isomap Algorithm

8,904657	-38,2757	-38,2757	-21,3143	29,25406	29,25406	-9,48069	26,02306
19,67622	16,47583	16,47583	6,461175	17,58479	17,58479	-1,43506	4,548702

0,03026	0,03026	31,01803	18,18208	43,88201	43,88201	-7,81298	-7,81298
3,236395	3,236395	11,73487	11,58695	-6,26967	-6,26967	-20,2557	-20,2557

49,48015	39,17203	-30,432	-30,432	-13,4529	-13,4529	-34,6252	-34,6252
10,82767	14,58085	15,00112	15,00112	14,05896	14,05896	9,950731	9,950731

-24,1863	-7,82906	-30,447	-48,3527	10,99216	35,13722	24,26682	7,009493
-5,66545	-20,1814	15,73666	20,92165	18,22521	18,14839	16,65995	-5,73741

5,982242	8,960189	8,960189	21,45968	40,27446	51,46008	-12,7415	19,07437
0,408573	-4,68855	-4,68855	-36,8122	-25,8929	17,8003	15,33338	17,7781

54,90136	54,90136	0,785066	0,785066	13,99109	-10,5316	-15,0706	-41,2194
3,418793	3,418793	-12,0263	-12,0263	13,75581	-10,3023	-23,159	13,59174

-38,5851	6,331359	6,331359	-10,7439	-25,4153	11,30748	11,30748	-44,3594
1,557339	7,991642	7,991642	-2,18026	-18,0317	-20,1819	-20,1819	19,00204

11,80054	-20,4539	-20,9687	-20,9687	39,08455	23,98341	48,06616	48,06616
10,71626	-1,08785	-36,8329	-36,8329	13,87682	22,52802	-3,63625	-3,63625

26,70551	26,70551	-1,1045	-25,1165	-17,0624	-17,0624	-19,6357	-22,2816
-22,1789	-22,1789	12,46417	2,930689	8,027264	8,027264	-9,29755	-35,0564

-41,1461	-41,1461	-25,7072	0,87565	-14,1089	-11,3191	-42,1341	-8,9281
1,027233	1,027233	-3,95671	-35,9254	-28,0018	7,484794	15,61349	-32,6316