Persian Handwriting Analysis Using Functional Principal Components

Yasaman Hosseinkashi[†,*] and Khalil Shafie[‡]

[†] University of Waterloo
[‡] University of Northern Colorado

Abstract. Principal components analysis is a well-known statistical method in dealing with large dependent data sets. It is also used in functional data for both purposes of data reduction as well as variation representation. On the other hand "handwriting" is one of the objects, studied in various statistical fields like pattern recognition and shape analysis. Considering time as the argument, the handwriting would be an infinite dimensional data; a functional object. In this paper we try to use the *functional principal components analysis* (FPCA) to the Persian handwriting data, analyzing the word *Mehr* which is the Persian term for Love.

Keywords. Principal components analysis; functional data analysis; B-spline smoothing; on-line handwriting.

1 Introduction

Handwriting is not a new topic in statistical literature. There are various attempts to analyze handwriting in order to signature verification (see for example Plamond and Lorette, 1989) or biometric identification (see for example Soldek, et al., 1997) or other psychological purposes. These works are mostly zoomed on off-line handwriting. All the information needed in off-line studies is the final pattern of written items. For a detailed example of off-line studies see McKeague (2004). Progress in technology has been fascinating for researchers to work with on-line data; consisting the velocity

---

[*] Corresponding author

and pen position information. Ramsay and Silverman (1997) applied functional methods for an example of on-line handwriting data. Such data has been studied in similar way also by Hasti et al. (1991) and Kashi et al. (1997). Also Plamondo and Srihari (2000), Dolfing, et al. (1998), Nelson, et al. (1994) and Jain, et al. (2002) give examples of both methods. In this paper, we continue Ramsey and Silverman's line of work and investigate the variational structure of a set of Persian on-line handwriting data by the help of functional principal components.

Persian, the literary language of Iran and much of Afghanistan, Tajikistan and Pakistan, is written in a version of the Arabic script called Persian or Farsi. Persian handwriting has been considered as an art for a long period and Persian calligraphy is an art in which laws of mathematics and nature are obeyed. It has a mysterious power that enables the artist to create beautiful pieces of calligraphy by using several forms of the same letter, or by employing various forms of the words and using them in different compositions. But, there has been not much study on statistical analysis of Persian handwriting. An off-line recognition system for Persian handwriting, using hidden Markov models is presented in Haji (2005). Khorsheed (2000), after transforming Arabic off-line handwriting into a normalized polar map applies $2D$ Fourier transform to extract features of the handwriting. Trenkle, et al. (1995) have used neural network for Arabic, off-line word recognition. Al-Yousefi and Upda (1992) used a quadratic Bayesian classifier to recognize Arabic characters. To the best of our knowledge there has been no study on online analysis of Persian handwriting.

The paper is organized as follows. Section 2 introduces the handwriting as a set of vector valued functions or curves. A quick review of bivariate functional principal components and some preliminaries are given in Section 3. Section 4 contains the detailed results of applying FPCA to Persian handwriting data.

2   Persian Handwriting and FDA Methods

Persian or Farsi, descended from Ancient Persian, the literary language of the Persian Empire and one of the great classical languages, is a member of the Iranian branch of Indo-European languages and the main language of Iran and some other neighbor countries. Modern Persian appeared during the 9th century and under Mongolian and Turkish rulers was adopted as the language of government in Turkey, central Asia and India, where it was

used for centuries, and until after 1900 in Kashmir. This language has been
written with a number of different scripts, including old Persian cuneiform,
Pahlavi, Aramaic, and Avestan. Today, this language is written in a version
of the Arabic script in which four extra letters are added to the original
Arabic script in order to write several pronunciations of a class of Asian
languages including Persian, Urdu, Pashto, Kashmiri and Sindhi. Under the
Ottoman Empire (before 1920), Turkish was also written in this script.

The Persian script is written from right to left in horizontal lines and
most letters are connected to the base line. Unlike English, there is no
upper and lower case for letters. But, every letter can have up to four
different shapes, depending on the location of the letter within a word and
some other different shapes depending on the style of handwriting. The
most distinguishing feature of the Persian script is that the words, whether
machine-printed or handwritten, are cursive and letters belonging to the
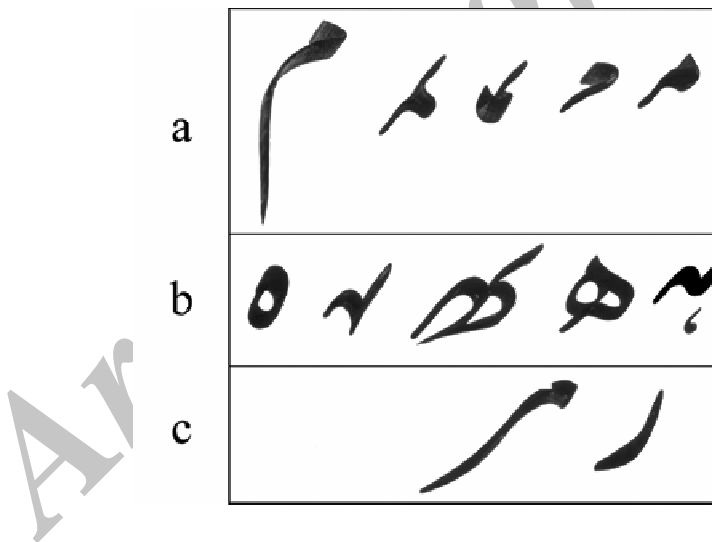same word get connected whenever possible.

Figure 1. All possible shapes of letters in *Mehr*. a) Five possible shapes of the first letter. b)
Five possible shapes of the second letter. c) Two possible shapes of the last letter. Two first
shapes from right in (a) and (b) are used at the beginning of a word, and two middle ones
are used at the middle part of a word and the last one is used at the end of a word. The first
shape from right in (c) is used separately at the beginning and end of a word and the second
shape in (c) is used for the pronunciation of R when there is a connectable letter before it.

To see the ability of functional principal component for analyzing Persian handwriting, a simple experiment is designed and 32 randomly chosen university student were asked to write the word *Mehr*. This word, as the first Persian sample to be analyzed by functional methods, is an old, deep meaning word. In ancient time it was used by the meaning of *sun* and nowadays it means to have affection to someone or somebody. *Mehr* consists of three letters. The first letter has the sound of *M* and can be written in five different shapes. The Second letter with the sound of *h* has also five different shapes. The third letter in *Mehr* pronounced like *r* and only has two different shapes. All possible shapes of these letters, using simple calligraphy, are illustrated in Figure 2. As it is seen from the figure, there would be different combinations for writing the word *Mehr*. One of these combinations, in a simple calligraphy, is displayed in Figure 2.

In off-line studies, "handwriting data" refers to the final result of writing. It means that the statistical object in such studies is a shape or pattern. Considering time as an influential factor results in alternative methods in handwriting analysis, called online methods.

From this point of view, statistical object is the manner of hand's movements, consisting velocity, timing and shifts during the period of writing a word. These information can be formulated by the help of couples: $(x(t), y(t))$. Horizontal movements are introduced by the function $x(t)$ and $y(t)$ stands for vertical movements where x and y are real functions defined on the interval $[a, b]$ which refers to the time period of writing. Figure 3 shows a sample $x(t)$ and $y(t)$ which are combined to a "Mehr" curve.

Figure 2. Three proper choices are connected to form *Mehr*.

3   Preliminaries

3.1   Registration and Smoothing

The most preliminary step in FDA is to produce a functional data set from discrete observations. The basic thought in most FDA researches, as in this work, is that the variables under study are smooth continuous functions. In practice a first record of functional data observations is a group of function snapshots on a finite set of argument values. Therefore the elementary observations have two undesired characteristics: discreteness and having some observational noise. The first characteristics is dealt with smoothing methods and the second one is captured by registration techniques. The most
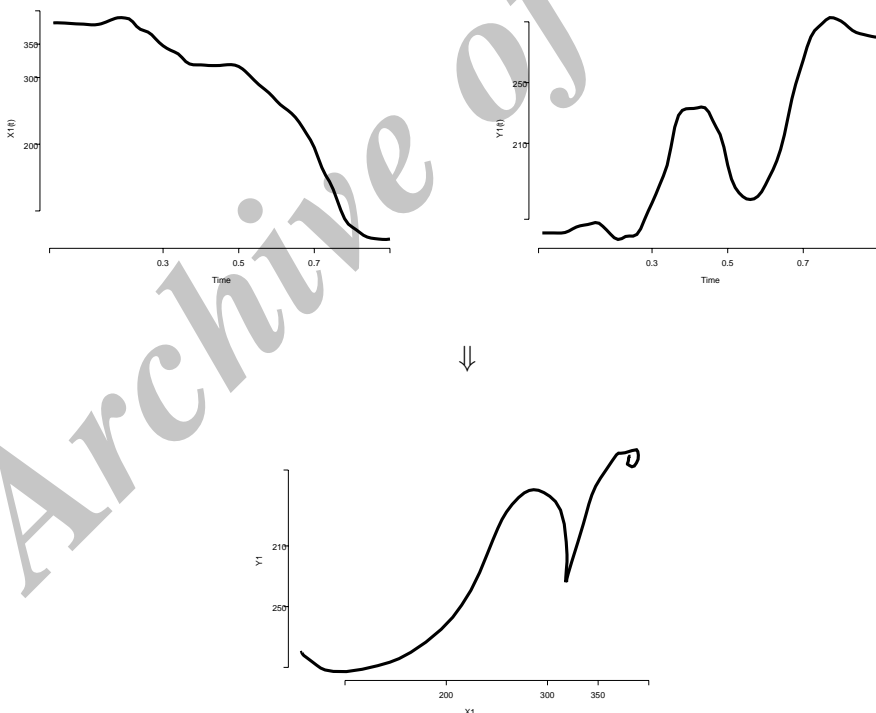
Figure 3. The conversion of simultaneous functions of time to a written word.

commonly used smoothing method in FDA context is representing the observed function $z(t)$ as a linear combination of a finite number of known basis functions $\phi_1, \ldots, \phi_K$. Basis functions can be fourier series, polynomials, regression splines, wavelet basis, etc. The method is to find the coefficient c_k in the expansion $\hat{z}(t) = \sum_{k=1}^{K} c_k \phi_k(t)$ such that the least square criterion:

$$\sum_{j=1}^{J} \left[z(t_j) - \sum_{k=1}^{K} c_k \phi_k(t) \right]^2$$

is minimized simultaneously for all observations (J in this formula is the number of discrete observations). The basis $\{\phi_k\}$ should be chosen as to have features matching those known to belong to the functions being estimated. In handwriting data the best choice for couples of curves (x, y) are B-splines basis functions.

Other important preliminary task is the registration of smoothed observed functions. In most functional data studies noise is a part of data. The term *noise* here means the type of variation among recorded data that is not of our interest and can frustrate even the simplest analysis of replicated functional data set. For example in some cases the domain varies from one individual to another and makes it impossible to compare special characteristics of observed functions. Undesired variation can have more complex effects that should be considered in the choice of applied registration method. Different registration methods as listed in Ramsay and Silverman (1997) are: shift, feature or landmarks and more general transformations. In shift registration method the domain of observed functions are shifted to a standard domain (for example in curve registration the domains are shifted to a standard interval). The method of feature or landmark registration requires a suitable transformation h_i applied to the domain of the ith observation in order to find registered function $x_i^*(t) = x_i(h_i(t))$ such that x_i^* has more or less identical argument values for any given landmark. Landmarks are some characteristic that one can associate with a specific argument value t. The third method, more general transformations, consists of applying a strictly increasing or monotonic function h_i to transform the argument values and satisfy some criterion such as having aligned local feature or minimizing a known measure. For more details on this see Ramsay and Silverman (1997).

3.2 Bivariate Functional Principal Components

Principal components analysis is a technique to produce weight functions ξ_i called eigenfunctions to clarify different types of variations and provide an optimum basis set to reproduce original data. Expansion of each observation in terms of these basis functions approximates original function as closely as possible. This is an important practical matter in FDA, specially in huge data sets and complex computations.

Principal components analysis for bivariate functional data is based on decomposing the covariance function to its eigenfunctions. In the bivariate functional data set $\{(x_i(t), y_i(t)); i : 1, \ldots, n\}$ covariance is defined as

$$V = \begin{bmatrix} v_x & v_{xy} \\ v_{yx} & v_y \end{bmatrix}$$

where:

$$v_x(s, t) = \frac{1}{n} \sum_{i=1}^{n} [x_i(t) - \bar{x}(t)][x_i(s) - \bar{x}(s)],$$

$$v_{xy}(s, t) = v_{yx}(t, s) = \frac{1}{n} \sum_{i=1}^{n} [x_i(t) - \bar{x}(t)][y_i(s) - \bar{y}(s)],$$

and $\bar{x}(t) = \frac{1}{n} \sum_{i=1}^{n} x_i(t)$, $\bar{y}(t) = \frac{1}{n} \sum_{i=1}^{n} y_i(t)$. Eigenfunctions of covariance function, classify the variation among observations in corresponding data set. These functions are ordered descending according to the magnitude of corresponding eigenvalues. The dominant modes of variation are detected by the first eigenfunctions.

Eigenvalues and eigenfunctions in a bivariate functional data set $\{z_i = (x_i(t), y_i(t)); i : 1, \ldots, n, t \in [0, 1]\}$ have the form of $\{(\lambda_k, \xi_k)\}_{k=1}^{\infty}$ where $\xi_k = (\xi_k^x, \xi_k^y)$ is the solutions of integral equation:

$$\begin{cases} \int v_x(s, t) \xi_x(s) ds + \int v_{xy}(s, t) \xi_y(s) ds = \lambda \xi_x(t) \\ \\ \int v_{yx}(t, s) \xi_x(s) ds + \int v_y(s, t) \xi_y(s) ds = \lambda \xi_y(t). \end{cases} \quad (1)$$

As an approximation to these equations, (1) can be converted to a matrix version $V\underline{\xi} = \lambda\underline{\xi}$. In this notation the matrix V is defined as:

$$V = \begin{bmatrix} V_x & V_{xy} \\ V_{yx} & V_y \end{bmatrix},$$

where $V_x = [v_x(t_i, t_j)]_{i,j=1}^K$ and other covariances V_y and V_{xy} are defined in a similar manner. For more details see Ramsay and Silverman (1997).

4   Analysis of Handwriting Data

In this section bivariate FPCA method is applied to the Persian handwriting data. This method has two exploratory advantages. The first advantage is to understand the nature of variation in the handwriting of a group of people. The whole information contained in second moment can be classified, saved and interpreted through eigenfunctions. The second one is to produce an empirical orthonormal basis that can be used in future computations.
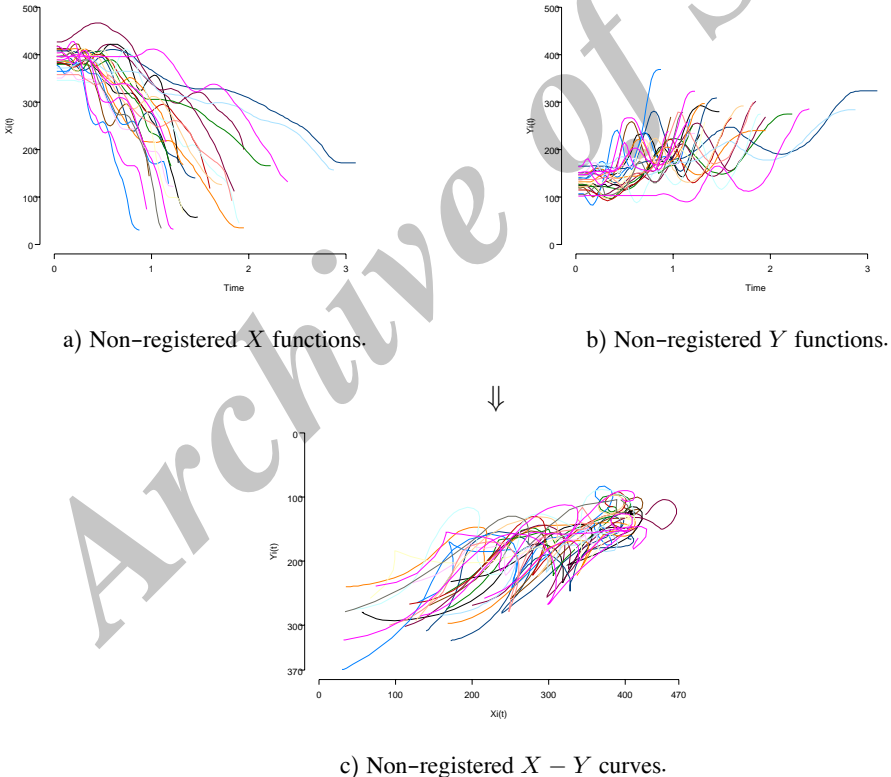
a) Non−registered X functions.   b) Non−registered Y functions.

$\Downarrow$

c) Non−registered $X − Y$ curves.

Figure 4. Non-registered data.

4.1 Row Data and Registration

In our experiment 32 college students were asked to write the form of the term *Mehr* as in 2 on a draw-pad using a special pen while the pen position was recorded 40 times per second. As it is clear from Figure 2 the curvature of this term is not too high so the frequency of 40 times per second seems quite enough to reconstruct the written word efficiently. Resulting data are 32 three column matrices; consisting of (t_{ij}, x_{ij}, y_{ij}); $j = 1, \ldots, J_i$ and $i = 1, \ldots, 32$ where t_{ij} is the jth recorded time for ith observation and (x_{ij}, y_{ij}), respectively, show the horizontal and vertical position of the pen at time t_{ij}. B-spline basis functions are used to smooth these observations.

The row curves in Figure 4 suggest that a common pattern of handwriting exist for all individuals, but it is also shown in this figure that what we call *common pattern* can not be presented by the mean functions $(\bar{x}(t), \bar{y}(t))$. This comes from the fact that differences among these row curves are caused by a combination of linear and nonlinear variations. Linear part is the one that we are interesting to explore. In other words the main question in this study is that how the total time is distributed in the writing period. Nonlinear variation or noise will be removed by applying landmark registration method. Four landmarks are defined as follows (illustrated in Figure 5):
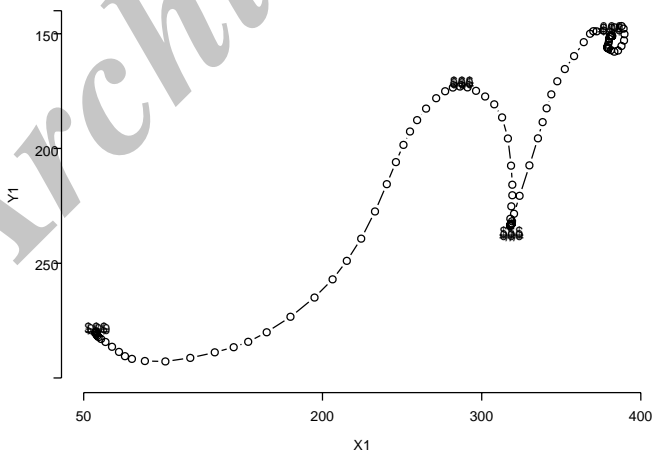

Figure 5. Landmarks in the first observation.

t_1^* the starting time

t_2^* the time of writing the lowest part of the middle letter

t_3^* the time of passing the connection of the middle and the last letter

t_4^* the ending time

Landmark vector ($l.d_i = (t_{i1}^*, t_{i2}^*, t_{i3}^*, t_{i4}^*)$) for the ith subject obtained and the means of t_{i2} and t_{i3} are computed as follows: $\bar{t}_2^* = 0.3$, $\bar{t}_3^* = 0.7$. The values of t_1^* and t_4^* was decided to be 0 and 1 before. Linear transformations:

$$s = h_i(t) = \begin{cases} \frac{0.3}{t_{i2}^* - t_{i1}^*} t - \frac{0.3 t_{i1}^*}{t_{i2}^* - t_{i1}^*} & t \in [t_{i1}^*, t_{i2}^*] \\[2mm] \frac{0.4}{t_{i3}^* - t_{i2}^*} t + \frac{0.3 t_{i1}^* - 0.7 t_{i2}^*}{t_{i3}^* - t_{i2}^*} & t \in [t_{i2}^*, t_{i3}^*] \\[2mm] \frac{0.3}{t_{i4}^* - t_{i3}^*} t + \frac{0.7 t_{i4}^* - t_{i3}^*}{t_{i4}^* - t_{i3}^*} & t \in [t_{i4}^*, t_{i3}^*] \end{cases}$$
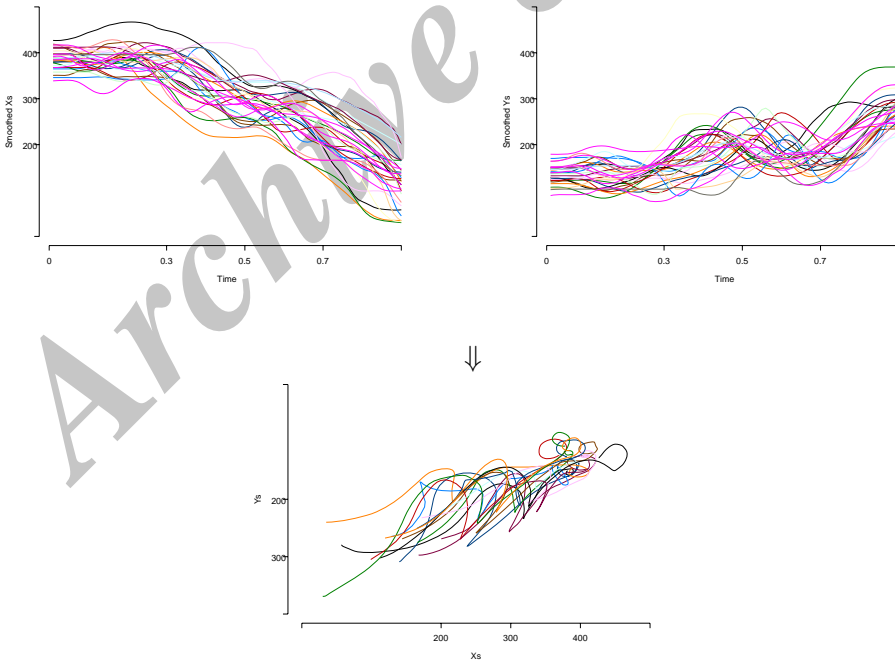

⇓

Figure 6. Registered and Smoothed Data.

are applied on the domains of each observed set $(x_i(t), y_i(t))$ pushing the second and third landmarks to be on the mean position computed before and transferring the starting and ending points to 0 and 1. Therefore the domain of all registered data are converted from $[t_{i1}, t_{iJ_i}]$ to standard interval $[0, 1]$.

4.2   Summary Statistics After Registration

To explore the handwriting data, we start with elementary statistics; mean and covariance functions. The mean functions $\bar{x}(t)$ and $\bar{y}(t)$ along with the mean curve $(\bar{x}(t), \bar{y}(t))$ for the data are computed and shown in Figure 7. As it is clear from the figure, because of registration, the mean curve resembles and preserves the common features of the observations. This is the center that we are going to study the linear variation about. The whole information about this variation is found in covariance functions or equivalently correlation functions (standardized covariance at times s and t), illustrated in Figure 7.

The figures represent unit correlation between functions at identical times and the correlation falls down near the boundaries where the t and s arguments are far from each other. Main similar fact about the two figures of X-correlation and Y-correlation is that high correlation values can be found on couples (s, t) in which the difference $|s - t|$ is about 0.6. Another similar fact is that when $0 < |s - t| < 0.6$ the correlation values fall dramatically. This fall is much more in Y-correlation function (Note that the whole period is standardized to the $[0, 1]$).

In conclusion, although on the very far distances ($|s - t|$ near to 1) the correlation is too low but there are long intervals, for example the time of writing the first and last letter of the word, that takes the biggest values of correlation (except for the case $t = s$).

All these statistics together, can make an elementary concept of the structure of this set of handwriting records. They describe the mean structure and the types of differences from this structure but because of the complexity of functional data sets, this description is not clear especially in variation explanation. More detailed and useful information will be represented by principal components or equivalently eigenfunctions.

J. Statist. Res. Iran **6** (2009): 141–160

www.SID.ir
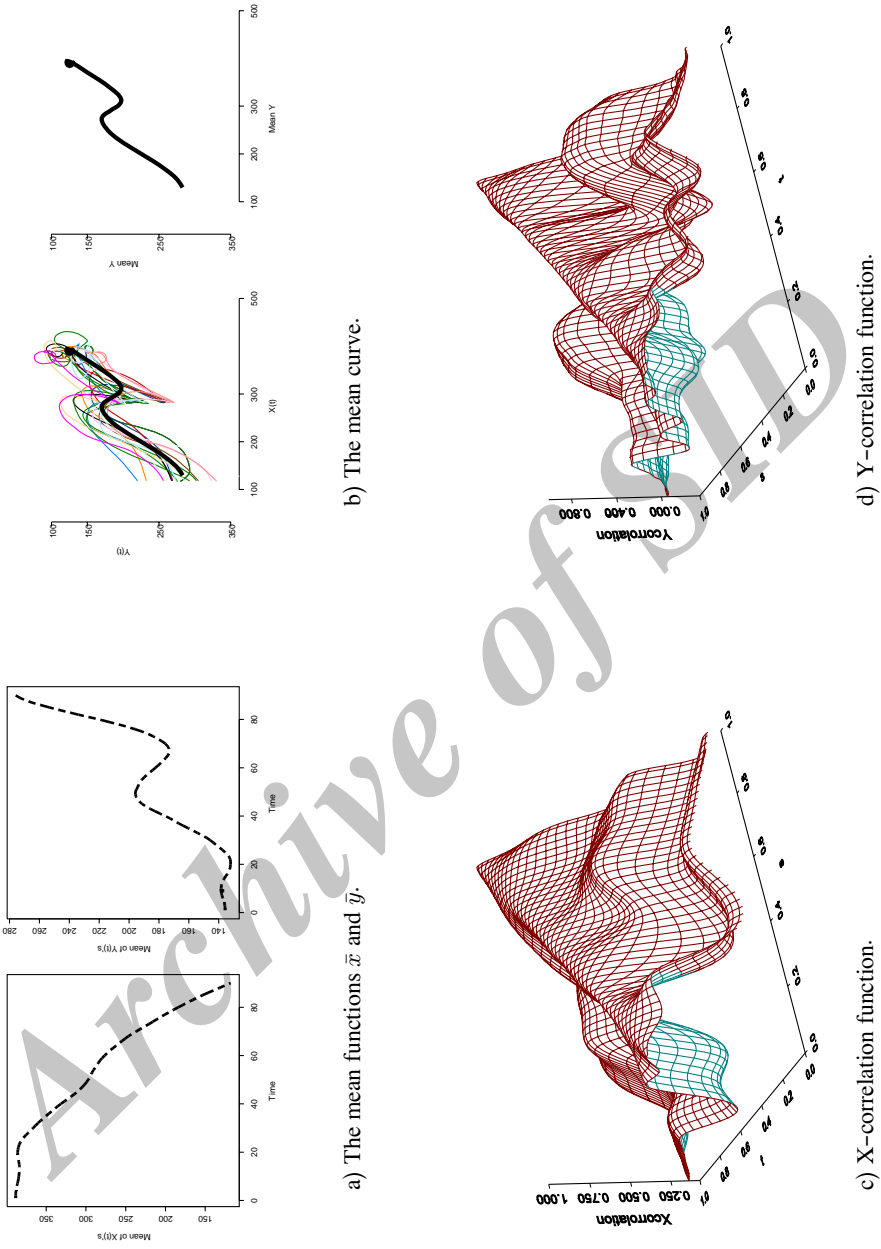
a) The mean functions $\bar{x}$ and $\bar{y}$.

b) The mean curve.

c) X-correlation function.

d) Y-correlation function.

Figure 7. Summary Statistics.

4.3 Principal Components

As mentioned before, principal components are computed using eigenfunctions which have some ideal properties. In this study the set of eigenfunctions consists of vectors (ξ_i^x, ξ_i^y) in which ξ_i^x and ξ_i^y are real functions and the couples (ξ_i^x, ξ_i^y) are ordered according to the magnitude of their corresponding eigenvalues $\{\lambda_i\}_{i=1}^{\infty}$. In practice we compute only a finite number of eigenfunctions by solving the matrix equation $V\xi = \lambda\xi$ where V is computed from discrete data. The sum on all eigenvalues gives the total variation so the proportion of each eigenvalue would be interpreted as the contribution of corresponding eigenfunction from the total variation. Proportions of the first four principal components in our data are:

$$\frac{\lambda_1}{\sum \lambda_i} = 0.496, \quad \frac{\lambda_2}{\sum \lambda_i} = 0.142, \quad \frac{\lambda_3}{\sum \lambda_i} = 0.105, \quad \frac{\lambda_4}{\sum \lambda_i} = 0.059. \quad (2)$$

That is the first eigenfunction, interpreted as the dominant mode of variation, accounts for 49.6% of the variability, first and second eigenfunctions simultaneously 63.8% of total variation and so forth.

On the other hand, since $\| \xi_i^x \|^2 + \| \xi_i^y \|^2 = 1$ the value $\| \xi_i^x \|^2$ (or $\| \xi_i^y \|^2$) gives the proportion of variability in the ith principal component accounted for by variation in the x (or y) functions. For example in the first and second eigenfunctions we have:

$$\| \xi_1^x \|^2 = 0.82, \qquad \| \xi_1^y \|^2 = 0.18,$$

$$\| \xi_2^x \|^2 = 0.27, \qquad \| \xi_2^y \|^2 = 0.73.$$

These proportions clarify that in the first eigenfunction, ξ_1^x plays a more important role (82% of total variability) and in the second one the ξ_2^x is less important. In other words, the first mode of variation is due to the differences between hand movements in horizontal direction and in second mode, it happens along vertical axes. These facts are displayed in Figure 8.

Curves of the eigenfunctions track down the occurrence duration of corresponding variability. Figure 8(a) shows that the first mode of variation, which is the most important one, is caused by the horizontal movements in the ending part of the writing period. This comes from the fact that in this part the absolute values goes far from the zero line. One can say that the biggest cause of variability in this data set is the difference between horizontal movements in writing the last letter.
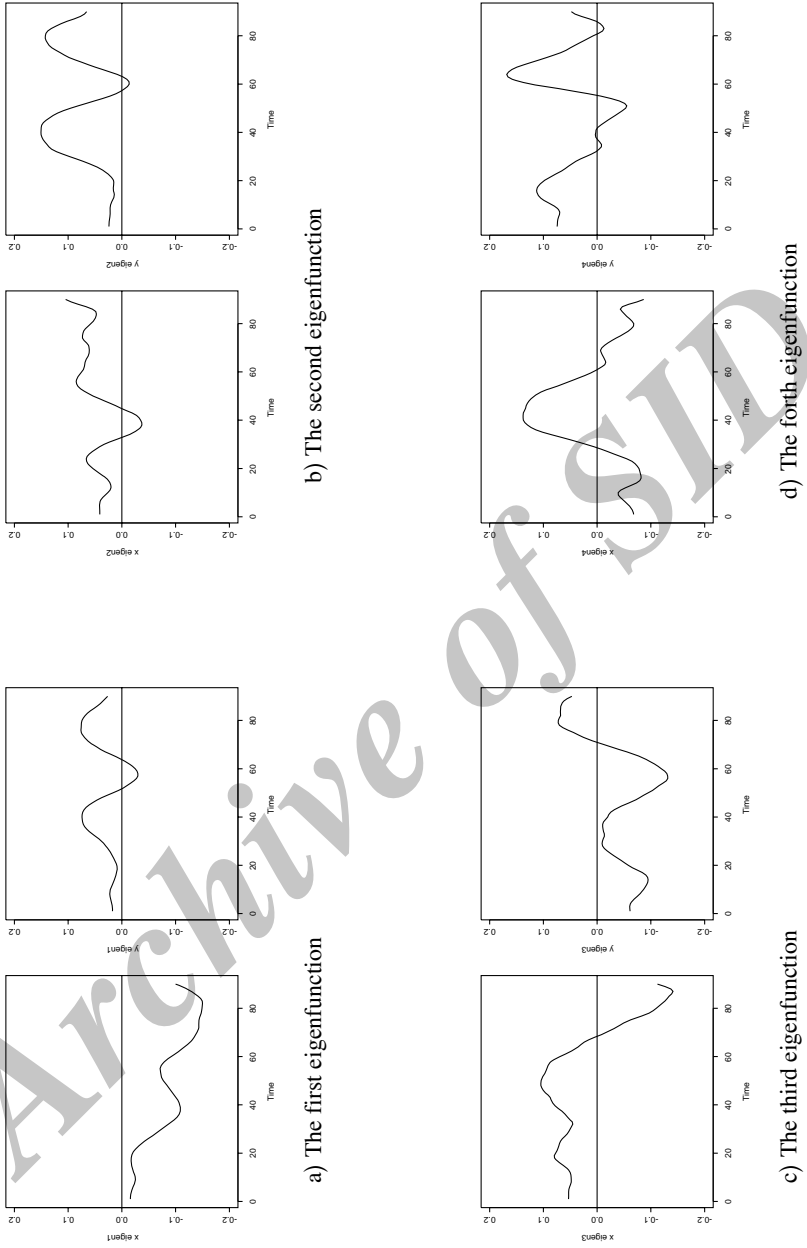
Figure 8. The first four eigenfunctions.

Second eigenfunction, shown in Figure 8(b), takes it's highest values on two periods; $[0.3, 0.5]$ and $[0.8, 0.9]$. These are the periods of starting to write the middle letter and connecting the middle and last letters respectively. So the second mode of variability is due to the vertical movements while writing the connecting parts of letters.

In the third eigenfunction, ξ_3^x and ξ_3^y play identical role and the higher values occur in the periods of writing the second parts of each letter. the last considered eigenfunction, the fourth one, accounts for a small proportion of the total variability since it is required to be orthogonal to the first three ones. This component resembles a time distortion during the whole period specially at the time of writing the ending parts of the word *Mehr*.

In conclusion, the dominant modes of variations are occured at the middle and ending parts of writing period, and the curves show almost similar structures at the starting part.

Up to now the variance partitioning made by the principal components has been studied, but not the nature or quality of these modes. Variation can be of two sorts. The first one is the temporal variation, caused by different timings in each letter. For interpreting this type of variation it is needed to plot the effects of eigenfunctions against time. The nature of second type of variation in a sense is some changes in the shape of each letter. This kind of effects can be illustrated by plotting x and y functions in the same plot, where the features related to time will disappear.

A good method to virtually illustrate both types of variations, offered by Ramsay and Silverman (1997), is plotting the effects of adding and subtracting the eigenfunctions to the mean function. For the first four types, these plots are shown in Figure 9. In this figure the solid line refers to the mean functions and dashed and dotted lines represent the functions obtained from adding and subtracting eigenfunctions to the mean, that is:

$$\bar{x}(t) \pm c_1 \xi_i^x, \qquad \bar{y}(t) \pm c_2 \xi_i^y \quad i : 1, \ldots, 4,$$

the constant coefficients c_1, c_2 are entered to clarify the plots and make them more interpretable. Here they are chosen to be $c_1 = 1.5 \times \|\bar{x}(t) - \bar{\bar{x}}\|$ and $c_2 = 1.5 \times \|\bar{y}(t) - \bar{\bar{y}}\|$, where $\bar{\bar{x}} = \int_0^1 \bar{x}(t)dt$ and $\bar{\bar{y}} = \int_0^1 \bar{y}(t)dt$.

Part (a) in Figure 9 shows the effects of first principal component. We saw that this type of variation is mainly caused by the x functions. Plots in Figure 9(a) reveals the effect of first principal component as a uniform increase or decrease on mean. It can be seen that the first mode of variation doesn't change mean pattern but just cause an overall rise or decline. We
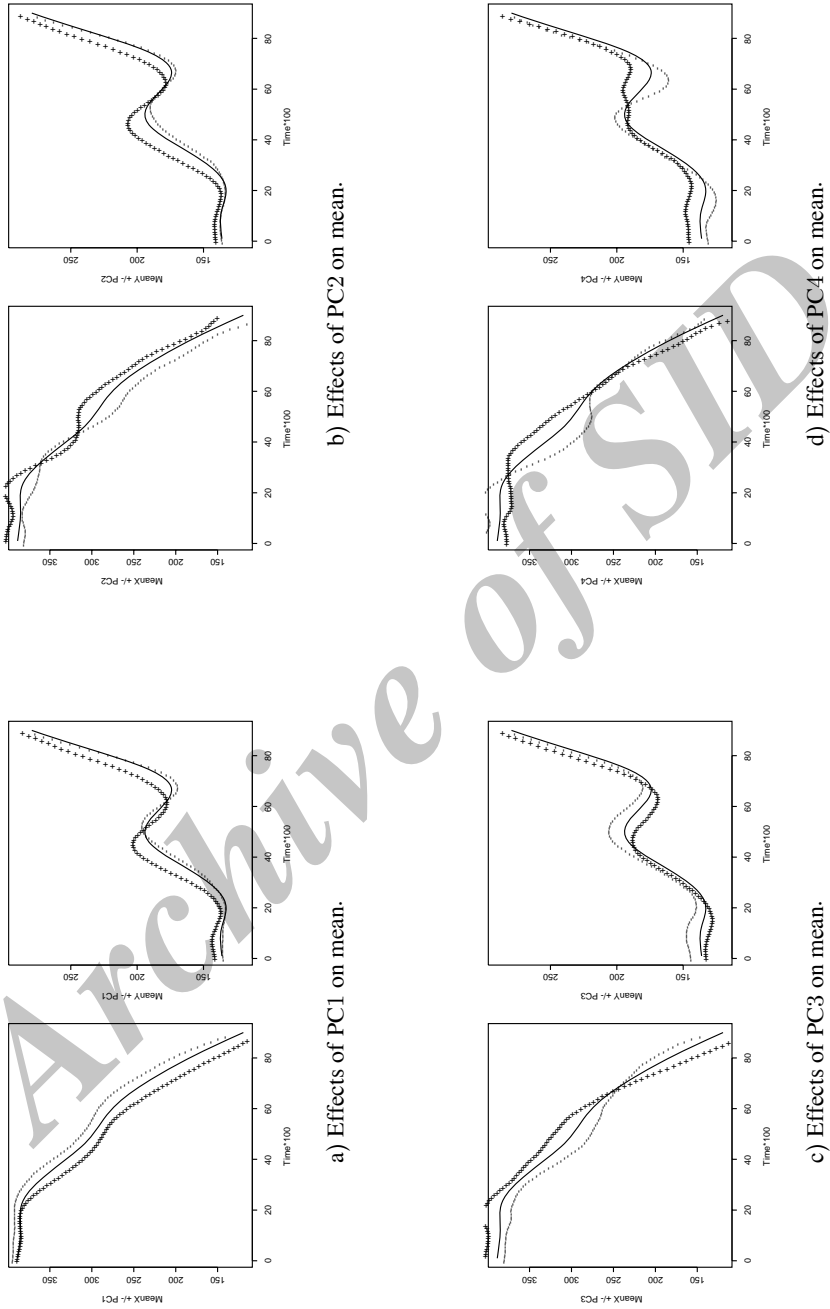
J. Statist. Res. Iran **6** (2009): 141–160

www.SID.ir

Figure 9. The effects of first four principal components.

can conclude that this type of variation is due to speed of hand movements along the x coordinate. These plots also illustrates that the effects of first principal component is little in the starting points and rockets at the rest of period. This feature comes from the fact that the differences in speed became more significant by the time passage.

Plots of Figure 9(b) indicate similar facts about the second principal component. Now the greater role is due to y functions. It is not surprising to find that the second principal component has the same effect as first one on vertical axis. Again the mean function changes at an approximately constant rate by adding or subtracting the eigenfunction. The only distinction is that this type creates its biggest effect at the middle of the period and this effect is weakened in the last parts. So the second type of variation is due to moving slower or faster than average along with the y axis. Note that this difference decreases in the last part of the period. The structure of both types (first and second) are simple increase or decrease during the period, but other principal components will offer more complex effects. Third eigenfunction affects through a combination of time shift and increase/decrease changes; illustrated in Figure 9(c). Note that the third mode of variation dues to both x and y functions, so it is to be analyzed simultaneously. Related x plot reveals a time shift in the second period, where in the first part it just contains the increase/decrease effect. Conversely, in y function, time shift happens in the first and middle part which is an interval around $t = 0.6$. This interval is related to the time of connecting the middle and last letters in the mean function. The fourth principal components, presented in the Figure 9(d), accounts for only 5% of variation. This effect can be interpreted as some fluctuations during the total period.

We have indicated two sorts of variations before. The second sort is the difference among final pattern of letters, caused by the modes of variation related to time. Figure 10 illustrates this type of effects. The best interpretation can be made by the plots themselves and in words we can just indicate some limited features. It is not astonishing to see that the first principal component has the effect of pushing the whole pattern along the x coordinate. In the upper-right plot of Figure 10 the same change is made in the vertical axes, but this time the direction of change is different in the starting and ending parts. Plot of third principal component emphasizes this effect. We have mentioned before that the third eigenfunction affects as a combination of time shift and position changes, but the effects which are directly related to time, can not be shown here. Forth principal component indicates some

distortions in the whole pattern. People with the high scores on this type of variation would probably have more irregular handwriting than others.

As a conclusion one could say that in the set of 32 university students greatest similarities is found in starting period of writing. At the middle period, they vary a lot in velocity. These differences mostly occur in writing the second part of the letters. Also people vary a lot in writing the last letter, specially in vertical direction. But at the almost ending instance, all functions shows a tendency to zero line. It means that despite of all variations even in writing the last letter, people tends to finish writing in the same manner and velocity.

5   Discussion

In this paper, principal components method was used to describe and analyze the online handwriting data. For this purpose PCA was applied in, $\mathbb{H}$, the linear space of vectors of real functions: $z = (x(t), y(t))$, where $\int_a^b x^2(t) dt$
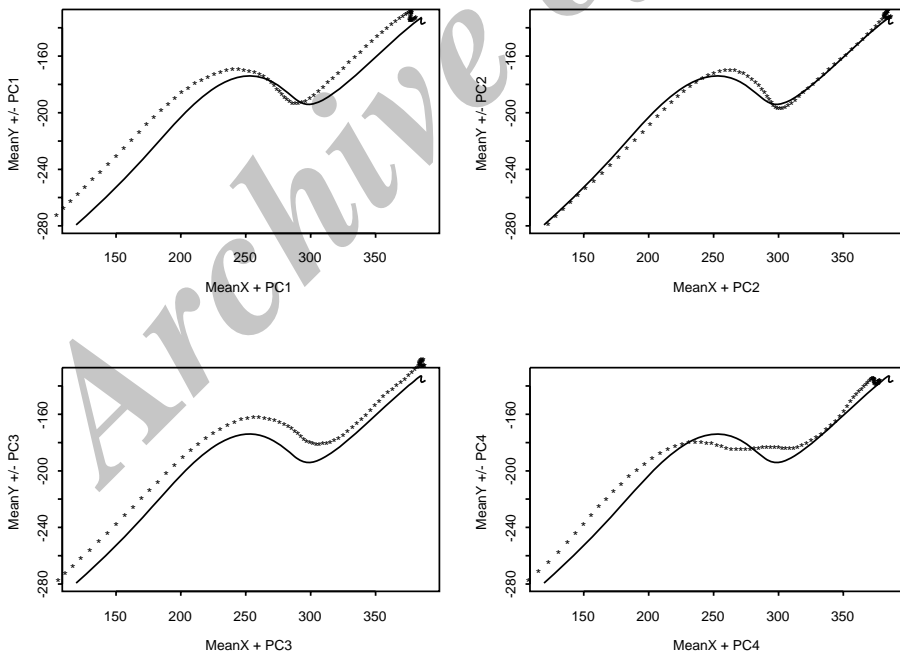

Figure 10. The effects of principal components on final pattern.

and $\int_a^b y^2(t)dt$ are finite. Inner product is the initial concept that principal components are built on. There are several ways to define an inner product in a bivariate context. The one that we used and makes $\mathbb{H}$ a Hilbert space is:

$$< z_1, z_2 >_1 = < x_1, x_2 > + < y_1, y_2 >,$$

where for two real valued functions f and g,

$$< f, g > = \int_a^b f(t)g(t)dt. \tag{3}$$

By this definition inner product always results in a scalar value and two vectors $z_1 = (x_1(t), y_1(t))$ and $z_2 = (x_2(t), y_2(t))$ are orthogonal if $< x_1, x_2 > = < y_1, y_2 > = 0$ or $< x_1, x_2 > = - < y_1, y_2 >$. This does not make sense to call two vectors orthogonal when the second situation happens. An alternative way to define an inner product on $\mathbb{H}$ is:

$$< z_1, z_2 >_2 = \begin{bmatrix} < x_1, y_1 > & < x_1, y_2 > \\ < x_2, y_1 > & < x_2, y_2 > \end{bmatrix}.$$

This inner product makes $\mathbb{H}$ a Hilbert C^*-module. In this space the principal components would have different meaning and characteristics that needs to be evaluated experimentally.

As the final words it should be mentioned that handwriting is a mechanical process which should be incorporated in the analysis. In other words the derivatives of functions and their relationships should be entered in the analysis. Finding relationships among derivatives is an alternative method to describe the process. For this purpose, differential equation techniques could be applied to replicated handwriting samples from a single individual (See Ramsey (2000) for a detailed example).

References

Dolfing, J.G.A., Aarts, E.H.L., van Osterhout, J.J.G.M. (1998). On-line signature verification with hidden Markov models. Proc. of IEEEE 14th International Conference on Pattern Recognition, Brisbane, Australia. **2**, 1309-1312.

Jain, K., Griess, F.D., Connell, S.D. (2002). On-line signature verification. *Pattern Recognition*, **35**, 2963-2972.

Haji, M.M. (2005). *Farsi Handwritten Word Recognition Using Continues Hidden Markov Models and Structural Features*, Shiraz University, Shiraz.

Kashi, R.S., Hu, J., Nelson, W.L., Turin, W. (1997). On-line handwriting signature verification using hidden Markov model features. Proc. of the 4th Intl. Conf. On Document Analysis and Recognition, **1**, 253-257.

McKeague, W. (2004). A Statistical Model for Signature Verification . *Journal of the American Statistical Association* (in press).

Nelson, W.L., Turin,W., Hastie, T. (1994). Statistical methods For on-line signature verification. *Int. Journ. Pattern Rec. and Mach. Intll.* **8**, 749-770.

Ramsay, J.O., Silverman, B.W. (1997). *Functional Data Analysis.* Springer, New York.

Ramsay, J.O. (2000). Functional components of variation in handwriting. *Journal of American Statistical Association*, **95**, 9-15.

Plamondo, R., Lorette, G. (1989). Automatic signature verification and writer identification-the state of art., *Pattern Recognition*, **22**, 107-131.

Plamondo, R., Srihari, S.N. (2000). On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Trans. Pattern Anal. and Machine Intell.*, **22**, 63-84.

Soldek, J., Shmerko, V., Phillips, P., Kukharev, G., Rogers, W., Yanushkevich, S. (1997). Image analysis and pattern recognition in biometric technologies. Proceedings International Conference On The Biometrics: Fraud Prevention, Enchanced Service, Las Vegas, Nevada, USA, 270-286.

Zuo, L., Wang, Y., Tan, T., (2002). Personal handwriting identification based on PCA. International Conference on Image and Graphics, Wei Sui; Ed. **4875**, 766-771.

Yasaman Hosseinkashi
200 Candlewood CR,
Waterloo, ON, Canada.
email: *yhossein@uwaterloo.ca*

Khalil Shafie
Applied Statistics and Research Methods,
University of Northern Colorado,
Colorado, U.S.A.
email: *Khalil.shafie@unco.edu*