Journal of Advances in Mathematics and Computer Science

23(6): 1-12, 2017; Article no.JAMCS.33436

Previously known as British Journal of Mathematics & Computer Science
ISSN: 2231-0851

Eigenvalue Problem with the Basis Exchange Algorithm

Leon Bobrowski[1,2*]

[1]Faculty of Computer Science, Białystok University of Technology, Poland.
[2]Institute of Biocybernetics and Biomedical Engineering, PAS, Warsaw, Poland.

Author's contribution

The sole author designed, analyzed and interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/JAMCS/2017/33436
Editor(s):
(1) Dariusz Jacek Jakóbczak, Chair of Computer Science and Management in this department, Technical University of Koszalin, Poland.
Reviewers:
(1) S. Graham Kelly, The University of Akron, USA.
(2) Anthony Aidoo, Eastern Connecticut State University, USA.
(3) Zhonggen wang, Anhui University, China.
(4) Okan Ozer, University of Gaziantep,Turkiye.
(5) Shenyang Tan, Taizhou institute of Sci.&Tec., National University of Sciences and Technology, Pakistan.
Complete Peer review History: http://www.sciencedomain.org/review-history/20424

Original Research Article

Received: 15th April 2017
Accepted: 3rd August 2017
Published: 8th August 2017

Abstract

The eigenvalue problem plays an important role in contemporary methods of exploratory data analysis. As an example, the *principal component analysis* (*PCA*) widely used in data exploration, is based on finding the eigenvalues and eigenvectors of the covariance matrix.

The paper presents a new method of the eigenvalue problem solution which uses the basis exchange algorithms. The basis exchange algorithms, similarly to the linear programming techniques are based on the Gauss-Jordan transformation of the inverted matrices. The proposed approach to the eigenvalue problem may also be connected to the regularization of feature vectors which constitute squared matrices by single unit vectors. The proposed approach is based on inducing a linear dependence among regularized vectors.

Keywords: Eigenvalue problem; data exploration; principal component analysis; basis exchange algorithms; linear dependency.

**Corresponding author: E-mail: l.bobrowski@pb.edu.pl;*

1 Introduction

Exploratory data analysis is typically aimed at discovering patterns (regularities) in large, multivariate data sests [1]. Patterns can be discovered by using variety of data mining tools [2]. Observed patterns may allow to extract useful knowledge from a given data set. Principal component analysis (*PCA*) is one of the fundamental, widely used method of data exploration [3]. The origins of the Principal Components Analysis (*PCA*) are linked to the works of Pearson [4], and Hotelling [5]. One of the best modern reference to *PCA* is provided by Jolliffe [6].

The principal component analysis (*PCA*) involves computation of the eigenvalues and eigenvectors of the covariance matrix created on the basis of the given data set [7]. Many computational techniques of the eigenvalue problems solutions and their applications in data exploration have been proposed and developed [8]. Apart from *PCA*, the eigenvalue solutions are used in Correspondence Analysis (*CA*), or Canonical Correlation Analysis (*CCA*). However, new computational methods are still developed which could allow to improve the efficiency of large, multivariate data exploration.

A new method of the eigenvalue problem solution is proposed in the paper. The presented method of the eigenvalue problem solution uses the matrices inversion techniques based on the basis exchange algorithm [9]. A crucial role in the proposed approach is played by induced linear dependency among regularized vectors [10]. The approach is linked to the regularization techniques of squared matrices by unit vectors, and it should be useful, among others, in collinear biclustering aimed on the flat pattern extraction [7,11].

2 Eignvalue Problem in Principal Component Analysis

Let us consider the data set $\boldsymbol{C}$ composed of m feature vectors $\mathbf{x}_j = [x_{j,1},...,x_{j,n}]^T$:

$$\boldsymbol{C} = \{\mathbf{x}_1,\ldots, \mathbf{x}_m\} \tag{1}$$

Feature vectors $\mathbf{x}_j$ can be considered as points in the n-dimensional feature space $F[n]$ ($\mathbf{x}_j \in F[n]$).

We assume that m objects O_j ($j = 1,\ldots, m$) are represented in a standardized manner by the n - dimensional feature vectors $\mathbf{x}_j$. Components $x_{j,i}$ of the feature vector $\mathbf{x}_j$ may represent numerical results of measurements of n different features x_i ($i = 1,\ldots,n$) of the j - th object O_j ($x_{j,i} \in \{0,1\}$ or $x_{j,i} \in R^1$).

The data set $\boldsymbol{C}$ (1) can be characterized by the mean vector $\mathbf{m}$ and the covariance matrix $\boldsymbol{S}$ [3]:

$$\mathbf{m} = \boldsymbol{\Sigma}_j \, \mathbf{x}_j \,/\, m \tag{2}$$

and

$$\boldsymbol{S} = \Sigma_j \, (\mathbf{x}_j - \mathbf{m})(\mathbf{x}_j - \mathbf{m})^T / (m - 1) \tag{3}$$

The *PCA* exploratory techniques are aimed at identifying unknown trends in multidimensional data sets. The *PCA* method is expected to reduce the dimensionality of multivariate data while preserving the variance as much as possible.

The basic idea of the *PCA* method is to describe the variation of a set of multivariate data in terms of a set of new uncorrelated variables, each of which is a particular linear combination of the n features x_i. A linear transformation of the feature vectors $\mathbf{x}_j$ (1) is applied that new variables are uncorrelated and have the greatest variability.

The *PCA* method is based on the solution of the eigenvalue problem with the symmetric covariance matrix $\boldsymbol{S}$ (3) of the dimension n x n [3]:

$$\boldsymbol{S}\ \mathbf{k}_i = \lambda_i\ \mathbf{k}_i \quad (4)$$

where $\mathbf{k}_i = [k_{i,1},...,k_{i,n}]^T$ is the *i*-th eigenvector ($i = 1,\ldots, n$) and λ_i is the *i*-th eigenvalue ($\lambda_i \geq 0$).

The eigenvectors $\mathbf{k}_i$ should have the unit lentgh:

$$(\forall i \in \{1,\ldots, n\}) \qquad \mathbf{k}_i^T\mathbf{k}_i = 1 \quad (5)$$

Two eigenvectors $\mathbf{k}_i$ and $\mathbf{k}_{i'}$ (4) of the symmetrical matrix $\boldsymbol{S}$ are perpendicular if their eigenvalues λ_i and $\lambda_{i'}$ are different ($\lambda_i \neq \lambda_{i'}$) [6]:

$$(\forall i \neq i') \quad \textit{if}\ \ \lambda_i \neq \lambda_{i'},\ \textit{then}\ \mathbf{k}_i^T\mathbf{k}_{i'} = 0 \quad (6)$$

The eigenvectors $\mathbf{k}_i$ are usually set according to decreasing eigenvalues λ_i:

$$\mathbf{k}_1,\ \mathbf{k}_2,\ \ldots,\ \mathbf{k}_n \quad (7)$$

where

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$$

The eigenvector $\mathbf{k}_i$ defines the *i*-th principal component y_i:

$$(\forall i \in \{1,\ldots, n\}) \quad y_i = \mathbf{k}_i^T\mathbf{x} = k_{i,1}x_1 + \ldots + k_{i,n}x_n \quad (8)$$

In accordance with the above transformation, each feature vector $\mathbf{x}_i$ can be projected on the point $y_{i,j}$ of the line y_i (8). It has been shown that the variance σ_i^2 of the vectors $\mathbf{x}_j$ projected on the line y_i (5) is equal to λ_i^2 [6]:

$$(\forall i \in \{1,\ldots, n\}) \quad \sigma_i^2 = \Sigma_j\ (y_{i,j} - \mu_i)^2 / (m - 1) = \lambda_i^2 \quad (9)$$

where μ_i is the mean value of the points $y_{i,j} = \mathbf{k}_i^T\mathbf{x}_j$ (5).

Based on the equation (7) we can see that the first principal components y_i (8) include the greatest parts of the vectors $\mathbf{x}_j$ variability. The covariance $Cov(y_i, y_{i'})$ of the new variables y_i and $y_{i'}$ (8) is equal to zero [6]:

$$(\forall i \neq i') \quad Cov(y_i, y_{i'}) = \Sigma_j\ (y_{i,j} - \mu_i)\ (y_{i',j} - \mu_{i'}) = 0 \quad (10)$$

It follows from the above equation that the new variables y_i and $y_{i'}$ (8) are uncorrelated.

The eigenvalues λ_i and the eigenvectors $\mathbf{k}_i$ are computed on the basis of the equation (4). The equation (4) can be presented in the below form [8]:

$$(\boldsymbol{S} - \lambda_i \boldsymbol{I})\ \mathbf{k}_i = \mathbf{0} \quad (11)$$

The determinant $|\boldsymbol{S} - \lambda_i \boldsymbol{I}|$ equal to zero is the condition necessary for a non-trivial solution $\mathbf{k}_i \neq \mathbf{0}$ of the equation (11):

$$|\ \boldsymbol{S} - \lambda_i \boldsymbol{I}\ | = 0 \quad (12)$$

The *characteristic equation* (12) is solved in order to find the matrix's $\boldsymbol{S}$ eigenvalues λ_i. Based on this, the eigenvectors $\mathbf{k}_i$ (7) can be found.

The Singular Value Decomposition (*SVD*) is currently the primary method of the eigenvalue problem solution for the Principal Components Analysis [8].

3 Eigenvectors Extracted from Inversed Matrix

Let us consider the non-singular matrix $\boldsymbol{S}$ composed of n rows $\mathbf{s}_j = [s_{j,1},\ldots, s_{j,n}]^T$ ($\mathbf{s}_j \in R^n$):

$$\boldsymbol{S} = [\mathbf{s}_1, \mathbf{s}_2,\ldots, \mathbf{s}_n]^T \qquad (13)$$

Referring to the eigenvalue equation (11) with the matrix $\boldsymbol{S}$ let us introduce the below *regularized matrix* $\boldsymbol{S}_\lambda$:

$$\boldsymbol{S}_\lambda = [\mathbf{s}_1 - \lambda\, \mathbf{e}_1, \mathbf{s}_2 - \lambda\, \mathbf{e}_2,\ldots, \mathbf{s}_n - \lambda\, \mathbf{e}_n]^T = [\mathbf{z}_1, \mathbf{z}_2,\ldots, \mathbf{z}_n]^T \qquad (14)$$

where λ is the *regularization parameter* ($\lambda \in R^1$), $\mathbf{e}_j$ is the j-th unit vector, and $\mathbf{z}_j$ is the *regularized vector*:

$$(\forall j \in \{1,\ldots, n\}) \quad \mathbf{z}_j = \mathbf{s}_j - \lambda\, \mathbf{e}_j \qquad (15)$$

The non-singular matrix $\boldsymbol{S}_\lambda$ (14) and the inverse matrix S_λ^{-1} can be created in an iterative manner from the unit matrix $\boldsymbol{I} = [\mathbf{e}_1,\ldots, \mathbf{e}_n]$ [9]. During the first step ($k = 1$) the matrix $\boldsymbol{S}_\lambda(1) = [\mathbf{z}_1, \mathbf{e}_2,\ldots, \mathbf{e}_n]^T$ can be created by changing the unit vector $\mathbf{e}_1$ to $\mathbf{z}_1$. Similarly, during step k, the matrix $\boldsymbol{S}_\lambda(k) = [\mathbf{z}_1,\ldots, \mathbf{z}_k, \mathbf{e}_{k+1},\ldots, \mathbf{e}_n]^T$ can be created by changing the vector $\mathbf{e}_k$ to $\mathbf{z}_k$ (15).

Let us assume that for some index i ($i \in \{1,\ldots, n\}$) the below non singular matrix $\boldsymbol{S}^i(n - 1)$ has been created in an iterative manner from the unit matrix $\boldsymbol{I} = [\mathbf{e}_1,\ldots, \mathbf{e}_n]$ during the first $n - 1$ steps of the matrix $\boldsymbol{S}_\lambda$ (14) inversion:

$$\boldsymbol{S}^i(n - 1) = [\mathbf{z}_1,\ldots, \mathbf{z}_{i-1}, \mathbf{e}_i, \mathbf{z}_{i+1}, \ldots, \mathbf{z}_n]^T \qquad (16)$$

where $(\forall j \in \{1,\ldots, n - 1\})$ $\mathbf{z}_j = \mathbf{s}_j - \lambda\, \mathbf{e}_j$ (15).

Only one unit vector $\mathbf{e}_i$ remains in the matrix $\boldsymbol{S}^i(n - 1)$ (16). The inverse matrix $\boldsymbol{S}^i(n - 1)^{-1}$ can be represented in the below manner [9]:

$$\boldsymbol{S}^i(n - 1)^{-1} = [\mathbf{r}_1^i(n - 1),\ldots, \mathbf{r}_i^i(n - 1),\ldots, \mathbf{r}_n^i(n - 1)] \qquad (17)$$

The regularized vectors $\mathbf{z}_j = \mathbf{s}_j - \lambda\, \mathbf{e}_j$ (15) and the i-th column $\mathbf{r}_i^i(n - 1)$ of the inverse matrix $\boldsymbol{S}^i(n - 1)^{-1}$ (17) fulfill the below equations:

$$(\forall i \in \{1,\ldots, n\}) \quad \mathbf{z}_i^T \mathbf{r}_i^i(n - 1) = (\mathbf{s}_i - \lambda\, \mathbf{e}_i)^T \mathbf{r}_i^i(n - 1) = 1 \qquad (18)$$

and

$$(\forall j \neq i) \quad \mathbf{z}_j^T \mathbf{r}_i^i(n - 1) = (\mathbf{s}_j - \lambda\, \mathbf{e}_j)^T \mathbf{r}_i^i(n - 1) = 0 \qquad (19)$$

The equations (18) and (19) result from the definition of the inverse matrix $\boldsymbol{S}^i(n - 1)^{-1}$ (17). These equations are similar to the eigenvalue equation (11) which can be represented in the below form:

$$(\forall j \in \{1,\ldots, n\}) \quad (\mathbf{s}_j - \lambda_i\, \mathbf{e}_j)^T \mathbf{k}_i = 0 \qquad (20)$$

where $\mathbf{s}_j$ is the j-th row of the matrix $\boldsymbol{S}$ (11), λ_i is the i-th eigenvalue and $\mathbf{k}_i$ is the i-th eigenvector of the matrix $\boldsymbol{S}$ (11) ($1 \leq i \leq n$).

Comparing the equations (20) with the equations (19) we can remark, that the unit length vector $\mathbf{r}_i^i(n - 1) / \| \mathbf{r}_i^i(n - 1)\|$ (17) would become the eigenvector of the matrix $\boldsymbol{S}$ (11) if the i-th equation (18) will be changed to:

$$(\mathbf{s}_i - \lambda\, \mathbf{e}_i)^T \mathbf{r}_i^i(n - 1) = 0 \tag{21}$$

The matrix $\boldsymbol{S}_\lambda$ (14) could be obtained from the matrix $\boldsymbol{S}^i(n - 1)$ (16) through the replacement of the *i*-th unit vector $\mathbf{e}_i$ by the regularized vector $\mathbf{s}_i - \lambda\, \mathbf{e}_i$. During such replacement the inverse matrix $\boldsymbol{S}^i(n - 1)^{-1}$ (17) should be transformed in the inverse matrix $\boldsymbol{S}_\lambda^{-1}$ (14). The Gauss-Jordan transformation can be used in a computation of the inverse matrices resulting from replacement of single vectors in the matrices $\boldsymbol{S}_\lambda(k) = [\mathbf{z}_1,\ldots, \mathbf{z}_k, \mathbf{e}_{k+1},\ldots, \mathbf{e}_n]^T$. The inverse matrices $S_\lambda(k)^{-1}$ can be computed efficiently in successive steps k by using the basis exchange algorithm based on the Gauss-Jordan transformation even in the case of high dimensional vectors [12].

The Gauss - Jordan transformation linked to the replacement of the the *i*-th unit vector $\mathbf{e}_i$ by the vector $\mathbf{s}_i - \lambda\, \mathbf{e}_i$ in the matrix $\boldsymbol{S}^i(n - 1)$ (16) can be given in the below manner [13]:

$$\mathbf{r}_i^i(n) = (1 / \mathbf{r}_i^i(n - 1)^T(\mathbf{s}_i - \lambda\, \mathbf{e}_i))\, \mathbf{r}_i^i(n - 1) \tag{22}$$

and $(\forall j \neq i)$

$$\begin{aligned}\mathbf{r}_j^i(n) &= \mathbf{r}_j^i(n - 1) - \mathbf{r}_j^i(n - 1)^T(\mathbf{s}_i - \lambda\, \mathbf{e}_i))\, \mathbf{r}_i^i(n) = \\ &= \mathbf{r}_j^i(n - 1) - (\mathbf{r}_j^i(n - 1)^T(\mathbf{s}_i - \lambda\, \mathbf{e}_i)) / \mathbf{r}_i^i(n - 1)^T(\mathbf{s}_i - \lambda\, \mathbf{e}_i))\, \mathbf{r}_i^i(n - 1)\end{aligned} \tag{23}$$

where $\mathbf{r}_j^i(n - 1)$ are the columns of the inverse matrix $\boldsymbol{S}^i(n - 1)^{-1}$ (17).

Remark 1: The condition (20) results in the division by zero in the equation (22).

The Gauss - Jordan transformation (22) cannot be used during the replacement of the unit vector $\mathbf{e}_i$ in the matrix $\boldsymbol{S}^i(n - 1)$ (16) by the vector $\mathbf{s}_i - \lambda\, \mathbf{e}_i$ if the condition (20) is met. The inverse matrix $\boldsymbol{S}_\lambda^{-1}$ (14) does not exists in this case. The condition (20) has also an interesting geometric interpretation as the move in the parameter space along the parallel hyperplane $h_j = \{\mathbf{w}: (\mathbf{s}_j - \lambda\, \mathbf{e}_j)^T\mathbf{w} = 1\}$ [12].

The condition (20) allows to compute the *prognosed values* λ_i^p of the parameter λ:

$$(\forall i \in \{1,\ldots, n\})\;\; \mathbf{r}_i^i(n - 1)^T(\mathbf{s}_i - \lambda_i^p\, \mathbf{e}_i) = 0 \tag{24}$$

thus

$$\lambda_i^p = \mathbf{r}_i^i(n - 1)^T\mathbf{s}_i \,/\, \mathbf{r}_i^i(n - 1)^T\mathbf{e}_i = \mathbf{r}_i^i(n - 1)^T\mathbf{s}_i \,/\, r_{i,i} \tag{25}$$

where $\mathbf{r}_i^i(n - 1)= [r_{i,1},\ldots, r_{i,n}]^T$ $(\mathbf{r}_i^i(n - 1)\in R^n)$.

Definition 1: The prognosed value λ_i^p (25) is *consistent* with the *i*-th eigenvalue λ_i if it is equal to the parameter λ used in the regularization $\mathbf{z}_j = \mathbf{s}_j - \lambda\, \mathbf{e}_j$ (15) of the vectors $\mathbf{s}_j$ (13).

$$\lambda_i^p = \lambda \tag{26}$$

Theorem 1: If the prognosed value λ_i^p (25) is equal to the regularized value λ (26), then the *i*-th eigenvalue λ_i of the matrix $\boldsymbol{S} = [\mathbf{s}_1,\ldots, \mathbf{s}_n]^T$ (13) is equal to λ_i^p

$$\lambda_i = \lambda_i^p \tag{27}$$

and is linked to the below eigenvector $\mathbf{k}_i$ (5):

$$\mathbf{k}_i = \mathbf{r}_i^i(n - 1) / \| \mathbf{r}_i^i(n - 1) \| \tag{28}$$

where $\mathbf{r}_i^i(n - 1)$ is *i*-th column of the inverse matrix $\boldsymbol{S}^i(n - 1)^{-1}$ (17).

The thesis of this theorem results directly from the construction described previously.

4 Itterative Fitting of Eigenvalues

Changing the vector $\mathbf{e}_i$ to $\mathbf{s}_i - \lambda_i\, \mathbf{e}_i$ in the non-singular matrix $\boldsymbol{S}^i(n - 1)$ (16) causes the matrix $\boldsymbol{S}_\lambda = [\mathbf{s}_1 - \lambda_i\, \mathbf{e}_1, \mathbf{s}_2 - \lambda_i\, \mathbf{e}_2, \ldots, \mathbf{s}_n - \lambda_i\, \mathbf{e}_n]^T$ (14) singularity if the condition (20) is fulfilled. The matrix $\boldsymbol{S}_\lambda$ obtained from the non singular matrix $\boldsymbol{S}^i(n - 1)$ (16) becomes singular if the *i*-th vector $\mathbf{s}_i - \lambda_i\, \mathbf{e}_i$ is a linear combination of the remaining vectors $\mathbf{s}_j - \lambda_i\, \mathbf{e}_j$ $(j \neq i)$ [12]:

$$(\forall i \in \{1, \ldots, n\}) \quad \mathbf{s}_i - \lambda_i\, \mathbf{e}_i = \alpha_{i,1}\, (\mathbf{s}_1 - \lambda_i\, \mathbf{e}_1) + \ldots + \alpha_{i,n}\, (\mathbf{s}_n - \lambda_i\, \mathbf{e}_n) \tag{29}$$

where $\mathbf{s}_i = [s_{i,1}, \ldots, s_{i,n}]^T$ $(\mathbf{s}_i \in R^n)$, $(\forall j \in \{1, \ldots, n\})$ $\alpha_{i,j} \in R^1$, and $\alpha_{i,i} = 0$.

Remark 2: The linear dependence (29) of the regularized vector $\mathbf{s}_i - \lambda_i\, \mathbf{e}_i$ on the remaining $n - 1$ vectors $\mathbf{s}_j - \lambda_i\, \mathbf{e}_j$ $(j \neq i)$ results in the appearance of the condition (20).

The problem of fitting the prognosed parameters λ_i^p (25) on the basis of the consistency condition (26) is now analysed. Let us first consider fitting the regularization parameter λ (15) in an iterative procedure.

The proposed iterative procedure is based on the *Theorem* 1. The procedure starts with inversion of the matrix $\boldsymbol{S}_\lambda$ (14) composed of the n regularized vectors $\mathbf{z}_i = \mathbf{s}_i - \lambda_0\, \mathbf{e}_i$ (15), where λ_0 is a large value of the paramter λ $(\lambda_0 \in R^1)$. A large value λ_0 of the parameter λ is expected to give the non-singular matrix $\boldsymbol{S}_\lambda$ (14):

$$\boldsymbol{S}_\lambda(n) = [\mathbf{s}_1 - \lambda_0\, \mathbf{e}_1, \ldots, \mathbf{s}_i - \lambda_0\, \mathbf{e}_i, \ldots, \mathbf{s}_n - \lambda_0\, \mathbf{e}_n]^T \tag{30}$$

The inverse matrix $\boldsymbol{S}_\lambda(n)^{-1}$ is represented below as:

$$\boldsymbol{S}_\lambda(n)^{-1} = [\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_n] \tag{31}$$

The rows $(\mathbf{s}_i - \lambda_0\, \mathbf{e}_i)^T$ (30) and the columns $\mathbf{r}_i$ of the matrix $\boldsymbol{S}_\lambda^{-1}(n)$ (31) fulfill the inverse equations (18) and (19).

As a result of the replacements of the n unit vectors $\mathbf{e}_i$ by the regularized vectors $\mathbf{z}_i = \mathbf{s}_i - \lambda_0\, \mathbf{e}_i$ (15) the complete inverse matrix $\boldsymbol{S}_\lambda(n)^{-1}$ (31) can be obtained. A temporary replacement of the vector $\mathbf{z}_i = \mathbf{s}_i - \lambda_0\, \mathbf{e}_i$ by the unit vector $\mathbf{e}_i$ in the matrix $\boldsymbol{S}_\lambda(n)$ (31) allows to compute efficiently the *i*-th column $\mathbf{r}_i^i(n - 1)$ of the matrix $\boldsymbol{S}_\lambda^i(n - 1)^{-1}$ (17) in accordance with the Gauss - Jordan transformation (22):

$$\mathbf{r}_i^i(n - 1) = (1 / \mathbf{r}_i^T \mathbf{e}_i)\, \mathbf{r}_i = (1 / r_{ii})\, \mathbf{r}_i \tag{32}$$

where $\mathbf{r}_i = [r_{i,1}, \ldots, r_{i,n}]^T$ is the *i*-th column $\mathbf{r}_i$ of the matrix $\boldsymbol{S}_\lambda(n)^{-1}$ (31).

The *i*-th column $\mathbf{r}_i^i(n - 1)$ of the matrix $\boldsymbol{S}_\lambda^i(n - 1)^{-1}$ (31) allows to compute the prognosis λ_i^p (25) for each eigenvalue λ_i:

$$(\forall i \in \{1, \ldots, n\}) \quad \lambda_i^p = \mathbf{r}_i^i(n - 1)^T \mathbf{s}_i \,/\, \mathbf{r}_i^i(n - 1)^T \mathbf{e}_i = \mathbf{r}_i^T \mathbf{s}_i \,/\, r_{i,i} \tag{33}$$

The proposed iterative procedure is based on the comparisons of the prognosed value λ_i^p (33) with the actual value λ_0 used in the definition of the regularized vectors $\mathbf{z}_i = \mathbf{s}_i - \lambda_0\, \mathbf{e}_i$ (15). The below rules based on comparison of the difference $|\lambda_i^p - \lambda_0|$ with a small parameter ε $(\varepsilon \geq 0)$ are proposed:

$$\textbf{\textit{if}}\ \ |\lambda_i^p - \lambda_0| > \varepsilon, \ \ \textbf{\textit{then}}\ \lambda_0 = \lambda_i^p, \ \ (\forall i \in \{1, \ldots, n\})\ \mathbf{z}_i = \mathbf{s}_i - \lambda_0\, \mathbf{e}_i, \ \textbf{\textit{and}} \tag{34}$$

the new matrix $\boldsymbol{S}_\lambda^{i}(n - 1)^{-1}$ (31) is designed from the vectors $\mathbf{z}_i$

$$\textbf{\textit{if}}\ |\lambda_i^p - \lambda_0| \leq \varepsilon,\ \textbf{\textit{then}}\ \lambda_i = \lambda_i^p,\ \textbf{\textit{and}}\ \text{the procedure is stopped.} \tag{35}$$

In accordance with above rules the iterative procedure is stopped if the difference $|\lambda_i^p - \lambda_0|$ becomes small. In this case the last prognosed value λ_i^p (25) can play the role of the *i*-th eigenvalue λ_i.

5 Fitting Eigenvalues through Inducing Linear Dependency

Alternative procedure of the eigenvalues λ_i computation is considered here. The proposed procedure is not iterative and is linked directly to the induced linear dependency (29) between the regularized vectors $\mathbf{s}_i - \lambda_i\mathbf{e}_i$ (15). The equation (29) describes the linear dependency of the *i*-th vector $\mathbf{s}_i - \lambda_i\ \mathbf{e}_i$ on the remaining $n - 1$ vectors $\mathbf{s}_i - \lambda_i\ \mathbf{e}_i$ ($j \neq i$). This equation can be formulated for each of the n vectors $\mathbf{s}_i = \mathbf{s}_i - \lambda_i\ \mathbf{e}_i$ ($i = 1,\ldots, n$) in the below manner:

$$\begin{array}{l}(\forall i \in \{1,\ldots, n\}) \\ \mathbf{s}_i - \alpha_{i,1}\ \mathbf{s}_1 - \ldots - \alpha_{i,n}\ \mathbf{s}_n = \lambda_i\ (\mathbf{e}_i - \alpha_{i,1}\ \mathbf{e}_1 - \ldots - \alpha_{i,n}\ \mathbf{e}_n)\end{array} \tag{36}$$

where $\alpha_{i,i} = 0$.

Multiplying both sides of the equality (36) for particular values of the index i by the unit vectors $\mathbf{e}_j$ ($j = 1,\ldots, n$) we obtain the below set of equations:

$$\begin{array}{l}(\forall i \in \{1,\ldots, n\}) \\ \quad s_{i,1} - \alpha_{i,1}s_{1,1} - \ldots - \alpha_{i,n}\ s_{n,1} = -\lambda_i\ \alpha_{i,1} \\ \quad s_{i,2} - \alpha_{i,1}s_{1,2} - \ldots - \alpha_{i,n}\ s_{n,2} = -\lambda_i\ \alpha_{i,2} \\ \ldots. \\ \ldots. \\ \quad s_{i,i} - \alpha_{i,1}s_{1,i} - \ldots - \alpha_{i,n}\ s_{n,i} = \lambda_i \\ \ldots. \\ \ldots. \\ \quad s_{i,n-1} - \alpha_{i,1}s_{1,n-1} - \ldots - \alpha_{i,n}\ s_{n,n-1} = -\lambda_i\ \alpha_{i,n-1} \\ \quad s_{i,n} - \alpha_{i,1}s_{1,n} - \ldots - \alpha_{i,n}\ s_{n,n} = -\lambda_i\ \alpha_{i,n}\end{array} \tag{37}$$

The *i*-th equation in the set (37) allows to determine the parameter λ_i in the below manner:

$$\begin{array}{l}(\forall i \in \{1,\ldots, n\}) \\ \quad \lambda_i = s_{i,i} - \alpha_{i,1}s_{1,i} - \ldots - \alpha_{i,n}\ s_{n,i}\end{array} \tag{38}$$

The sets of equations (37) can be represented without the parameter λ_i in the below manner:

$$\begin{array}{l}(\forall i \in \{1,\ldots, n\}) \\ \quad s_{i,1} - \alpha_{i,1}s_{1,1} - \ldots - \alpha_{i,n}\ s_{n,1} = -(s_{i,i} - \alpha_{i,1}s_{1,i} - \ldots - \alpha_{i,n}\ s_{n,i})\ \alpha_{i,1} \\ \quad s_{i,2} - \alpha_{i,1}s_{1,2} - \ldots - \alpha_{i,n}\ s_{n,2} = -(s_{i,i} - \alpha_{i,1}s_{1,i} - \ldots - \alpha_{i,n}\ s_{n,i})\ \alpha_{i,2} \\ \ldots. \\ \ldots. \\ \ldots. \\ \quad s_{i,n} - \alpha_{i,1}s_{1,n} - \ldots - \alpha_{i,n}\ s_{n,n} = -(s_{i,i} - \alpha_{i,1}s_{1,i} - \ldots - \alpha_{i,n}\ s_{n,i})\ \alpha_{i,n}\end{array} \tag{39}$$

The *i*-th set of $n - 1$ equations (39) contains $n - 1$unknown variables $\alpha_{i,j}$ for each value of the index *i*.The set of parameters contains $\alpha_{i,1},\ldots,$ and $\alpha_{i,n}$ unknown variables without the coefficient $\alpha_{i,i}$ which is equal to zero ($\alpha_{i,i} = 0$). The equations (39) contain both linear as well as quadratic variables $\alpha_{i,i}$.

Theorem 2: If the matrix $\boldsymbol{S}^{i}(n - 1)$ (16) is not singular, then a solution $\lambda_{i'}$ (38) of the *i*-th system of *n* equations (37) is equal to the eigenvalue λ_i (4) of the matrix $\boldsymbol{S} = [\mathbf{s}_1,\dots, \mathbf{s}_n]^T$ (11):

$$\lambda_i = \lambda_{i'} \tag{40}$$

The eigenvalue λ_i (40) can be linked to the below eigenvector $\mathbf{k}_i$ (28):

$$\mathbf{k}_i = \mathbf{r}_i^{i}(n - 1) \,/\, \| \mathbf{r}_i^{i}(n - 1) \| \tag{41}$$

where $\mathbf{r}_i^{i}(n - 1)$ is the *i*-th column of the inverse matrix $\boldsymbol{S}^{i}(n - 1)^{-1}$ (16).

Proof: If the parameter λ_i' is the solution of the *i*-th system of *n* equations (37), then the *i*-th vector $\mathbf{z}_i(i) = \mathbf{s}_i - \lambda_i' \mathbf{e}_i$ is the linear combination (29) of the remaining $n - 1$ vectors $\mathbf{z}_j(i) = \mathbf{s}_j - \lambda_i' \mathbf{e}_j$ $(j \neq i)$.

If the matrix $\boldsymbol{S}^{i}(n - 1)$ (16) is not singular, then the inverse matrix $\boldsymbol{S}^{i}(n - 1)^{-1}$ (17) can be computed by using the iterative basis exchange procedure based on the Gauss - Jordan transformation [12]. The *i*-th column $\mathbf{r}_i^{i}(n - 1)$ of the inverse matrix $\boldsymbol{S}^{i}(n - 1)^{-1}$ (17) fulfils the equations (19):

$$(\forall j \in \{1,\dots, n: j \neq i\}) \;\; \mathbf{r}_i^{i}(n - 1)^T(\mathbf{s}_j - \lambda_i' \mathbf{e}_j) = 0 \tag{42}$$

The solution λ_i' of the *i*-th system (37) allows to fulfill the equation (42) also by the *i*-th vector $\mathbf{z}_i(i) = \mathbf{s}_i - \lambda_i' \mathbf{e}_i$:

$$\mathbf{r}_i^{i}(n - 1)^T\mathbf{z}_i(i) = \mathbf{r}_i^{i}(n - 1)^T(\mathbf{s}_i - \lambda_i' \mathbf{e}_i) = 0 \tag{43}$$

The above equality results from the linear dependency (29) of the *i*-th vector $\mathbf{z}_i(i) = \mathbf{s}_i - \lambda_i' \mathbf{e}_i$ on the remaining $n - 1$ vectors $\mathbf{z}_j(i) = \mathbf{s}_j - \lambda_i' \mathbf{e}_j$ $(j \neq i)$.

Taking into account the equations (42) and (43) we realize that the parameter λ_i' obtained from the equation (37) is the eigenvalue λ_i of the matrix $\boldsymbol{S} = [\mathbf{s}_1,\dots, \mathbf{s}_n]^T$ (11) with the eigenvector $\mathbf{k}_i$ given by (41). □

6 Examples of Eigenvalues Calculations

Examples of computation of eigenvalues of two and tree dimensional matrices by using the induced linear dependency () are provided in this section. These simple examples should help to work out the intuition behind the calculation technique proposed in this paper.

Example 1: Let us consider the below symmetric matrix $\boldsymbol{A}_1$ and the regularized matrix $\boldsymbol{A}_1'$ (14):

$$\boldsymbol{A}_1 = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix} \quad \boldsymbol{A}_1' = \begin{bmatrix} 6 - \lambda & 2 \\ 2 & 3 - \lambda \end{bmatrix} \tag{44}$$

The induced linear dependency (29) of the columns of the regularized matrix A_1' leads to the below equations with the parameters α $(\alpha \in R^1)$ and λ $(\lambda \in R^1)$:

$$2 = \alpha\,(6 - \lambda) \;\; \textit{and} \;\; (3 - \lambda) = 2\,\alpha \tag{45}$$

thus

$$\lambda = 3 - 2\,\alpha \;\textit{and}\; 2\alpha^2 + 3\alpha - 2 = 0 \tag{46}$$

Two eigenvalues λ_1 and λ_2 of the matrix A_1 are obtained from the above equations:

$$\lambda_1 = 7 \ \ for \ \ \alpha_1 = -2 \ \ and \ \lambda_2 = 2 \ \ for \ \ \alpha_2 = 2\ \alpha_2 \tag{47}$$

Example 2: The non-symmetric matrix A_2 is considered:

$$A_2 = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} \quad A_2' = \begin{bmatrix} 2-\lambda & 1 \\ 2 & 3-\lambda \end{bmatrix} \tag{48}$$

The equations of the induced linear dependency (29) have now the below form:

$$1 = \alpha\,(2 - \lambda) \ \ and \ \ (3 - \lambda) = 2\,\alpha \tag{49}$$

Two eigenvalues λ_1 and λ_2 of the matrix A_2 are obtained from these equations:

$$\lambda_1 = 4 \ \ for \ \ \alpha_1 = -1/2 \ \ and \ \lambda_2 = 1 \ \ for \ \ \alpha_2 = 1 \tag{50}$$

Example 3: The below matrix A_3 of the dimension $3 * 3$ is considered:

$$A_3 = \begin{bmatrix} 2 & 2 & -1 \\ -5 & 9 & -3 \\ -4 & 4 & 1 \end{bmatrix} \quad A_3' = \begin{bmatrix} 2-\lambda & 2 & -1 \\ -5 & 9-\lambda & -3 \\ -4 & 4 & 1-\lambda \end{bmatrix} \tag{51}$$

The induced linear dependency (29) of the first column of the matrix A_3' from the other two columns gives the below equations:

$$\begin{aligned} 2 - \lambda &= 2\,\alpha_2 - \alpha_3 \\ -5 &= (9 - \lambda)\,\alpha_2 - 3\,\alpha_3 \\ -4 &= 4\,\alpha_2 + (1 - \lambda)\,\alpha \end{aligned} \tag{52}$$

This system of equations has the solution with two eigenvalues:

$$\begin{aligned} &\lambda_1 = 4 \ \ for \ \ \alpha_{2,1} = -1, \ and \ \alpha_{3,1} = 0 \ \ or \\ &\lambda_2 = 5 \ \ for \ \ \alpha_{2,2} = -2, \ and \ \ \alpha_{3,2} = 0 \end{aligned} \tag{53}$$

The induced linear dependency (29) of the second column of the matrix A_3' from the other two columns results in the equations:

$$\begin{aligned} 2 &= (2 - \lambda)\,\alpha_1 - \alpha_3 \\ 9 - \lambda &= -5\alpha_1 - 3\,\alpha_3 \\ 4 &= -4\,\alpha_1 + (1 - \lambda)\,\alpha_3 \end{aligned} \tag{54}$$

Three eigenvalues λ_1, λ_2 and λ_3 of the matrix A_3 can be obtained from the above equations:

$$\begin{aligned} &\lambda_1 = 3 \ \ for \ \ \alpha_{1,1} = -2, \ and \ \alpha_{3,1} = 0 \ \ or \\ &\lambda_2 = 4 \ \ for \ \ \alpha_{1,2} = 0, \ and \ \ \alpha_{3,2} = 1 \\ &\lambda_3 = 5 \ \ for \ \ \alpha_{1,3} = -½, \ \ and \ \ \alpha_{3,2} = ½ \end{aligned} \tag{55}$$

7 Repeated Eigenvalues

Let us now consider the below non-singular matrices $\boldsymbol{S}_\lambda^{i}(k)$ created in the successive k steps ($1 \le k \le n - 1$) from the unit matrix $\boldsymbol{I} = [\mathbf{e}_1,\ldots, \mathbf{e}_n]$ through exchange the unit vectors $\mathbf{e}_k$ by the regularized vectors $\mathbf{z}_k(i) = \mathbf{s}_k - \lambda_i'\,\mathbf{e}_k$ (15) [9].

$$\mathbf{S}_\lambda^{i}(k) = [\mathbf{z}_1(i), \ldots, \mathbf{z}_k(i), \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n]^T \tag{56}$$

The k-th matrix (*basis*) $\mathbf{S}_\lambda^{i}(k)$ contains k regularized basic vectors $\mathbf{z}_j(i) = \mathbf{s}_j - \lambda_i' \, \mathbf{e}_j$ ($j \in J_k$), and $n - k$ unit vectors $\mathbf{e}_i$. It is assumed here that the regularized vectors $\mathbf{z}_k(i) = \mathbf{s}_k - \lambda_i' \, \mathbf{e}_k$ (15) are defined by the solution λ_i' of the i-th system of equations (37).

It can be seen, that the k'-th unit vector $\mathbf{e}_{k'}$ ($k' > k$) in the non-singular matrix $\mathbf{S}_\lambda^{i}(k)$ (56) should not be replaced by the vector $\mathbf{z}_{k'}(i) = \mathbf{s}_{k'} - \lambda_i' \, \mathbf{e}_{k'}$ (15) if the below condition similar to (24) is met:

$$\mathbf{r}_{k'}^{i}(k)^T(\mathbf{s}_{k'} - \lambda_i' \, \mathbf{e}_{k'}) = 0 \tag{57}$$

where $\mathbf{r}_{k'}^{i}(k)$ is the k'-th column of the inverse matrix $\mathbf{S}_\lambda^{i}(k)^{-1}$ (56):

$$S_\lambda^{i}(k)^{-1} = [\mathbf{r}_1^{i}(k), \ldots, \mathbf{r}_{k'}^{i}(k), \ldots, \mathbf{r}_n^{i}(k)] \tag{58}$$

Lemma 1: The regularized vector $\mathbf{z}_{k'}(i) = \mathbf{s}_{k'} - \lambda_i' \, \mathbf{e}_{k'}$ ($k' \neq i$) can-not be inserted into the matrix (basis) $\mathbf{S}_\lambda^{i}(k)$ (56) while preserving this matrix non-singularity, if the vector $\mathbf{z}_{k'}(i)$ is a linear combination (29) of the k basic vectors $\mathbf{z}_j(k) = \mathbf{s}_j - \lambda_i' \, \mathbf{e}_j$ ($j = 1, \ldots, k$):

$$\mathbf{s}_{k'} - \lambda_i' \, \mathbf{e}_{k'} = \alpha_{k',1} (\mathbf{s}_1 - \lambda_i' \, \mathbf{e}_1) + \ldots + \alpha_{k',k} (\mathbf{s}_k - \lambda_i' \, \mathbf{e}_k) \tag{59}$$

Proof: In accordance with the Gauss-Jordan transformation (22) the regularized vector $\mathbf{z}_{k'}(i) = \mathbf{s}_{k'} - \lambda_i' \, \mathbf{e}_{k'}$ (15) can not be inserted into the matrix $\mathbf{S}_\lambda^{i}(k)$ (56) if the condition (57) occurs. The Gauss Jordan transformation (22) cannot be used with the condition (56) because the division by zero appears.

The columns $\mathbf{r}_{j'}^{i}(k)$ ($j' > k$) of the inverted matrix $\mathbf{S}_\lambda^{i}(k)^{-1}$ (58) and the basic vectors $\mathbf{z}_j(k) = \mathbf{s}_j - \lambda_i' \, \mathbf{e}_j$ ($j \in J_k$) in the matrix $\mathbf{S}_\lambda^{i}(k)$ (56) fulfill the equations (19):

$$(\forall j \in J_k) \; (j' > k) \;\; \mathbf{r}_{j'}^{i}(k)^T(\mathbf{s}_j - \lambda_i' \, \mathbf{e}_j) = 0 \tag{60}$$

Therefore, if the vector $\mathbf{z}_{k'}(i) = \mathbf{s}_{k'} - \lambda_i' \mathbf{e}_{k'}$ ($k' > k$) is a linear combination (59) of the basic vectors $\mathbf{z}_j(k) = \mathbf{s}_j - \lambda_i' \, \mathbf{e}_j$ ($j \in J_k$) then the condition (60) appears. We can also infer that the linear dependency (59) is necessary for the condition (60) appearing [12]. □

Theorem 3: If the k'-th regularized vector $\mathbf{z}_{k'}(i) = \mathbf{s}_{k'} - \lambda_i' \, \mathbf{e}_{k'}$ ($k' > k$) defined by the solution λ_i' of the i-th system (37) fulfills the condition (60) then then the k'-th eigenvalue $\lambda_{k'}$ of the matrix $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_n]^T$ (13) is equal to λ_i':

$$\lambda_{k'} = \lambda_i' \tag{61}$$

and the k'-th eigenvector $\mathbf{k}_{k'}$ of the matrix $\mathbf{S}$ can be determined as

$$\mathbf{k}_{k'} = \mathbf{r}_{k'}^{i}(k) \, / \, \| \, \mathbf{r}_{k'}^{i}(k) \, \| \tag{62}$$

where $\mathbf{r}_{k'}^{i}(k)$ is the k'-th column of the inverse matrix $\mathbf{S}_\lambda^{k}(k)^{-1}$ (58) which is linked to the k'-th unit vector $\mathbf{e}_{k'}$ ($k' > k$) in the basis $\mathbf{S}_\lambda^{i}(k)$ (56).

Remark 4: Each regularized vector $\mathbf{z}_{k'}(i) = \mathbf{s}_{k'} - \lambda_i' \, \mathbf{e}_{k'}$ ($k' \notin J_k$) ($k' > k$) which is a linear combination (59) of the k basic vectors $\mathbf{z}_i(k) = \mathbf{s}_i - \lambda_i' \, \mathbf{e}_i$ ($j \in J_k$) allows to determine the k'-th eigenvalue $\lambda_{k'}$ (61) and the eigenvector $\mathbf{k}_{k'}$ (62) of the matrix $\mathbf{S}$ (13).

If none of n - k regularized vector $\mathbf{z}_i(i) = \mathbf{s}_i - \lambda_i' \mathbf{e}_i$ $(j > k)$ can be inserted into the basis $\boldsymbol{S}_\lambda^{i}(k)$ (56), then the repeated eigenvalues $\lambda_{k'} = \lambda_j'$ (61) are linked to n - k eigenvectors $\mathbf{k}_{k'}$ (61) of the matrix $\boldsymbol{S}$ (13).

8 Concluding Remarks

The eigenvalue problem (4) has been decomposed in the proposed method into subproblems linked to single regularized vectors $\mathbf{s}_i - \lambda\, \mathbf{e}_i$ $(i = 1,\ldots, n)$ (15) The equation (29) describes the induced linear dependency of the i-th regularized vector $\mathbf{s}_i - \lambda\, \mathbf{e}_i$ $(i = 1,\ldots, n)$ on the remaining n - 1 vectors $\mathbf{s}_j - \lambda\, \mathbf{e}_j$ $(j \neq i)$. The equation of linear dependency (29) of the i-th regularized vector $\mathbf{s}_i - \lambda\, \mathbf{e}_i$ (14) allows to find the i-th eigenvalue λ_i (38) through solving the set of quadratic equations (39). In accordance with *Theorem* 2 the i-th eigenvalue λ_i (40) allows to determine the eigenvector $\mathbf{k}_i$ (41) on the basis of the i-th column $\mathbf{r}_i^{i}(n - 1)$ of the inverse matrix $\boldsymbol{S}^{i}(n - 1)^{-1}$ (17).

The induced linear dependency (29) between the regularized vectors $\mathbf{s}_i - \lambda\, \mathbf{e}_i$ (15) plays a crucial role in the proposed solution of the eigenvalue problem. The considered approach to the eigenvalue problem can be linked to the regularization techniques of squared matrices by single unit vectors [7]. The presented approach should be useful, among others, in enlarging possibilties of collinear biclustering aimed at flat patterns extraction from large, high dimensional data sets [11].

Acknowledgements

The presented study was supported by the grant S/WI/2/2013 from Bialystok University of Technology and funded from the resources for research by Polish Ministry of Science and Higher Education.

Competing Interests

Author has declared that no competing interests exist.

References

[1] Duda OR, Hart PE, Stork DG. Pattern classification. J. Wiley, New York; 2001.

[2] Hand D, Smyth P, Mannila H. Principles of data mining. MIT Press, Cambridge; 2001.

[3] Johnson RA, Wichern DW. Applied multivariate statistical analysis. Prentice-Hall, Inc., Englewood Cliffs, New York; 1991.

[4] Pearson K. On lines and planes of closest fit to systems of points in space. Philosophical Magazine. 1901;2(11):559–572.

[5] Hotelling H. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology. 1933;24:417–441.

[6] Jolliffe IT. Principal component analysis. Series: 2nd ed., Springer, NY; 2002.

[7] Bishop, Christopher M. Pattern recognition and machine learning. Springer, New York; 2007.

[8] Golub GH, Van Loan Ch. F. Matrix computations. Johns Hopkins University Press (Fourth Edition), Baltimore; 2013.

[9] Bobrowski L. Large matrices inversion using the basis exchange algorithm. British Journal of Mathematics & Computer Science. 20017;21:1.
Available:http://www.sciencedomain.org/abstract/18203
ISSN: 2231-0851

[10] Bobrowski L. Discovering main vertexical planes in a multivariate data space by using CPL functions. In: Advances in Data Mining: Applications and Theoretical Aspects, Ed. Perner P., Springer Verlag, Berlin. 2014;200-213.

[11] Bobrowski L. Biclustering based on collinear patterns. IWBBIO (1), Springer, New York. 2017;134-144.

[12] Bobrowski L. Data mining based on convex and piecewise linear criterion functions (in Polish). Technical University Białystok; 2005.

[13] Bobrowski L. Design of piecewise linear classifiers from formal neurons by some basis exchange technique. Pattern Recognition. 1991;24(9):863-870.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://sciencedomain.org/review-history/20424