Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 114 (2017) 234–241

Procedia Computer Science

www.elsevier.com/locate/procedia

Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems, CAS
October 30–November 1, 2017, Chicago, Illinois, USA

Partial Least-Squares Method for Three-Mode Three-Way Datasets Based on Tucker Model

Jun Tsuchida[a], Hiroshi Yadohisa[b]*

[a]*Graduate School of Culture and Information Science ,Doshisha University, Kyoto, 610-0394, Japan*
[b]*Department of Culture and Information Science ,Doshisha University, Kyoto, 610-0394, Japan*

Abstract

When analyzing two three-mode three-way datasets (object × variable × condition), the objective is to obtain common factors that show the relationships between the two datasets. The partial least-squares (PLS) method has been applied to such datasets to investigate the common factors. However, the PLS method was proposed for two-mode two-way datasets, such as multivariate datasets. Therefore, this method does not consider the condition when searching for relationships between datasets; that is, it tends to regard the same variable under different conditions as different variables. To address this problem, we extended the PLS method to three-mode three-way datasets by using the Tucker model so that the same variable under different conditions is regarded as the same. Moreover, we can apply the proposed method to three-mode three-way datasets with different dimensions for the conditions and variables, and the output is obtained in the form of three-mode three-way datasets. We show the advantage of the proposed method by applying it to a multicollinearity case as a numerical example.

Peer-review under responsibility of the scientific committee of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.

Keywords: Alternative least squares; Dimensional reduction; Linear regression model

* Corresponding author. Tel.:+81-074-65-7857;
E-mail address: jun.tsuchida0328@gmail.com

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.
10.1016/j.procs.2017.09.065

1. Introduction

Sophisticated observation technology means that three-mode three-way data are obtained in many research areas. One example of three-mode three-way data is longitudinal data, where values are determined according to the combination of the object, variable, and time. In general, three-mode three-way data are defined as the combination of three sets, such as objects, variables, and conditions. There are problems with applying multivariate analysis such as linear regression to three-mode three-way datasets. For example, multicollinearity tends to occur when linear regression is applied to a three-mode three-way dataset. This is caused by the characteristics of multivariate analysis. Multivariate analysis for two-mode two-way data, such as multivariate data, regards the same variable under other conditions as different variables. The correlation between the same variables under other conditions tends to be strong.

To address this problem, many researchers have proposed multivariate methods for three-mode three-way data. For example, the Tucker model [4] is a well-known principal component analysis method for three-mode three-way data. PARAFAC [1] is a commonly used factor analysis method for three-mode three-way data. The Tucker model has many extensions, such as sparse modeling [2] and three-way clustering with dimensional reduction [6]. However, these methods are applicable to one dataset. When two three-mode three-way datasets are being analyzed, the interest is in obtaining common factors that show the relationships between the two datasets.

The partial least-squares (PLS) method [7] has been applied to datasets in order to investigate the common factors. However, the PLS method was proposed for two-mode two-way datasets. Thus, we propose a PLS method for three-mode three-way datasets. The proposed method is based on the Tucker model, which is a principal component analysis method for three-mode three-way data. The PLS method is defined as maximizing the square covariance between response and independent variables. We used the concept of principal component analysis by modifying maximizing the variance to maximizing the square covariance. Another reason for using the Tucker model is the number of dimensions. The Tucker model can set different numbers of dimensions for the variables and conditions. This characteristic is an advantage for analysis when datasets have large numbers of variables and conditions because variables and conditions often cannot be assumed to have the same dimensions.

The importance of the proposed PLS method for three-mode three-way datasets is that it outputs three-mode three-way data from the response and independent variables. That is, we can obtain an output with the same structure as the data matrix.

The remainder of this paper is organized as follows. Section 2 defines the proposed model; we introduce a regression model for three-mode three-way datasets. Section 3 presents numerical examples to demonstrate the advantages of the proposed model. Section 4 presents the concluding remarks and suggestions for future study.

2. Partial least squares for three-mode three-way data

In this section, we explain the proposed model. First, we introduce linear regression for three-mode three-way datasets. We assume that response and independent variables are three-mode three-way datasets. Then, we introduce the PLS method for three-mode three-way datasets. Finally, we explain the algorithm and the update formula for the parameters.

2.1. Three-mode three-way regression

Given two three-mode three-way datasets $\boldsymbol{X} \in \mathbb{R}^{n\times p\times k}$ and $\boldsymbol{Y} \in \mathbb{R}^{n\times q\times \ell}$, the linear regression model is written as follows:

$$\boldsymbol{Y}_1 = \boldsymbol{X}_1 (\boldsymbol{C} \otimes \boldsymbol{B}) + \boldsymbol{E},$$

where $\boldsymbol{C} \in \mathbb{R}^{k\times \ell}$ is the coefficient matrix for conditions, $\boldsymbol{B} \in \mathbb{R}^{p\times q}$ is the coefficient matrix for variables, and $\boldsymbol{E} \in \mathbb{R}^{n\times q\ell}$ is the error matrix. $\otimes$ denotes the Kronecker product. $\boldsymbol{X}_1 \in \mathbb{R}^{n\times pk}$ and $\boldsymbol{Y}_1 \in \mathbb{R}^{n\times q\ell}$ are data matrices that are $\boldsymbol{X}$ and $\boldsymbol{Y}$ unfolded by mode 1. That is, the rows of $\boldsymbol{X}_1$ and $\boldsymbol{Y}_1$ are objects. When we set $\boldsymbol{A} = (\boldsymbol{C} \otimes \boldsymbol{B})$, this model is the same as a multivariate linear regression model. By decomposing the coefficient matrices to conditions and variables, this model determines the same variable under different conditions to be the same variable. Moreover, we obtain response variables as a three-mode three-way dataset. That is, this model does not corrupt the structure of the datasets.

When we use the least-squares method for parameter estimation, the objective function f is obtained as follows:

$$\boldsymbol{f}(\boldsymbol{C},\boldsymbol{B}|\boldsymbol{X},\boldsymbol{Y}) = ||\boldsymbol{Y}_1 - \boldsymbol{X}_1\,(\boldsymbol{C} \otimes \boldsymbol{B})||^2 \rightarrow \text{minimize},$$

where $||\cdot||$ is the Frobenius norm.

This objective function is rewritten as follows:

$$\begin{aligned}\boldsymbol{f}(\boldsymbol{C},\boldsymbol{B}|\boldsymbol{X},\boldsymbol{Y}) &= ||\boldsymbol{Y}_2 - \boldsymbol{B}'\boldsymbol{X}_2\,(\boldsymbol{C} \otimes \boldsymbol{I}_n)||^2 \\ &= ||\boldsymbol{Y}_3 - \boldsymbol{C}'\boldsymbol{X}_3\,(\boldsymbol{B} \otimes \boldsymbol{I}_n)||^2,\end{aligned} \tag{1}$$

where $\boldsymbol{X}_2 \in \mathbb{R}^{p\times nk}$ and $\boldsymbol{Y}_2 \in \mathbb{R}^{q\times n\ell}$ are matrices unfolded by mode 2 and $\boldsymbol{X}_3 \in \mathbb{R}^{k\times np}$ and $\boldsymbol{Y}_3 \in \mathbb{R}^{\ell\times nq}$ are matrices unfolded by mode 3. The second term of Equation (1) is the same as the regression given $\boldsymbol{C}$. The third term of Equation (1) is also the same as the regression given $\boldsymbol{B}$. Thus, by using an alternative least-squares algorithm, we can solve the objective function. We discuss the details of the update formula in Section 2.3.

2.2. Partial least-squares method for three-mode three-way datasets

Before presenting the proposed PLS method for three-mode three-way datasets, we introduce the two-mode two-way PLS method. Many researchers have proposed objective functions for the PLS method. We adopted the objective function proposed by Sun et al. [3] because it is defined as a dimensional reduction problem. The objective function of the PLS method is as follows:

$$\boldsymbol{f}(\boldsymbol{A}|\boldsymbol{X},\boldsymbol{Y}) = ||\boldsymbol{Y}_1{}'\boldsymbol{X}_1\,\boldsymbol{A}||^2 \rightarrow \text{maximize},$$

$$\text{Subject to } \boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I},$$

where $\boldsymbol{A} \in \mathbb{R}^{pk\times r}$ is the weight matrix for dimensional reduction and r is the number of dimensions. We use this objective function to obtain parameters that maximize the square covariance. We use $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I}$ and ten Berge's theorem [3] to obtain the parameters by eigenvalue decomposition. We modify this model to be applicable to three-mode three-way datasets in the same manner as a linear regression model. We set $\boldsymbol{A} = \boldsymbol{C}_x \otimes \boldsymbol{B}_x$. Then, the objective function of the PLS method is as follows:

$$\boldsymbol{f}(\boldsymbol{B}_x,\boldsymbol{C}_x|\boldsymbol{X},\boldsymbol{Y}) = ||\boldsymbol{Y}_1{}'\boldsymbol{X}_1\,(\boldsymbol{C}_x \otimes \boldsymbol{B}_x)||^2 \rightarrow \text{maximize},$$

$$\text{Subject to } \boldsymbol{B}_x{}'\boldsymbol{B}_x = \boldsymbol{I}_{\mathrm{r_b}},\ \boldsymbol{C}_x{}'\boldsymbol{C}_x = \boldsymbol{I}_{r_c},$$

where $\boldsymbol{B}_x \in \mathbb{R}^{p\times r_b}$ is the weight matrix of variables and $\boldsymbol{C}_x \in \mathbb{R}^{k\times r_c}$ is the weight matrix of conditions. The weight matrices of the variables and conditions are orthogonal because $(\boldsymbol{C}_x \otimes \boldsymbol{B}_x)'(\boldsymbol{C}_x \otimes \boldsymbol{B}_x) = \boldsymbol{I}$ holds. By decomposing the weight matrix for dimensional reduction with the Kronecker product, we obtain three-mode three-way data with reduced dimensions. That is, we obtain the PLS score as a three-mode three-way dataset. Therefore, we can apply linear regression to three-mode three-way datasets where the independent variables are PLS scores.

When we define the PLS scores $\boldsymbol{T}_1 = \boldsymbol{X}_1(\boldsymbol{C}_x \otimes \boldsymbol{B}_x)$, the formula of the PLS model for three-mode three-way datasets is written as follows:

$$\begin{aligned} \boldsymbol{Y}_1 &= \boldsymbol{X}_1 (\boldsymbol{C}_x \otimes \boldsymbol{B}_x) \\ &= \boldsymbol{T}_1 (\boldsymbol{C} \otimes \boldsymbol{B}) \\ &= \boldsymbol{X}_1 \boldsymbol{W}, \end{aligned}$$

where $\boldsymbol{W} = (\boldsymbol{C_x C}) \otimes (\boldsymbol{B_x B})$ is the coefficient matrix of the PLS method. When we calculate the predictor from new datasets, we use this coefficient matrix. Because this matrix is obtained by three-mode three-way linear regression, the predictor is three-mode three-way dataset. That is, we can interpret the same variable under other conditions as the same variable.

There are two advantages of our method compared to the method for two-mode two-way data. First, our method does not corrupt the three-mode three-way data structure. When we apply a method for two-mode two-way data to three-mode three-way data, the same variable under other conditions is regarded as different variables. Therefore, it is difficult to interpret the parameters. With the proposed method, however, we can easily interpret the parameters of variables and conditions. Second, our method reduces the numbers of variable and conditions. In contrast, the method for two-mode two-way data tends to aggregate the same variables under other conditions because the correlation coefficients between the same variables under other conditions tend to be higher.

2.3. Algorithm

The overall algorithm for parameter estimation is as follows: first, we calculate the PLS scores; then, we apply linear regression to the three-mode three-way datasets. The steps of the algorithm are as follows:

- Step 1: Set the number of dimension r_b and r_c
- Step 2: Estimate the PLS scores by using three-mode three-way data
- Step 3: Estimate the coefficients matrix by using the PLS scores.

We use the alternative least-squares algorithm to estimate the parameters for the PLS method and linear regression because deriving the Kronecker product is complex. We discuss the update formula for the PLS method in Subsections 2.3.1 and 2.3.2 and the update formula for linear regression of three-mode three-way data in Subsections 2.3.3 and 2.3.4.

2.3.1. *Update* $\boldsymbol{C_x}$

In this subsection, we explain the update formula for the weight matrix of conditions $\boldsymbol{C_x}$. Given $\mathbf{B_x}$, the objective function of the PLS method for three-mode three-way datasets as follows:

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{C}_x|\boldsymbol{B}_x, \boldsymbol{X}, \boldsymbol{Y}) &= ||\boldsymbol{Y}_1{}'\boldsymbol{X}_1 (\boldsymbol{C}_x \otimes \boldsymbol{B}_x)||^2 \\ &= ||\boldsymbol{C}_x{}'\boldsymbol{X}_3 (\boldsymbol{B}_x \otimes \boldsymbol{Y}_1)||^2 \\ &= \mathrm{tr}(\mathbf{C}_x \boldsymbol{X}'{}_3 (\boldsymbol{B}_x' \boldsymbol{B}_x \otimes \boldsymbol{Y}_1' \boldsymbol{Y}_1) \boldsymbol{X}_3 \mathbf{C}_x') \end{aligned}$$

Subject to $\boldsymbol{C}_x{}'\boldsymbol{C}_x = \boldsymbol{I}_{\boldsymbol{r_c}}$.

By using Lagrange multipliers, this objective function is modified as follows:

$$\mathrm{tr}(\mathbf{C}_x \boldsymbol{X}'{}_3 (\boldsymbol{B}_x' \boldsymbol{B}_x \otimes \boldsymbol{Y}_1' \boldsymbol{Y}_1) \boldsymbol{X}_3 \mathbf{C}_x') - \mathrm{tr}(\boldsymbol{\Lambda}(\mathbf{C}_x' \mathbf{C}_x - \boldsymbol{I})).$$

This objective function is the same as that of principal component analysis. Thus, the update formula of $\boldsymbol{C_x}$ is as follows:

$$\boldsymbol{C}_x = \boldsymbol{U}_c,$$

where $\boldsymbol{U_c}$ is the eigenvector matrix of $\boldsymbol{X'}_3(\boldsymbol{B}'_x\boldsymbol{B}_x \otimes \boldsymbol{Y}'_1\boldsymbol{Y}_1)\boldsymbol{X}_3$.

2.3.2. Update $\boldsymbol{B_x}$

In this subsection, we explain the update formula for the weight matrix of variable $\boldsymbol{B}_x$. Given $\boldsymbol{C_x}$, the objective function of the PLS method for three-mode three-way datasets is as follows:

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{B}_x|\boldsymbol{C}_x, \boldsymbol{X}, \boldsymbol{Y}) &= ||\boldsymbol{Y}_1{}'\boldsymbol{X}_1\ (\boldsymbol{C}_x \otimes \boldsymbol{B}_x)||^2 \\ &= ||\boldsymbol{B}_x{}'\boldsymbol{X}_2\ (\boldsymbol{C}_x \otimes \boldsymbol{Y}_1)||^2 \\ &= \mathrm{tr}(\boldsymbol{B}_x\boldsymbol{X'}_2(\boldsymbol{C}'_x\boldsymbol{C}_x \otimes \boldsymbol{Y}'_1\boldsymbol{Y}_1)\boldsymbol{X}_2\boldsymbol{B}'_x) \\ \text{subject to } \boldsymbol{B}_x{}'\ \boldsymbol{B}_x &= \boldsymbol{I}_{\mathbf{r_b}}. \end{aligned}$$

This objective function is the same as that of principal component analysis. Thus, the update formula of $\boldsymbol{B_x}$ is as follows:

$$\boldsymbol{B}_x = \boldsymbol{U}_b,$$

where $\boldsymbol{U_b}$ is the eigenvector matrix of $\boldsymbol{X'}_2(\boldsymbol{C}'_x\boldsymbol{C}_x \otimes \boldsymbol{Y}'_1\boldsymbol{Y}_1)\boldsymbol{X}_2$.

2.3.3. Update $\boldsymbol{C}$

In this subsection, we explain the update formula for the coefficient matrix of conditions. Given $\boldsymbol{B}$ and the three-mode three-way PLS score $\boldsymbol{T}$, then the objective function of the linear regression is as follows:

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{C}|\boldsymbol{B}, \boldsymbol{T}, \boldsymbol{Y}) &= ||\boldsymbol{Y}_3 - \boldsymbol{C}'\boldsymbol{T}_3\ (\boldsymbol{B}\ \otimes \boldsymbol{I_n})||^2 \\ &= ||\boldsymbol{Y}'_3 - (\boldsymbol{B}\ \otimes \boldsymbol{I_n})\boldsymbol{T}'_3\boldsymbol{C}||^2. \end{aligned}$$

This objective function is the same as that of linear regression. Therefore, the update formula is as follows:

$$\boldsymbol{C} = (\boldsymbol{T_3}(\boldsymbol{B}'\boldsymbol{B}\ \otimes \boldsymbol{I_n})\boldsymbol{T}'_3)^{-1}\boldsymbol{T}_3(\boldsymbol{B}' \otimes \boldsymbol{I_n})\mathbf{Y}_3{}'.$$

*2.3.4. Update **B***

In this subsection, we explain the update formula of the coefficient matrix for variables. Given $\boldsymbol{T}$ and $\boldsymbol{C}$, the objective function of the linear regression is as follows:

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{B}|\boldsymbol{C}, \boldsymbol{T}, \boldsymbol{Y}) &= ||\boldsymbol{Y}_2 - \boldsymbol{B}'\boldsymbol{T}_2\ (\boldsymbol{C}\ \otimes \boldsymbol{I_n})||^2 \\ &= ||\boldsymbol{Y}'_2 - (\boldsymbol{C}\ \otimes \boldsymbol{I_n})\boldsymbol{T}'_2\boldsymbol{B}||^2. \end{aligned}$$

This objective function is the same as that of linear regression. Therefore, the update formula is as follows:

$$\boldsymbol{B} = (\boldsymbol{T_2}(\boldsymbol{C}'\boldsymbol{C}\ \otimes \boldsymbol{I_n})\boldsymbol{T}'_2)^{-1}\boldsymbol{T}_2(\boldsymbol{C}' \otimes \boldsymbol{I_n})\mathbf{Y_2}'.$$

3. Numerical example

In this section, we present numerical examples to compare the prediction errors of the proposed method and the two-mode two-way method. The first case is the three-mode three-way settings. The weight matrix for dimensional reduction is obtained by the Kronecker product for the weight matrices of the conditions and variables. The second case is a special case for the three-mode three-way settings. The weight matrices of the conditions and variables have a simple structure.

3.1. Settings

We generate the PLS scores $\boldsymbol{T}_1 \in \mathbb{R}^{\boldsymbol{n}\times\mathbf{3}*\mathbf{3}}$. The row vectors of $\boldsymbol{T}_1$ follow an identical and independent multivariate normal distribution with the mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{I}$. In the first case, the weight matrices of $\boldsymbol{C}_x \in \mathbb{R}^{16\times 3}$ and $\boldsymbol{B}_x \in \mathbb{R}^{15\times 3}$ generate uniform distributions with min = -1 and max = 1. However, these matrices are not orthogonal. By using singular value decomposition, we obtain the orthogonal matrices. In the second case, the weight matrices of $\boldsymbol{C}_x \in \mathbb{R}^{16\times 3}$ and $\boldsymbol{B}_x \in \mathbb{R}^{15\times 3}$ are generated as follows:

$$\boldsymbol{C}_x = (\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3), \boldsymbol{c}'_1 = (\mathbf{1}_5, \mathbf{0}_{11}), \boldsymbol{c}'_2 = (\mathbf{0}_5, \mathbf{1}_6\ \mathbf{0}_5), \boldsymbol{c}'_3 = (\mathbf{0}_{11}, \mathbf{1}_5),$$
$$\boldsymbol{B}_x = (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3), \boldsymbol{b}'_1 = (\mathbf{1}_5, \mathbf{0}_{10}), \boldsymbol{b}'_2 = (\mathbf{0}_5, \mathbf{1}_5\ \mathbf{0}_5), \boldsymbol{b}'_3 = (\mathbf{0}_{10}, \mathbf{1}_5),$$

where $\mathbf{1_m}$ is the *m*-dimensional vector whose elements are 1 and $\mathbf{0_m}$ is the *m*-dimensional vector whose elements are 0. The independent variables $\boldsymbol{X}_1 \in \mathbb{R}^{\boldsymbol{n}\times\mathbf{15}*\mathbf{16}}$ are defined as $\boldsymbol{X}_1 = \boldsymbol{T}_1(\boldsymbol{C}_x \otimes \boldsymbol{B}_x)' + \mathbf{E}$, where $\mathbf{E}$ is a noise matrix with elements following an identical and independent normal distribution with a mean of 0 and standard deviation of *sd*. The response variables $\mathbf{Y}_1 \in \mathbb{R}^{\boldsymbol{n}\times\mathbf{15}*\mathbf{16}}$ are generated as $\boldsymbol{T}_1(\boldsymbol{C}_x \otimes \boldsymbol{B}_x)' + \boldsymbol{E}_y$. $\boldsymbol{E}_y$ is generated in the same manner as $\mathbf{E}$. We set the number of training samples *n* = (300,500) and the standard deviation of the noise *sd* = (0.1, 0.5, 1). We used the following evaluation criterion for the mean squared error:

$$\frac{1}{n_p q \ell} ||\mathbf{Y}_{1_\mathrm{p}} - \widehat{\boldsymbol{Y}}_{1_p}||^2,$$

where n_p is the number of test datasets. We set $n_p = 300$ in all cases. $\mathbf{Y}_{1_\mathrm{p}}$ is generated in the the same manner as $\mathbf{Y}_1$, and $\widehat{\boldsymbol{Y}}_{1_p}$ is the predictor when using the estimated parameters.

The following methods were compared: the combination of three-mode three-way PLS and two-mode two-way linear regression, three-mode three-way regression, two-mode two-way PLS, and two-mode two-way linear regression. The combined method of three-mode three-way PLS and two-mode two-way linear regression uses the PLS score estimated for the three-mode three-way dataset as independent variables.

3.2. Results

Table 1 and Table 2 show the mean and standard deviation in both cases. The value in cell is mean of predict error. The bracketed value shows standard deviation. The Highlighted cells show the best result among the same situation. Table shows, from the left, Three-Three: Three-mode three-way regression after applying PLS method for three-mode three-way, Three-two: two-mode two-way regression after PLS method for three-mode three-way, Three-way reg: three-mode three-way regression, Two-Two: two-mode two-way regression after two-mode two-way PLS method, Two-way reg: two-mode two-way regression. From Table 1 and 2, proposed method is best result among comparing methods.

Fig 1 and Fig 2 show the boxplots. The x-axis of Figure1 and 2 show, from the left, 3-3: Three-mode three-way regression after applying PLS method for three-mode three-way, 3-2: two-mode two-way regression after PLS method for three-mode three-way, 3: three-mode three-way regression, 2-2: two-mode two-way regression after two-mode two-way PLS method, 2: two-mode two-way regression. From Figures, two-mode two-way models are the wide range of the predict error. On the other hand, three-mode three-way models including our model are the narrow range.

Table 1. the mean and standard deviation of predict error in first case.

	Three-Three	Three-Two	Three-way reg	Two-Two	Two-way reg
***n*=300, *sd*= 0.1**	1.171(0.041)	1.204(0.041)	1.185(0.042)	2.970(0.496)	6.000(0.303)
***n*=300, *sd*= 0.5**	0.120(0.005)	0.123(0.005)	0.121(0.005)	2.326(0.629)	0.616(0.031)
***n*=300, *sd*= 1**	0.598(0.022)	0.615(0.023)	0.606(0.023)	2.693(0.543)	3.071(0.151)
***n*=500, *sd*= 0.1**	0.202(0.008)	0.205(0.008)	0.203(0.008)	4.141(1.017)	0.389(0.016)
***n*=500, *sd*= 0.5**	1.962(0.063)	1.994(0.064)	1.975(0.064)	5.093(0.822)	3.783(0.125)
***n*=500, *sd*= 1**	0.994(0.041)	1.010(0.042)	1.001(0.042)	4.674(1.238)	1.919(0.084)

Table 2. the mean and standard deviation of predict error in second case

	Three-Three	Three-Two	Three-way reg	Two-Two	Two-way reg
***n*=300, *sd*= 0.1**	0.121(0.005)	0.124(0.005)	0.122(0.006)	2.321(0.631)	0.616(0.032)
***n*=300, *sd*= 0.5**	1.177(0.043)	1.211(0.044)	1.190(0.043)	2.916(0.534)	6.030(0.275)
***n*=300, *sd*= 1**	0.598(0.027)	0.615(0.027)	0.606(0.027)	2.608(0.603)	3.073(0.163)
***n*=500, *sd*= 0.1**	0.200(0.008)	0.203(0.008)	0.202(0.008)	3.588(0.874)	0.385(0.016)
***n*=500, *sd*= 0.5**	1.948(0.071)	1.981(0.072)	1.962(0.072)	4.839(0.878)	3.757(0.139)
***n*=500, *sd*= 1**	0.996(0.041)	1.012(0.042)	1.003(0.042)	4.344(0.950)	1.916(0.080)

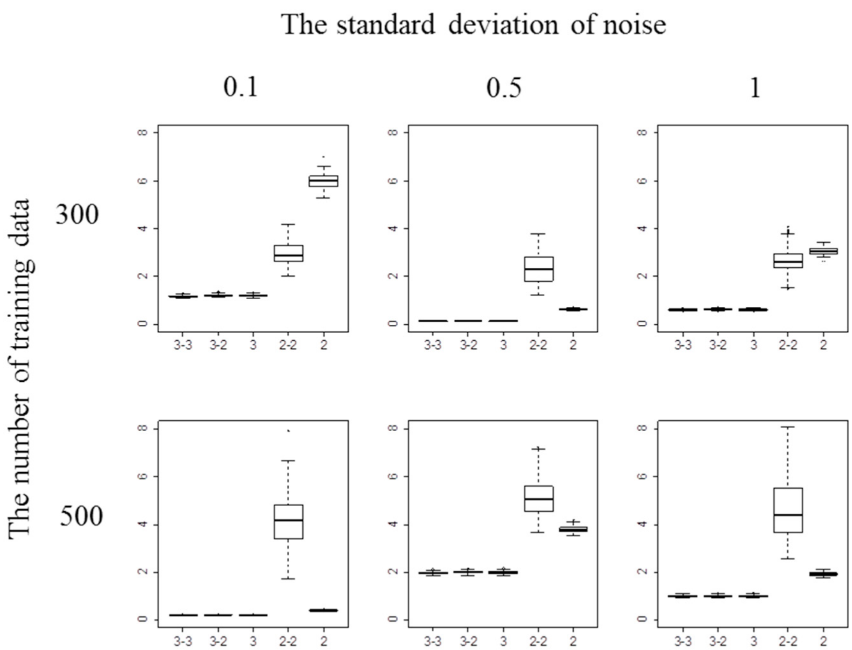

Fig 1. the boxplots of prediction error in the first case

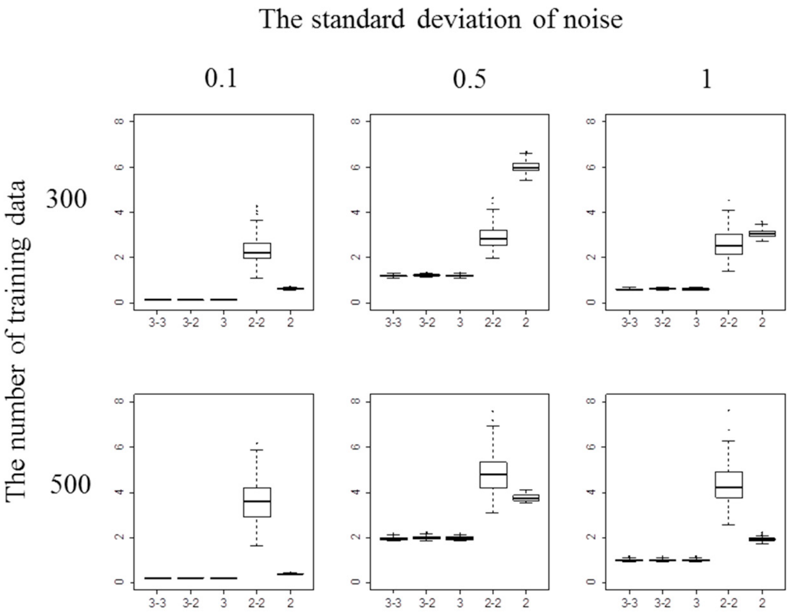

Fig 2. the boxplots of prediction error in the first case

4. Conclusion

We propose a PLS method for three-mode three-way datasets. Because we define the estimation of the PLS score as maximizing the square covariance between dependent and independent variables, we can easily estimate the parameters by using eigenvalue decomposition. The numerical examples showed that the three-mode three-way method performed better than the two-mode two-way method in both cases. Therefore, we demonstrated the advantages of our model for the analysis of three-mode three-way datasets.

We have two topics of future studies planned. The first is developing a method for determining the number of dimension. Because our model have two weight matrices, we need a method to determine the number of dimensions for the pair. One suggestion is using the scree plot variance. The second is determining the relationship between our method and other methods. Orthogonal PLS is equivalent to canonical correlation analysis in the two-mode two-way case [3]. However, the relationship between our method and other three-mode three-way methods is not obvious because of the Kronecker product and equivalence of strongly constrained parameters.

References

[1] Harshman, R. A. (1970). Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis. *UCLA Working Papers in Phonetics*, **16**, 1–84.

[2] Ikemoto, H., and Adachi, K. (2016). Sparse Tucker2 analysis of three-way data subject to a constrained number of zero elements in a core array, *Computational Statistics and Data Analysis*, **98**, 1–16.

[3] Sun, L., Ji, S., Yu, S., and Ye, J. (2009, July). On the equivalence between canonical correlation analysis and orthonormalized partial least squares. In *IJCAI*, **9**, 1230–1235.

[4] ten Berge, J. M. (1993). *Least squares optimization in multivariate analysis*. Leiden, The Netherlands: DSWO Press, Leiden University.

[5] Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, **31**, 279–311.

[6] Vichi, M., Rocci, R., and Kiers, A.H. (2007) Simultaneous component and clustering models for three-way data: within and between approaches. *Journal of Classification*, **24**, 71–98.

[7] Wold, H. (1975). *Soft modeling by latent variables: the nonlinear iterative partial least squares approach*. Perspective in Probability and Statistics, Papers in Honour of M. S. Bartlett, 520–540, Academic Press.