Tensor Completion Based on Tucker Factorization for Missing Value Imputation

Thao Nguyen Thieu[1], Hyung-Jeong Yang[1*], Tien Duong Vu[1], Sun-Hee Kim[2]

[1] Dept. of Electronics and Computer Engineering, Chonnam National University, Gwangju, South Korea.
[2] Dept. of Brain and Cognitive Engineering, Korea University, Seoul, South Korea.

* Corresponding author. Email: hjyang@jnu.ac.kr

Abstract: In this paper, we propose an algorithm to estimate missing values of electroencephalogram (EEG) signals in tensor form. EEG signals is a type of brain data which is recorded from many electrodes placed on the scalp. EEGs are easily missing values if recording in the bad condition or transforming from raw data into more understandable structure. The method is proposed to solve the problem in tensor form and it is built on recent studies about tensor completion using decomposition method. First, we give a definition for tensor completion using the component of decomposing process. Second, the tensor completion is formulated as a convex optimization problem. Therefore, a first order optimization method, Nonlinear Conjugate Gradient, is applied to solve it. In this paper, we also show several numerical experiments that demonstrate the effective of the proposed method.

Key words: Tucker decomposition, tensor completion, electroencephalogram (EEG), optimization.

1. Introduction

Electroencephalogram (EEG) is a signal recording the electrical activity of the brain through many sensors (electrodes) placed on the scalp. It measures voltage fluctuations resulting from ionic current flows within the neurons of the brain. Recording the activity of the human brain is one of the most important works in the not only health care field but also other areas such as bio-information, signal processing, and data mining. It can be used to analyze and study the extraordinary mutation of the human brain, even in the canine brain, and its effect on motion and action of subjects [1]. An EEG dataset can be reconstructed into a three-dimensional or four-dimensional data. This form is considered as a tensor.

Tensor is a kind of data in multidimensional space. More formally, an N-way or Nth-order tensor is an element of the tensor product of N vector spaces [2]. Tensor analysis is an approach that has many applications in many fields such as physics, mechanics, and information sciences. Two most important applications of tensor are Tensor Completion and Tensor Decomposition. Tensor decomposition is used to divide a tensor into many smaller parts, so that the process will become faster and simpler. Tensor Completion is a method to recover missing data of a tensor. As mathematical definition, tensor completion methods can reconstruct a tensor from a smaller part of its entries.

In recent years, recovering a tensor with low rank is an open problem. The main idea of low rank is to solve the problems related to the lost information in a matrix that has low rank [3]. Lost information or also called "missing values" usually occurs when collecting data as recording signal or taking a picture in a bad

condition or when processing data from the raw form into the more logical structure. Generally, for more understandable, low-rank matrix completion is a method that can recover a matrix from a subset of its entries. Low-rank matrix is a very useful method for various fields such as signal processing, computer vision, biomedical and many others. However, a real dataset is presented not only in second-order form as matrix but also in higher-order form as a tensor. Therefore, low-rank tensor completion is being focused on studying more and more commonly.

The low-rank tensor completion problem is a difficult non-convex problem [4]. Therefore, Nuclear Norm Approximation (NNA) is approached to tensor completion. Most of previous papers related to tensor completion used NNA to replace the rank of tensor. The idea of NNA is to compute the sum of some singular values of that matrix. The nuclear norm regularization promotes a low-rank solution, which is a key idea that can be applied to many applications such as recommender system, dimension reduction in multivariate regression, and multi-task learning. Generally, tensor completion is developed based on optimization idea: minimizing or maximizing a function to obtain the optimal solution. The optimal solution of the optimization problem is the reconstructed tensor. Some researchers proposed another idea to recover a tensor by using the components after decomposing. Those are using the weight least square matrix of tensor [5].

In this paper, we propose an application of Tensor Completion using Tucker factorization for recovering lost values in EEG signals. We use Tucker factorization to build the objective function for the optimization problem. Nonlinear Conjugate Gradient [6] is applied to find the optimal solution. We evaluate the effect of the proposed method on EEG data by experiments.

The rest of this paper is organized as follows. Section 2 briefly reviews the related works. Section 3 introduces the proposed method. Experimental results are shown in Section 4 to demonstrate the performance of the proposed algorithm and the final conclusions are extracted in Section 5.

2. Related Works

Tensor completion is an extension of matrix completion which is usually used to recover a data in a matrix form. Matrix which has low rank become the subject for simple recovery [3], [7], [8]. Low-rank matrix completion arises in many experiments because it can recover a matrix from several entries. In the other hand, many papers are been studying to recover a data in multidimensional space, in tensor form, because of its applications [4], [5], [9]-[11]. Tensor with low rank is a difficult non-convex problem. Therefore, nuclear norm approximation continued approaching to tensor completion. Most of previous papers related to tensor completion used nuclear norm to replace the rank of tensor to build the projective function for an optimization problem. Ref. [9] proposed an extension of low-rank tensor. It focuses on n-rank and proposed three algorithms for tensor completion problem. Simple low rank tensor completion (SiLRTC) is a block coordinate descent method. A faster method be FaLRTC (Fast low-rank tensor completion) is also proposed. Finally, the authors also proposed the highly accurate low rank tensor completion (HaLRTC). Similar approaches were used in papers of Gandy [4] and Tan [11]. However, there are some problems with these approaches. Generally, NNA is the sum of singular values of a matrix.Therefore, when applying these approaches, every iteration of the SVD needs to be computed. This makes these algorithm slower and the complexity of problem will be increased.

Acar [5] also built a model to recover a low rank tensor with missing data but using different approach to recover a low rank tensor with missing values to avoid using SVD. In [5], an objective function formed from the components is proposed after PARAFAC decomposing and its weight least matrix which just contains 0 and 1. Tucker is better than PARAFAC when decomposing a tensor. Marko [10] showed the effects of his approach by some simulation datasets and 3D images with assuming rank of tensor. This makes the results

be not real and still lose many information. The approach used in this paper is similar to the above mentioned papers [5], [10]. However, the rank of tensor used in this paper will be computed by a detail equation to obtain the best one. We also use the error after decomposing as a way to obtain the better result. Our experiments are adapted into EEG signals while the others applied into 3D images.

3. The Proposed Method

These are some notations used in this paper. Tensors are denoted by bolded Euler script letters $(\mathcal{X}, \mathcal{Y})$, matrices are denoted by capital letters (A, B, C). The letter A_n aredenoted the matrices along to mode n of tensor. (a, b, c) are the vectors and (a, b, c) are the scalars. The $(i_1, i_2, \ldots, i_N)$ entry of an N-order tensor $\mathcal{X}$ is denoted by $x_{i_1 i_2 \ldots i_N}$.

Let us define the multiplication of tensor and matrix as follows; The *n-mode product* of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ with matrix $U \in \mathbb{R}^{J \times I_n}, 1 \leq n \leq N$ is denoted by $\mathcal{X} \times_n U$ which is of size $I_1 \times I_2 \times \ldots \times J \times I_{k+1} \times \ldots \times I_N$. Therefore, we have

$$(\mathcal{X} \times_n U)_{i_1 \ldots i_{n-1} j i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \ldots i_N} u_{j i_n}$$

Let assume that the tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$ given in this paper is low-rank tensor. Ranks of a tensor can be chose as a lowest rank or chose randomly. However, in these case, we can assume that we have some approximations of true ranks. In this paper, we used a popular way to compute the rank of tensor [3]. $\mathcal{W}$ is denoted as a tensor with missing observations and $\mathcal{W}$ has the same size as $\mathcal{X}$. Tensor $\mathcal{W}$ can be defined as:

$$w_{i_1 i_2 \ldots i_N} = \begin{cases} 1, \text{if} x_{i_1 i_2 \ldots i_N} \in \Omega \\ 0, \text{if} x_{i_1 i_2 \ldots i_N} \in \Omega^C \end{cases}$$

Here, Ω is a subset of $\{1, \ldots, I_1\} \times \{1, \ldots, I_2\} \times \ldots \times \{1, \ldots, I_N\}$, containing positions of known entries of tensor and Ω^C is its complement.

Following Tucker model, tensor $\mathcal{X}$, which has rank of each modes be R_n $(n = 1, \ldots, N)$, can be factorized as:

$$\mathcal{X} = \mathcal{G} \times_1 A_1 \times_2 A_2 \times_3 \ldots \times_N A_N + error$$

where $\mathcal{G}$ is $r_1 \times r_2 \times \ldots \times r_N$ core tensor and $A_n \in \mathbb{R}^{I_n \times R_n}$ are factor matrices along each mode of tensor.

Our goal is to find core tensor $\mathcal{G}$ and matrices $A_n (n = 1, \ldots, N)$ by optimizing or minimizing the objective function which will be defined as below. Most of the optimization problem is computed based on an objective function and its gradient. As the definition of Tucker decomposition, when decomposing a tensor in $\mathbb{R}^n$, we will get a core tensor and some factor matrices along to each mode. However, in reality, the result of the mode-n product of the core tensor and n factor matrices is not same as the original tensor, just an approximate tensor. A tensor will be decomposed into a core tensor, factor matrices and an "error" tensor.

The N-way objective function for the optimization problem is defined by Marko [10] as (1).

$$f_{\mathcal{W}}(\mathcal{G}, A_1, A_2, \ldots, A_N) = \|\mathcal{W} * (\mathcal{X} - \mathcal{G} \times_1 A_1 \times_2 A_2 \times_3 \ldots \times_N A_N)\|_F^2 \tag{1}$$

The gradient can be obtained by computing the partial derivatives of objective function $f_{\mathcal{W}}$ to each element of the core tensor and the factor matrices after decomposing. The gradient equation can be

rewritten in the matrix notation case as follows:

$$\frac{\partial f_W}{\partial A_n} =$$

$$2\left\{[\mathcal{W} * (\mathcal{G} \times_1 A_1 \times_2 A_2 \times_3 \dots \times_N A_N - \mathcal{X})]_{(n)} \cdot [(\mathcal{G} \times_1 A_1 \times_2 \dots \times_{n-1} A_{n-1} \times_{n+1} A_{n+1} \times_{n+2} \dots \times_N A_N)_{(n)}]^T\right\} \tag{2}$$

Gradient of the objective function on the core tensor is:

$$\frac{\partial f_W}{\partial \mathcal{G}} = 2\{\mathcal{W} * (\mathcal{G} \times_1 A_1 \times_2 A_2 \times_3 \dots \times_N A_N - \mathcal{X}) \times_1 A_1^T \times_2 \dots \times_N A_N^T\} \tag{3}$$

In this paper, we used Nonlinear Conjugate Gradient (NCG) because of its speed.

4. Experiments

In this paper, we experimented two EEG datasets. The first EEG dataset recorded from four dogs with naturally occurring epilepsy using an ambulatory monitoring system. The recording process was acquired on 16 channels. Data are recorded continuously at a sampling frequency of 400 Hz and referenced to the group average. This data set is gotten freely at the International Epilepsy Electrophysiology portal and was developed by the University of Pennsylvania and the Mayo Clinic [12].

The dataset will be set into a tensor of size $16 \times 400 \times 200$ as $channel \times time-point \times segments$ for each subject. Each segment is the signals recorded in one second. In this paper, 200 segments recorded in 200 first seconds are used. The low-n-rank of these tensor is $15 \times 213 \times 200$. It will be computed based on this equation:

$$n\text{-}rank(\mathcal{X}) = (rankX_{(1)},\ rankX_{(2)},\dots,\ rankX_{(N)}) \tag{4}$$

The relative error will be computed for determining the difference of the original data and the new one after recovering.

$$r_{err} = \frac{\left\|\hat{\mathcal{X}} - (\mathcal{X} - e)\right\|_F}{\|\mathcal{X}\|_F} \tag{5}$$

where $\hat{\mathcal{X}}$ denote the output of the proposed method and e is the error after decomposing. The relative error is always nonnegative and the best possible score is 0. The core and factor matrices are initialized by Tucker algorithm in N-way Toolbox [13] and the optimization algorithm will be calculated by Poblano toolbox [14].

The results are shown in Table 1. The table displays the results of the relative error of two tensors, the original tensor and the output of the proposed method, and the time for each run. This correlation is computedfrom equation (5). The rank of core tensor is $15 \times 213 \times 200$ as the low feasible rank of this tensor. The results of four dogs approximately obtain less than 4% error. It means the similarity of the original tensor and its reconstruction using Tucker model is about 96%.

Table 1. The Relative Error

Missing data	Dog 1	Dog 2	Dog 3	Dog 4
30%	1.2%	1.69%	1.59%	1.85%
50%	1.76%	3.3%	1.59%	3.34%
70%	1.49%	4%	2.88%	3.5%

As an example of the original signals of the first dog, its missing versions and the recovered signals are shown in Fig. 1 a) is the figure of the first EEG segment with missing values of Dog 1. Fig. 1 b) is the recovery of (a). These figures show that the proposed method can successfully recover a signal with 70% missing values in a short time.

In this experiment, we also compare the effectiveness of the proposed method with the others that are from Ji Liu's paper [5]. A simple method of Ji Liu, Simple low-rank tensor completion (SiLRTC), is used as an optimization method that can be appliedto a simple convex structure which can be solved by block coordinate descent. A faster version of SiLRTCis also proposed as another comparison method, called FaLRTC. This method uses the smooth model of the convex problem to solve.

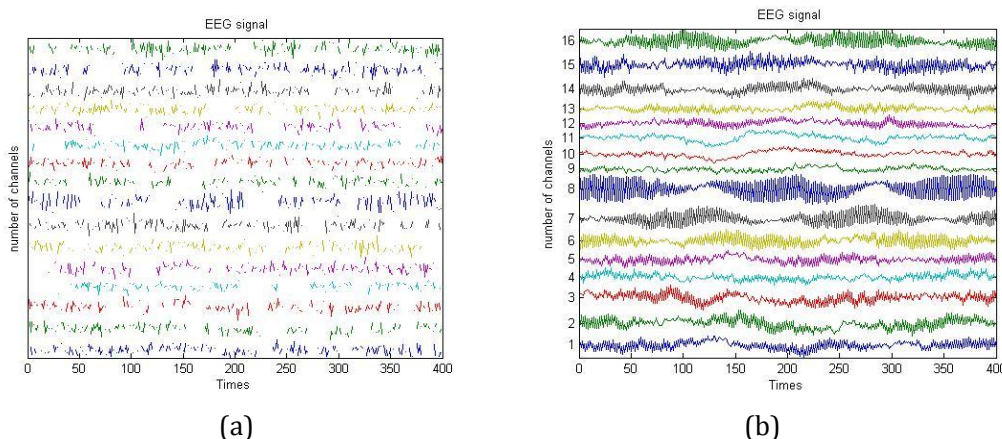(a)   (b)
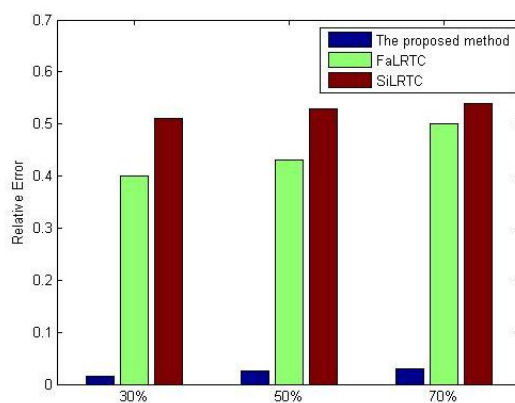
Fig. 1. The EEG segment of dog 1.

Fig. 2. Comparison of the proposed method and the others.

The comparison of the proposed method and two Ji Liu's methods is showed in Fig. 2 with average related errors of four dogs. This figure is plotted by bar function in Matlab. This also shows the big difference output between these methods. These scores are computed by the relative score for the whole tensor. The relative scores computed after applying SiLRTC and FaLRTC are close 50%. It is too high for a completion method. It also shows that these two methods could not complete when applying for EEG data.

For the second experiment, we used the same dataset as Acar [4] for comparison. This dataset contains 64 channels recorded from 14 subjects during stimulation of left and right hand. The whole dataset will be arranged as a *channels* by *time-frequency* by *measurements* tensor of size $64 \times 4392 \times 28$. The rank of this tensor computed by equation (4) is $64 \times 71 \times 28$ [15]. Same as Acar's work, we also randomly remove one channel or more of 64 channels (1, 10, 20, respectively) for each measurement of the whole dataset.

Relative errors to compare to Acar's method are showed in Table 2. The error between two tensors: the original tensor and the reconstructed tensor is computed by formula (5). This experiment shows that the proposed method still give better results than Acar's method if removing 20 channels, about 30% entries of the data.

Table 2. The Comparison of the Relative Error between the Proposed Method and Acar's

Number of missing channels	The proposed method	Acar's method
1	$1.19e - 4$	0.0009
10	$1.53e - 4$	0.0065
20	0.0137	0.1656

5. Conclusion

The aim of this paper is to focus on the problem of low-rank tensor completion used to reconstruct an EEG tensor based on a subset of its entries. In reality, this approach is applied to recover missing values of the EEG dataset in multidimensionality. By using tensor form of data and its application, tensor decomposition, the proposed method can recover a tensor with missing values successfully. Tucker model was considered as a method to decompose a tensor in smaller parts. We also used the first-order optimization, Nonlinear Conjugate Gradient algorithm, to solve the optimization problem which the objective function is built from the outputs of the decomposing process. Furthermore, the rank of tensor is computed by a detail mathematical formula. In this paper, some values of the EEG tensor was removed randomly. The proposed method can recover it with less than 4% relative error. We also showed the effect of the proposed method by comparing with the other algorithms. In future works, we will modified this method to apply into the other datasets by focusing into minimizing the error of the decomposition process.

Acknowledgement

References

[1] Saeid, S., & Jonathon, C. (2007). *EEG Signal Processing*.

[2] Tamara, G. K., & Brett, W. B. (2009). Tensor decomposition and applications. *SIAM Review, 51(3)*, 455-500.

[3] Emmanuel, J., & Benjamin, R. (2009). Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, *9(6),* 717-772.

[4] Silvia, G., Benjamin, R., & Isao, Y. (2011). Tensor completion and low-n-rank tensor recovery via convex optimization. *Inverse Problems*.

[5] Acar, E., *et al*. (2010). Scalable tensor factorizations with missing data. *Proceedings of the Tenth SIAM International Conference on Data Mining* (pp. 701–712).

[6] Wikipedia. Retrieved from the website: https://en.wikipedia.org/wiki/Main_Page

[7] Silvia, G., & Isao, Y. (2010). Convex optimization techniques for the efficient recovery of a sparsely corrupted low-rank matrix. *Journal of Math-for-Industry*, *2*, 147-156.

[8] Vandereycken, B. (2013). Low-rank matrix completion by Riemannian Optimization. *SIAM Journal of Optimization*, *23(2)*, 1214-1236.

[9] Ji, L., Przemyslaw, M., *et al*. (2009). Tensor completion for estimating missing values in visual data. *IEEE*

ICCV.

[10] Marko, F., & Ante, J. (2015). Tucker factorization with missing data with application to low-n-rank tensor completion. *Multidimensional Systems and Signal Processing*, *26*, 677-692.

[11] Cheng, B., *et al.* (2014). Tensor completion via a multi-linear low-n-rank factorization model. *Neurocomputing*, *133*, 161-169.

[12] The International Epilepsy Electrophysiology portal. Retrieved from the website: https://www.ieeg.org/

[13] Andersson, C. A., & Bro, R. (2000). The N-way toolbox for matlab. *Chemometrics and Intelligent Laboratory Systems*.

[14] Daniel, M. D., Tamara, G. K., & Evrim, A. (2010). *Poblano v1.0: A Matlab Toolbox for Gradient-Based Optimization*. Sandia National Laboratories, Albuquerque, NM and Livermore, CA.

[15] Mørup, M., *et al.* (2007). ERPWAVELAB a toolbox for multi-channel analysis of time-frequency transformed event related potentials. *Journal of Neuroscience Methods, 161(2),* 361–368.

Thao Nguyen Thieu received the B.S. from Ho Chi Minh University of Science, Vietnam. She is studying the master's course at Chonnam National University, Korea. She researches interests are in to bio-signals as EEG and related areas.

Hyung-Jeong Yang received her B.S., M.S. and Ph.D. degrees from Chonbuk National University, Korea. She was a post-doc researcher at Carnegie Mellon University, USA. She is currently an associate professor at Dept. of Electronics and Computer Engineering, Chonnam National University, Gwangju, Korea. Her main research interests include multimedia data mining, pattern recognition, artificial intelligence, e-learning, and e-design.

Tien Duong Vu received his B.S. from Ho Chi Minh University of Technology. He is studying the M.S. degree in Chonnam National University, Korea.

Sun-Hee Kim received her M.S. from in Dongguk University, Korea. She received the Ph.D. degree at the Dept. Electronics and Computer Engineering, Chonnam National University, Korea. She is currently a research professor at the Department of Brain and Cognitive Engineering, Korea University, Korea. Her research interests focus on datamining, sensor mining and stream mining, bioinformatics.