Geometric Deep Learning

Contents

1 Combining higher-order PLS with CCN to predict eye-movement

1.1 Abstract

A generalized multilinear regression model, termed the Higher-Order Partial Least Squares (HOPLS) [1], is introduced with the aim to predict a tensor $\underline{\mathbf{Y}}$ from a tensor $\underline{\mathbf{X}}$ through projecting the data onto the latent space and performing regression on the corresponding latent variables. This method could be applied to a wide range of datasets. To solve the problem of predicting eye movement we get time series of frames passed through Convolutional Neural Network (CNN).

1.2 Problem statement

We are given a dataset, that consists of several videos with eye movements and its positions. Let $\underline{\mathbf{X}}$ is an output of CNN, $\underline{\mathbf{Y}}$ — eye coordinate. Assume $\underline{\mathbf{X}} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ and $\underline{\mathbf{Y}} \in \mathbb{R}^{J_1 \times \ldots \times J_M}$. We assume $\underline{\mathbf{X}}$ is decomposed as a sum of rank-$(1, L_2, ..., L_N)$ Tucker blocks, while $\underline{\mathbf{Y}}$ is decomposed as a sum of rank-$(1, K_2, ..., K_M)$ Tucker blocks, which can be expressed as

$$
\begin{aligned}
\underline{\mathbf{X}} &= \sum_{r=1}^{R} \underline{\mathbf{G}}_r \times_1 \mathbf{t}_r \times_2 \mathbf{P}_r^{(1)} \times_3 \ldots \times_n \mathbf{P}_r^{(n-1)} + \underline{\mathbf{E}}_R \\
\underline{\mathbf{Y}} &= \sum_{r=1}^{R} \underline{\mathbf{D}}_r \times_1 \mathbf{t}_r \times_2 \mathbf{Q}_r^{(1)} \times_3 \ldots \times_n \mathbf{Q}_r^{(n-1)} + \underline{\mathbf{F}}_R
\end{aligned}
\tag{1.1}
$$

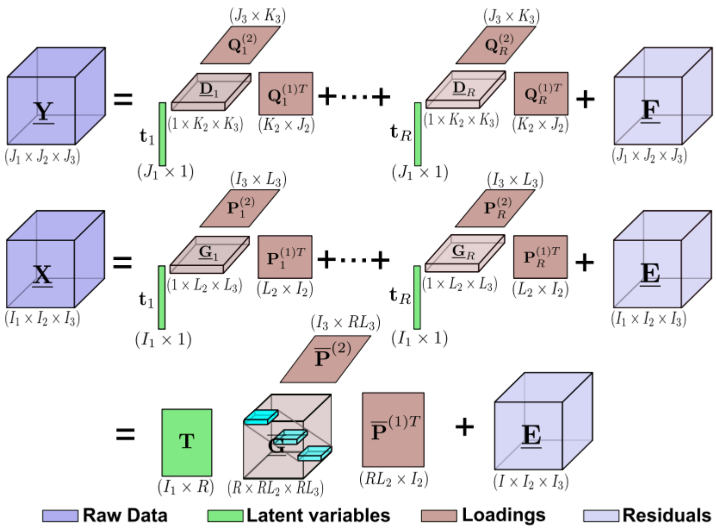

Figure 1: Schematic diagram of the HOPLS model

where R is the number of latent vectors, $t_r \in \mathbb{R}^{I_1}$ is the r-th latent vector, $\left\{\mathbf{P}_r^{(n)}\right\}_{n=1}^{N-1} \in \mathbb{R}^{I_{n+1} \times L_{n+1}}$ and $\left\{\mathbf{Q}_r^{(m)}\right\}_{m=1}^{M-1} \in \mathbb{R}^{J_{n+1} \times K_{n+1}}$ are loading matrices on mode-n and mode-m respectively, and $\underline{\mathbf{G}}_r \in \mathbb{R}^{1 \times L_2 \times \ldots \times L_N}$ and $\underline{\mathbf{D}}_r \in \mathbb{R}^{1 \times K_2 \times \ldots \times K_M}$ are core tensors.

To make a prediction we should use

$$\hat{\underline{\mathbf{Y}}} = \underline{\mathbf{X}}\mathbf{W}\mathbf{Q}^{*\top} \tag{1.2}$$

where $\mathbf{W}$ and $\mathbf{Q}^*$ have R columns, represented by

$$\begin{aligned} \mathbf{w}_r &= \left(\mathbf{P}_r^{(N-1)} \otimes \ldots \otimes \mathbf{P}_r^{(1)}\right) \underline{\mathbf{G}}_r^{+} \\ \mathbf{q}_r^* &= \underline{\mathbf{D}}_r \left(\mathbf{Q}_r^{(M-1)} \otimes \ldots \otimes \mathbf{Q}_r^{(1)}\right) \end{aligned} \tag{1.3}$$

1.3 Problem solution

Use CNN that converts 480x640 pixels images to 24x32 resolution. Its output is $\underline{\mathbf{X}}$. So $N = 3, I_2 = 34, I_3 = 39$. $\underline{\mathbf{Y}}$ tensor is given and it is 2 dimensional tensor with $J_2 = 2$ (Fig. 2). For these tensors we use Higher-Order Partial Least Squares algorithm described in [1] and make a prediction according (1.2).

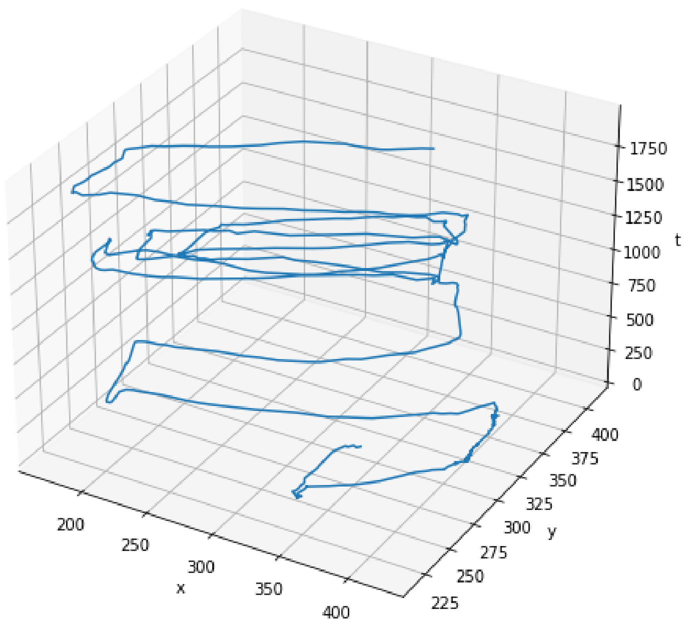

Figure 2: Trajectory of the center of eye

1.4 Code

HOPLS algorithm was made by Arthur Dehgan, co-author of [1] and avalible at https://github.com/arthurdehgan/HOPLS. The code for computation experiment is located on https://github.com/artem062/ForecastingMethods.

1.5 Experiment

It was used LPW dataset, that consist of 66 videos with eye moving. This videos was compressed and converted to the tensors. To archive the best R plot the value of $Q^2 = 1 - \|\underline{\mathbf{Y}} - \hat{\underline{\mathbf{Y}}}\|_F^2 / \|\underline{\mathbf{Y}}\|_F^2$.

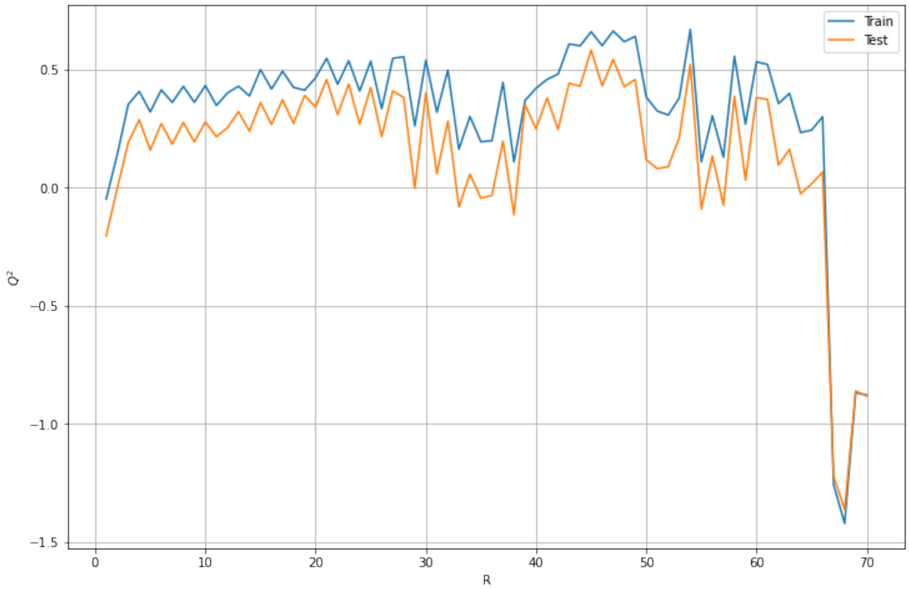

Figure 3: Graph of dependence of Q^2 on R

This graph shows that optimal value of R by train dataset is 54, but for test it is 45.

References

[1] Andrzej Cichocki. "Higher-Order Partial Least Squares (HOPLS): A Generalized Multi-Linear Regression Method". In: *arXiv:1207.1230v1* 13 (2012).