Canonical Correlation Analysis in Tensor Representation

Abstract

Canonical Correlation Analysis (CCA) is a multivariate statistical method used to identify relationships between two sets of variables. While traditionally applied to matrix data, modern applications often involve high-dimensional structured data, which can be more naturally represented as tensors. This paper presents a mathematical formulation of CCA in the context of tensor representation, offering a framework for exploring correlations between datasets structured as tensors.

1 Introduction

CCA was first introduced by Hotelling in 1936 and has since been widely used in various fields such as signal processing, neuroscience, and machine learning. The goal of CCA is to find linear combinations of variables from two datasets, $\mathbf{X} \in \mathbb{R}^{n\times p}$ and $\mathbf{Y} \in \mathbb{R}^{n\times q}$, such that the correlation between the resulting projections is maximized. Given the growing use of tensor data (multi-dimensional arrays), CCA has been extended to handle tensors in order to exploit the structural information embedded in these multi-way data.

2 Review on Canonical Correlation Analysis

Given two random vectors $\mathrm{x} \in \mathbb{R}^{m_1}, \mathrm{y} \in \mathbb{R}^{m_2}$, a pair of transformations $\mathbf{u}, \mathbf{v}$, called canonical transformations, is found to maximize the correlation of $x' = \mathbf{u}^T\mathbf{x}$ and $y' = \mathbf{v}^T\mathbf{y}$ as

$$\rho = \max_{\mathbf{u},\mathbf{v}} \frac{\hat{\mathbb{E}}\left[x'y'\right]}{\sqrt{\hat{\mathbb{E}}\left[x'^2\right]\hat{\mathbb{E}}\left[y'^2\right]}} = \frac{\mathbf{u}^T\mathbf{C}_{\mathbf{xy}}\mathbf{v}}{\sqrt{\mathbf{u}^T\mathbf{C}_{\mathbf{xx}}\mathbf{u}\mathbf{v}^T\mathbf{C}_{\mathbf{yy}}\mathbf{v}}}, \tag{1}$$

where $\hat{\mathbb{E}}[f]$ denotes empirical expectation of function f and ρ is called the canonical correlation. Multiple canonical correlations $\rho_1, \ldots, \rho_d$, where $d \leq \min(m_1, m_2)$, are defined by the next pairs of $\mathbf{u}, \mathbf{v}$, which are orthogonal to the previous ones. Canonical correlations are affine-invariant to inputs, i.e.,

$\mathbf{Ax}+\mathbf{b}, \mathbf{Cy}+\mathbf{d}$ for arbitrary (nonsingular) $\mathbf{A} \in \mathbb{R}^{m_1 \times m_1}$, $\mathbf{b} \in \mathbb{R}^{m_1}, \mathbf{C} \in \mathbb{R}^{m_2 \times m_2}, \mathbf{d} \in \mathbb{R}^{m_2}$. The proof is straightforward from (1) as $\mathbf{C}_{\mathrm{xy}}, \mathbf{C}_{\mathrm{xx}}, \mathrm{C}_{\mathrm{yy}}$ are covariance matrices and are multiplied by canonical transformations $\mathbf{u}, \mathbf{v}$.

Given two vector sets as matrices $\mathbf{X} \in \mathbb{R}^{N \times m_1}$ and $\mathbf{Y} \in \mathbb{R}^{N \times m_2}$, Goloub's SVD solution is given as follows: If $\mathbf{P}_1, \mathbf{P}_2 \in \mathbb{R}^{N \times d}$ denote two eigenvector matrices of $\mathbf{X}, \mathbf{Y}$, respectively, where $N \gg m_1, m_2 \geq d$, canonical correlations are obtained as singular values of $(\mathbf{P}_1)^T \mathbf{P}_2$ by

$$(\mathbf{P}_1)^T \mathbf{P}_2 = \mathbf{Q}_1 \mathbf{\Lambda} \mathbf{Q}_2^T, \quad \mathbf{\Lambda} = \operatorname{diag}(\rho_1, \ldots \rho_d) \tag{2}$$

3 Multilinear Algebra and Notations

Let $\mathbf{A} = (\mathbf{A})_{ijk} \in \mathbb{R}^{I \times J \times K}$. The inner product of any two tensors is defined as $\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{i,j,k} (\mathbf{A})_{ijk} (\mathbf{B})_{ijk}$. The mode-j vectors are the column vectors of matrix $\mathbf{A}_{(j)} \in \mathbb{R}^{J \times (IK)}$ and the j-mode product of a tensor $\mathbf{A}$ by a matrix $\mathbf{U} \in \mathbb{R}^{N \times J}$ is

$$(\mathbf{B})_{ink} \in \mathbb{R}^{I \times N \times K} = (\mathbf{A} \times_j \mathbf{U})_{ink} = \Sigma_j (\mathcal{A})_{ijk} \mathbf{u}_{nj}. \tag{3}$$

The j-mode product in terms of j-mode vector matrices is $\mathbf{B}_{(j)} = \mathbf{U} \mathbf{A}_{(j)}$.

4 Tensor Representation of Standard CCA

Standard CCA is represented by tensor notations. Given two vector sets as matrices $\mathbf{X} \in \mathbb{R}^{N \times m_1}, \mathbf{Y} \in \mathbb{R}^{N \times m_2}$ $(N \gg m_1, m_2)$, CCA is written as

$$\rho = \max_{\mathbf{u}, \mathbf{v}} \mathbf{x}'^T \mathbf{y}', \quad \text{where } \mathbf{x}' = \mathbf{Xu}, \mathbf{y}' = \mathbf{Yv} \tag{4}$$

Note that the canonical transformations $\mathbf{u}, \mathbf{v}$ are hereinafter defined to be such that $\mathbf{XU} = \mathbf{P}_1 \mathbf{Q}_1, \mathbf{YV} = \mathbf{P}_2 \mathbf{Q}_2$, where $\mathbf{U}, \mathbf{V}$ have $\mathbf{u}, \mathbf{v}$ in their columns, respectively, and $\mathbf{P}, \mathbf{Q}$ are eigenvector and rotating matrices defined in (2), respectively. If we take $\mathbf{X}, \mathbf{Y}$ as second-order tensors $(\mathcal{X})_{ij}, (\mathcal{Y})_{ij}$, the standard CCA is then represented as

$$\rho = \max_{\mathbf{u}, \mathbf{v}} \left\langle \mathcal{X} \times_j \mathbf{u}^T, \mathcal{Y} \times_j \mathbf{v}^T \right\rangle \tag{5}$$

CCA has one shared mode (index i) and mode products by canonical transformations (index j), which is illustrated in Fig. 1. The two data matrices, for which $\mathbf{P}_1, \mathbf{P}_2$ are computed, can be written with respect to the j-mode vector matrices such that $\mathbf{X} = \mathbf{X}_{(j)}^T, \mathbf{Y} = \mathbf{Y}_{(j)}^T$. The j-mode products $\mathcal{X} \times_j \mathbf{U}^T, \mathcal{Y} \times_j \mathbf{V}^T$ in terms of j-mode vector matrices are $\mathbf{U}^T \mathbf{X}_{(j)} = (\mathbf{P}_1 \mathbf{Q}_1)^T, \mathbf{V}^T \mathbf{Y}_{(j)} = (\mathbf{P}_2 \mathbf{Q}_2)^T$, respectively. The canonical transformations are obtained by $\mathbf{U} = \left(\mathbf{X}_{(j)} \mathbf{X}_{(j)}^T \right)^{-1} \mathbf{X}_{(j)} \mathbf{P}_1 \mathbf{Q}_1, \quad \mathbf{V} = \left(\mathbf{Y}_{(j)} \mathbf{Y}_{(j)}^T \right)^{-1} \mathbf{Y}_{(j)} \mathbf{P}_2 \mathbf{Q}_2$.

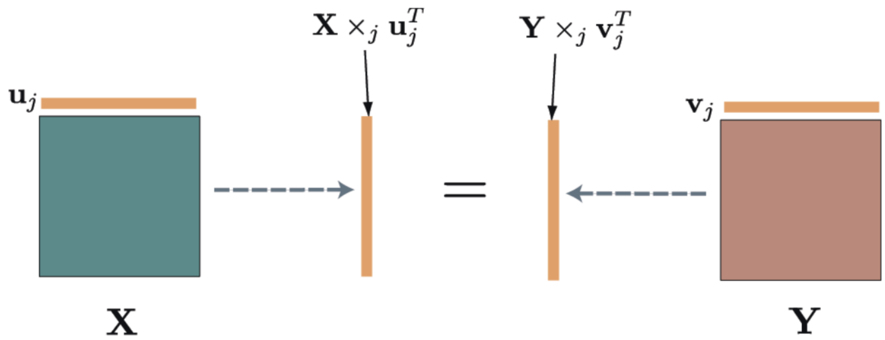

Figure 1: Tensor representation of standard CCA. A pair of canonical transformations, $\mathbf{u}$ and $\mathbf{v}$, are applied to the two data matrices $\mathcal{X}$ and $\mathcal{Y}$ to yield maximally correlated vectors (called canonical vectors).

5 Discussion

CCA in tensor form allows us to generalize the correlations to multi-dimensional data, retaining the intrinsic relationships across multiple modes (dimensions). This makes Tensor CCA particularly powerful for analyzing complex data, such as those encountered in neuroimaging, genomics, and social network analysis. By projecting tensors instead of matrices, Tensor CCA can capture richer interactions that are not available in traditional CCA models.

6 Conclusion

Canonical Correlation Analysis has been extended to tensor data to handle multi-way structures inherent in modern datasets. Tensor CCA offers a more flexible and powerful approach for exploring correlations between higher-order data representations, and it has wide applications in areas involving complex, structured datasets. Further research could focus on computational efficiency and application to specific domains.

References

[1] Fabien Girka et al. *Tensor Generalized Canonical Correlation Analysis*. 2023. arXiv: `2302.05277 [stat.ML]`. URL: `https://arxiv.org/abs/2302.05277`.

[2] Tae-Kyun Kim and Roberto Cipolla. "Canonical Correlation Analysis of Video Volume Tensors for Action Categorization and Detection". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (2009), pp. 1415–1428. URL: `https://api.semanticscholar.org/CorpusID:1073613`.