Multiview Canonical Correlation Analysis (MCCA)

Mansi Verma (0801CS171041)
Monali Koshta (0801CS171044)

December 13, 2020

Contents

Chapter 1

Introduction

1.1 Multiview Learning

Multi-view learning is an emerging direction in machine learning which considers learning with multiple views to improve the generalization performance. To understand the multiple views Multiview dimensionality reduction has become one of the active topic to avoid the curse of dimensionality. As a powerful tool for multimodal feature fusion, canonical correlation analysis (CCA) has received widespread attentions.

1.2 CCA

Canonical correlation analysis (CCA)is the main technique for two-set data dimensionality reduction such that the correlation between the pairwise variables in the common subspace is mutually maximized. CCA was originally proposed by H . Hotelling in 1935. For finding correlation between more than two views CCA fails so in order to solve the problem MCCA is introduced.

1.3 MCCA

Multiview CCA is an extension of CCA from two view scenario to include three or more number of data channels. Defining a measure of cross correlation for more than two random variables is not straightforward and many possible measures have been proposed.

According to the definition of cross-view correlation for multi-view variables, we roughly split the multi-view CCA models into three groups, i.e., 1) the pairwise-correlation based methods 2) the "zero-order-correlation models and 3) the high-order-correlation based methods. Here only pairwise-correlation based method is mentioned.

For the pairwise-correlation-based group, there are still many different ways to measure the correlation degree, such as, by maximizing 1) the sum of all entries in the correlation matrix (SUMCOR), 2) the sum of squares of all entries in the correlation matrix (SSQCOR) and 3) the largest eigenvalue of the correlation matrix (MAXVAR); or by minimizing 4) the smallest eigenvalue of correlation matrix (MINVAR) and 5) the determinant of the correlation matrix (GENVAR).

Considering the objective functions of these models are highly related, this subsection will take only the SUMCOR as an example for elaboration.

The SUMCOR aims to find a projection for each view such that the sum of all possible pairwise correlation, after projecting each view onto the subspace is maximized:

Chapter 2

Mathematical Formulation

2.1 Formulation

The MCCA finds a set of m projections so that the correlation between paired data sets is maximized in the common feature subspace. Given the m sets of variables $X = (X^1, X^2, \cdots, X^m)$, where $X^i = (x_1^i, x_2^i, \ \ , \cdots, x_N^i, \ \ ) \in R^{d_i \times N}$ is the i^{th} view data with the dimensionality of d_i. The SUMCOR aims to find a projection $\mathrm{w_i}$ for each view such that the sum of all possible pairwise correlation, after projecting each view onto the subspace $X^{i'} w_i$, is maximized

$$\arg\max \rho = \sum_{i=1}^{m} \sum_{j=1}^{m} w_i' \, C_{ij} w_j \qquad (2.1)$$

$$s.t. \ w_i' C_{ij} w_j = 1 \, , i = 1, 2, \ldots\ldots, m.$$

Where the C_{ij} denotes the within-view covariance matrix$(i = j)$ or between-view covariance matrix$(i \neq j)$.Solving this problem by Lagrange multiplier technique, we can obtain a generalized multivariate eigenvalue problem (MEP).

Existing study has proven that the MEP problem has no analytical solutions when m ≥ 3. In [11], [14], the approximate solutions are found by resorting to the power iteration method with several restarts the initial vector for approaching to the local optimum.

The second form of SUMCOR reduces the couple constraint as follows

$$\arg\max \rho = \sum_{i=1}^{m} \sum_{j=1}^{m} w_i' \, C_{ij} w_j \qquad (2.2)$$

$$s.t. \sum_{i=1}^{m} w_i' C_{ii} w_i = 1$$

This model can be easily solved by the Lagrange multiplier technique, and it is equivalent to optimizing a generalized eigenvalue problem below

$$\begin{bmatrix} C_{11} & \cdots & C_{1m} \\ \vdots & \ddots & \vdots \\ C_{m1} & \cdots & C_{mm} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} = \lambda \begin{bmatrix} C_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C_{mm} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix} \qquad (2.3)$$

These two MCCA models measure the multi-view correlation degree by defining the sum of all possible pairwise correlation, which is a popular manner but not the straightforward way.

2.2 Universe subspace

MCCA is viewed as learning a common subspace such that correlation between multiple data views is maximized. The common subspaces between m views can be formulated as

$$U = \frac{1}{m} \sum_{v=1}^{m} (X_v w_v) \qquad (2.4)$$

Where, m is total number of views, and U is universe subspace of all views.

Chapter 3

Algorithm

In this section, we give an overview of the MCCA algorithms where we formulate the optimization problem as a standard eigen problem.

3.1 MCCA algorithm

Algorithm 1: MCCA algorithm

Input : Multiview training datasets $X = (X^1, X^2, \cdots, X^m)$, where $X^i = (x_1^i, x_2^i, \quad, \cdots, x_N^i, \ ) \in R^{d_i \times N}$ and m are number of views d_i is dimensionality of X^i.

Output: $w_i \quad s.t. \, i = 1,2, \cdots$m

1. Normalize all the X^i i.e. X^i=$X^i - \overline{X^i}$
2. Calculate within-view co-variance matrix of all X^i

   $C_{ii} = \frac{1}{N} X^i X^{i^T} \ for \ i = 1,2, \ldots . m$

3. Calculate between-view covariance matrix of all pairs

   $C_{ij} = \frac{1}{N} X^i X^{j^T} \ for \ i = 1,2, \ldots . m, j = 1,2, \ldots m$

4. Compute left(A) and right(B) matrix of generalized eigen value problem (GEP): $A \times w = \lambda \times B \times w$

   $$A=\begin{bmatrix} C_{11} & \cdots & C_{1m} \\ \vdots & \ddots & \vdots \\ C_{m1} & \cdots & C_{mm} \end{bmatrix} \qquad B=\begin{bmatrix} C_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C_{mm} \end{bmatrix}$$

5. After putting value of A and B eigen values are obtained.
6. Using eigen values compute eigen vectors which are:
   $w_i \ for \ i = 1,2, \cdots$m

References

- A Survey on Canonical Correlation Analysis - IEEE Journals & Magazine
- https://www.researchgate.net/publication/228836443_Multi-View_Canonical_Correlation_Analysis
- https://arxiv.org/abs/1907.01693