RESEARCH STATEMENT

ZEQUN ZHENG

My research interests lie in tensor decomposition, tensor approximation, machine learning, and data science. Tensors or multidimensional arrays are higher order generalizations of matrices. They are natural structures for expressing data that have inherent higher order structures. Tensor decompositions and Tensor approximations play an important role in learning those hidden structures. They have many applications in machine learning, statistical learning, data science, signal processing, neuroscience, and more. The following is a summary of my research projects during the Ph.D. program.

The project *Higher Order Correlation Analysis for Multi-View Learning*[1] proposes a new method to solve the higher order tensor correlation maximization problem. In the project *Low Rank Tensor Decompositions and Approximations*[2], we study the relation between generating polynomials and tensors, then use generating polynomials to compute tensor decompositions and low rank approximations. *Generating Polynomials and Tensor CP Decompositions*[3] project develops a novel algorithm that detects tensor decompositions when the rank is higher. Our algorithm is the first linear algebra based algorithm that can detect CP decomposition when the tensor's rank is greater than the largest dimension, to the best of the author's knowledge.

1. Higher Order Correlation Analysis For Multi-View Learning

To analyze multi-view or multidimensional data, multi-view learning methods are frequently used in data science. The tensor canonical correlation analysis (TCCA) method is one of the most important multi-view learning methods and it aims at maximizing the higher order tensor correlation. The traditional TCCA method often uses the alternating least square (ALS) method to maximize the higher order tensor correlation. The ALS is convenient for implementation, but its performance is generally not reliable. Therefore, we propose a new method for solving the higher order tensor correlation maximization problem.

Let $\{(\mathbf{y}_{i,1},\dots,\mathbf{y}_{i,m})\}_{i=1}^{N}$ be a multi-view data set, with m views and N points. We are looking for a r-dimensional latent space $\mathbb{R}^r$ such that each $\mathbf{y}_{i,j}$ is projected to $\mathbf{z}_{i,j}\in\mathbb{R}^r$. The projection for the jth view can be represented by a matrix P_j. Then the problem is equivalent to finding optimal projection matrices $P_1,\dots,P_m$ that maximize

$$\rho \coloneqq \sum_{i=1}^{N}\sum_{s=1}^{r}\prod_{j=1}^{m}(P_j^{\mathrm{T}}\mathbf{y}_{i,j})_s. \tag{1.1}$$

After some reformulation and relaxation, we may construct a tensor $\mathcal{M}$ from those data and solve the following optimization,

$$\begin{cases} \min_{\mathbf{u}^{s,j},\lambda_s} & \left\| \mathcal{M} - \sum_{s=1}^r \lambda_s \cdot \mathbf{u}^{s,1} \otimes \cdots \otimes \mathbf{u}^{s,m} \right\|^2, \\ \text{s.t.} & \|\mathbf{u}^{s,j}\|_2 = 1, s = 1, \ldots r, \; j = 1, \ldots, m. \end{cases} \tag{1.2}$$

After the vectors $\mathbf{u}^{s,j}$ are obtained by solving (1.2), the projection matrices P_j can be chosen based on $\mathbf{u}^{s,j}$.

Solving (1.2) is a low rank tensor approximation problem. We propose the following method using generating polynomials to solve for the optimizer.

Algorithm 1.1. (A generating polynomial method for TCCA)

Input: a multi-view data set $\{(\mathbf{y}_{i,1}, \ldots, \mathbf{y}_{i,m})\}_{i=1}^N$ and an approximating rank $r \leq n_1$.

Step 1. Generate the tensor $\mathcal{M} \in \mathbb{R}^{n_1 \times \cdots n_m}$ using the data $\{(\mathbf{y}_{i,1}, \ldots, \mathbf{y}_{i,m})\}_{i=1}^N$.

Step 2. Solve linear least squares for coefficients of the generating polynomials for $\mathcal{M}$.

Step 3. Use the Schur Decomposition to find the roots of the generating polynomials.

Step 4. Use those roots to compute a starting point for (1.2).

Step 5. Compute an improved solution $\mathbf{u}^{s,j}$ of

$$\min_{\mathbf{u}^{s,j} \in \mathbb{R}^{n_j}} \left\| \sum_{s=1}^r \mathbf{u}^{s,1} \otimes \mathbf{u}^{s,2} \otimes \ldots \otimes \mathbf{u}^{s,m} - \mathcal{M} \right\|^2.$$

Step 6. Compute projection matrices $P_j, \ldots, P_m$ based on the improved solution $\mathbf{u}^{s,j}$.

Output: The matrices $P_j, \ldots, P_m$.

We implemented Algorithm 1.1 and ran numerical experiments on two real-world image data sets. We compared our method with Multiset CCA and TCCA using ALS. The computational results show that our proposed method consistently outperforms the prior existing methods.

In conclusion, we proposed a new method for solving the higher order tensor correlation maximization problem. The generating polynomial method is introduced to compute low rank approximating tensors with promising performance from the higher order correlation tensor of multi-view input data. Consequently, the proposed method can achieve better performance than earlier methods based on the ALS.

2. Low Rank Tensor Decompositions And Approximations

Let m and $n_1, \ldots, n_m$ be positive integers. A tensor $\mathcal{F}$ of order m and dimension $n_1 \times \cdots \times n_m$ can be labelled such that

$$\mathcal{F} = (\mathcal{F}_{i_1,\ldots,i_m})_{1 \leq i_1 \leq n_1, \ldots, 1 \leq i_m \leq n_m}.$$

For every tensor $\mathcal{F} \in \mathbb{C}^{n_1 \times \cdots \times n_m}$, there exist vector tuples $(v^{s,1}, \ldots, v^{s,m})$, $s = 1, \ldots, r$, $v^{s,j} \in \mathbb{C}^{n_j}$, such that

$$\mathcal{F} = \sum_{s=1}^r v^{s,1} \otimes \cdots \otimes v^{s,m}. \tag{2.1}$$

The smallest such r is called the rank of $\mathcal{F}$. When r is minimum, the equation is called a rank-r tensor decomposition or CANDECOMP/PARAFAC (CP) tensor decomposition. The low rank tensor approximation (LRTA) problem is to find a low rank tensor that is close to a given one. This is equivalent to solving the following nonlinear least squares optimization

$$\min_{v^{s,j}\in\mathbb{C}^{n_j},\, j=1,\ldots,m} \big\|\mathcal{F} - \sum_{s=1}^{r} v^{s,1}\otimes\cdots\otimes v^{s,m}\big\|^2. \tag{2.2}$$

This work proposes a new method for low rank tensor decompositions and approximations. In this paper, we extend the generating polynomial method to compute tensor rank decompositions and low rank tensor approximations for nonsymmetric tensors.

Without loss of generality, assume the dimensions are decreasing as $n_1 \geq n_2 \geq \cdots \geq n_m$. For an order m tensor whose rank $r \leq n_1$, we show there is a bijection relation between tensor decompositions and generating polynomials. Using this property our algorithm can be applied to find tensor decompositions when the rank is low. In this case, the following theorem guarantees our algorithm produces a tensor decomposition for the tensor

Theorem 2.1. *Suppose $n_1 \geq n_2 \geq \cdots \geq n_m$ and $r \leq \min(n_1, \frac{n_2\cdots n_m}{n_3})$. For a generic tensor $\mathcal{F}$ of rank-r, our algorithm produces a rank-r tensor decomposition for $\mathcal{F}$.*

One advantage of our algorithm is that it only requires linear algebra computations. And our numerical experiment shows its speed is faster than the generalized eigenvalue decomposition (GEVD) method which is a classical one for computing tensor decomposition when the rank $r \leq n_2$.

When the rank $\frac{n_2\cdots n_m}{n_3} \leq r \leq n_1$, finding CP decompositions is harder and we need to apply optimization methods like nonlinear least square to find the generating polynomials.

Based on our algorithm for tensor decomposition, we also propose an algorithm for computing low rank tensor approximations. The error analysis gives the following theorem

Theorem 2.2. *Let $\mathcal{X}^{gp}$ be produced by our algorithm, Suppose the tensor $\mathcal{F}$ has the best (or nearly best) rank-r approximation $\mathcal{X}^{bs}$. Under a generic condition, if the distance $\epsilon = \|\mathcal{F} - \mathcal{X}^{bs}\|$ is sufficiently small, then*

$$\|\mathcal{X}^{bs} - \mathcal{X}^{gp}\| = \mathcal{O}(\epsilon) \quad \text{and} \quad \|\mathcal{F} - \mathcal{X}^{gp}\| = \mathcal{O}(\epsilon).$$

where the constants in the above $\mathcal{O}(\cdot)$ only depend on $\mathcal{F}$.

Theorem 2.2 concludes that if the tensor to be approximated is sufficiently close to a low rank one, then the obtained low rank tensor is a quasi-optimal low rank approximation. Numerical experiments demonstrate the outstanding performance of our algorithm.

3. Generating Polynomials And Tensor CP Decompositions

As mentioned in section 2, detecting tensors' CP decompositions is of great importance and has many applications. For order 3 tensors, when the rank $r > n_2$, it is hard to find tensors' decompositions. For algorithms that are linear algebra based, only a few of them can detect the tensor decompositions when $n_1 \geq r > n_2$.

Those methods require the construction of larger matrices compared with tensors' size. Therefore, the running time is not so satisfying for large tensors. And when $r > n_1$, none of them can detect the tensor decompositions.

In this work, we focus on order-3 tensors and study the tensor decomposition problem especially when the rank $r > n_2$. We propose a novel algorithm to find the CP decompositions by utilizing generating polynomials. Our algorithm successfully find tensors' decompositions when $(\lfloor \frac{n_2}{2} \rfloor n_3) \geq r$ or $\frac{(3n_3-8)n_2}{2n_3-5} \geq r > n_1$ using linear algebra computations only. When the order is greater than 3, we may apply the flattening trick in [2] and apply our algorithm. We have the following result.

Theorem 3.1. *Let $\mathcal{F} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$ be a rank-r tensor.*

(1) If $\frac{\lfloor n_2 \rfloor n_3}{2} \geq r > n_2$, then our algorithm can find the decomposition of $\mathcal{F}$ generically.

(2) If $\frac{(3n_3-8)n_2}{2n_3-5} \geq r > n_1$, under the assumptions of a Proposition in [3], our algorithm finds the decomposition $\mathcal{F}$.

For the $\frac{\lfloor n_2 \rfloor n_3}{2} \geq r$ case, we compare our algorithm with the nonlinear least square method and the state of art algorithm that has the running time $\mathcal{O}(n_1^6 n_2^6 n_3^6)$ in this case. While our method has running time $\mathcal{O}(n_1 n_2^2 n_3^5)$. The numerical experiment also shows the running time difference. Both our algorithm and the state of art algorithm find tensor decompositions successfully. The nonlinear least square method failed to find the correct tensor decompositions.

For the $\frac{(3n_3-8)n_2}{2n_3-5} \geq r > n_1$ case, there is no linear algebra based method before, to the best of the author's knowledge. So we compare our algorithm with the nonlinear least square method. Our algorithm finds tensor decompositions successfully. While the nonlinear least square method failed to find the correct tensor decompositions. Numerical examples successfully demonstrate the robustness and efficiency of our algorithm.

4. Future Work

In the future, I will continue working on machine learning, statistical learning, tensor computation, and applications of tensor computation. Here is a brief plan for my future work.

- High-dimensional tensor-valued data are observed in many fields such as personalized recommendation systems and imaging research. We are interested in studying the estimation and inference of conditional independence structure within tensor data. Applying tensor computation methods to the tensor graphical model and other high dimensional statistics problems is part of my future research plan.
- Tensor computation has gained more and more interest because of its broad applications especially in data science and machine learning. However, there are still lots of unsolved problems in this relatively new research area. In the future, I will continue to work on tensor computation problems, like tensor decomposition, tensor approximation and so on.
- Tensors have broad applications in real-world and data science, including tensor regression, Multi-view learning, tensor neural network, etc. For example, applying CP decomposition to the convolutional kernel of a pre-trained network can speed up deep neural networks. we are also interested

in connecting the tensor decomposition methods with deep tensor canonical analysis. Working on these applications is also part of my future research plan.

References

[1] J. Nie, L. Wang, and Z. Zheng. Higher Order Correlation Analysis for Multi-View Learning, *Pacific Journal of Optimization*, to appear, 2022.

[2] J. Nie, L. Wang, and Z. Zheng. Low Rank Tensor Decompositions And Approximations, *Preprint*, 2022. arXiv:2208.07477.

[3] J. Nie, Z. Yang, and Z. Zheng. Generating Polynomials And Tensor CP Decompositions, In preparation.