CP-ALS-QR report

Alex Zhang

July 2023

1 Introduction

The CANDECOMP/PARAFAC or canonical polyadic (CP) decomposition for multidimensional data, or tensors, is a popular tool for analyzing and interpreting latent patterns that may be present in multidimensional data. Basically CP decomposition of a tensor refers to its expression as a sum of r rank-one components and each of them is a vector outer product. One of the most popular methods used to compute a CP decomposition is the alternating least squares (CP-ALS) approach, which solves a series of linear least squares problems. Usually to solve these linear leaste squares problems, normal equations are used for CP-ALS. This approach may be sensitive for ill-conditioned inputs. Based on this idea, there are already a more stable approach which is solving the linear least sqaures problems using QR decomposition instead.
For my summer research project, I basically follows the QR apprach but trying to improve the efficiency for QR decomposition when assuming the input tensor is in Kruskal structure, that is, a tensor stored as factor matrices and corresponding weights. By exploiting this structure, we improve the computation efficiency by not forming Multi-TTM tensor. The problem left is when doing CP-ALS, QR-based methods is exponential in N, the number of modes. The normal equations approach is linear in N for Kruskal tensor. During the summer I tried to revise and implement former QR method which archieved better stability than normal equations but computation time increases linearly with respect to N.

2 Background

CP Decomposition

Given a d-way tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$, Its CP decomposition of rank $r \in \mathbb{N}$ can be represented as

$$\mathcal{X}(i_1, i_2, \ldots, i_d) \approx \sum_{j=1}^{r} \mathbf{A_1}(i_1, j)\mathbf{A_2}(i_2, j) \ldots \mathbf{A_r}(i_r, j)$$

$$\text{for all } (i_1, i_2, \ldots, i_d) \in [n_1] \otimes [n_2] \otimes [n_3] \otimes \cdots \otimes [n_d]$$

Where $\mathbf{A_k} \in \mathbb{R}^{n_k \times r}$ is a factor matrix for all $k \in [d]$.

CP-ALS

For doing CP decomposition, one method is about working on CP-ALS technique to factorize tensor, which is solving a bunch of linear least squares problems.

Linear Least Square Problem

In mathmatical form, a sample least square problem is like

$$\min_{\mathbf{X}} ||\mathbf{B} - \mathbf{X}\mathbf{A}^\top||_F$$

Solving this least sqaure with QR decomposition will be

$$\begin{aligned}
\mathbf{X}\mathbf{A}^\top &= \mathbf{B} \\
\mathbf{X}(\mathbf{Q}\mathbf{R})^\top &= \mathbf{B} \\
\mathbf{X}\mathbf{R}^\top\mathbf{Q}^\top\mathbf{Q} &= \mathbf{B}\mathbf{Q} \\
\mathbf{X}^\top &= \mathbf{R}^{-\top}\mathbf{B}\mathbf{Q}
\end{aligned}$$

In terms of the work in CP-ALS, we need to solve least square problem in the form

$$\min_{\hat{\mathbf{A}}_n} ||\mathbf{X}_{(\mathbf{n})} - \hat{\mathbf{A}}_\mathbf{n}\mathbf{Z}_n^\top||$$

where $\mathbf{X}_{(\mathbf{n})}$ is the matricized tensors, $\hat{\mathbf{A}}_\mathbf{n}$ is the factor matrix we are about to solve and $\mathbf{Z}_n^\top$ is transpose of the Khatri-Rao product of all factor matrices except n-mode which

$$\mathbf{Z}_n^\top = (\mathbf{A}_N \odot \mathbf{A}_{N-1} \odot \cdots \odot \mathbf{A}_{n+1} \odot \mathbf{A}_{n-1} \odot \cdots \odot \mathbf{A}_1)^\top$$

CP-Rounding

CP-ALS-QR

Result

Conclusion