Supplementary Material - Fast and Accurate PARAFAC2 Decomposition for Time Range Queries on Irregular Tensors

Anonymous Author(s)

ABSTRACT

In this supplementary document, we provide theoretical proofs omitted from the main paper.

1 PROOFS OF LEMMAS AND THEOREMS

We provide proofs of Lemmas and Theorems described in the main document.

1.1 Proof of Lemma 1

Proof. The factor matrix $\tilde{\mathbf{H}}$ is updated by efficiently computing $\tilde{\mathbf{Y}}_{(1)}(\tilde{\mathbf{W}} \odot \tilde{\mathbf{V}})(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{V}}^T\tilde{\mathbf{V}})^{-1}$.

$$\begin{aligned}\tilde{\mathbf{H}} &\leftarrow \tilde{\mathbf{Y}}_{(1)}(\tilde{\mathbf{W}} \odot \tilde{\mathbf{V}})(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{V}}^T\tilde{\mathbf{V}})^{-1} \\ &= \left(\|_{k=S}^{E}\mathbf{F}_k\mathbf{V}_{(b)}^T\right)\left(\vee_{k=S}^{E}(\tilde{\mathbf{W}}(k,:) \odot \tilde{\mathbf{V}})\right)(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{V}}^T\tilde{\mathbf{V}})^{-1}\end{aligned} \tag{1}$$

where $\|$ and $\vee$ are the horizontal and vertical concatenation operations, respectively. With block matrix multiplication (i.e., $\begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}\begin{bmatrix}\mathbf{C}\\ \mathbf{D}\end{bmatrix} = \mathbf{AC} + \mathbf{BD}$), we re-express the above equation similarly to incremental update techniques [1, 2]:

$$\begin{aligned}&\left(\|_{k=S}^{E}\mathbf{F}_k\mathbf{V}_{(b)}^T\right)\left(\vee_{k=S}^{E}(\tilde{\mathbf{W}}(k,:) \odot \tilde{\mathbf{V}})\right)(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{V}}^T\tilde{\mathbf{V}})^{-1} \\ &= \left(\sum_{k=S}^{E}\mathbf{F}_k\mathbf{V}_{(b)}^T(\tilde{\mathbf{W}}(k,:) \odot \tilde{\mathbf{V}})\right)(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{V}}^T\tilde{\mathbf{V}})^{-1}\end{aligned} \tag{2}$$

We obtain Equation (7) of the main paper from Equation (1) of this document using $(\tilde{\mathbf{W}}(k,:) \odot \mathbf{V}_{(b)}^T\tilde{\mathbf{V}})$ instead of $\mathbf{V}_{(b)}^T(\tilde{\mathbf{W}}(k,:) \odot \tilde{\mathbf{V}})$ since $\mathbf{A}(\mathbf{b}^T \odot \mathbf{C})$ is equal to $(\mathbf{b}^T \odot \mathbf{AC})$ where $\mathbf{b}$ is a column vector. □

1.2 Proof of Lemma 2

Proof. The factor matrix $\tilde{\mathbf{V}}$ is updated by $\tilde{\mathbf{Y}}_{(2)}(\tilde{\mathbf{W}} \odot \tilde{\mathbf{H}})(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1}$:

$$\begin{aligned}\tilde{\mathbf{V}} &\leftarrow \tilde{\mathbf{Y}}_{(2)}(\tilde{\mathbf{W}} \odot \tilde{\mathbf{H}})(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1} \\ &= \left(\|_{k=S}^{E}\mathbf{V}_{(b)}\mathbf{F}_k^T\right)\left(\vee_{k=S}^{E}\tilde{\mathbf{W}}(k,:) \odot \tilde{\mathbf{H}}\right)(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1}\end{aligned} \tag{3}$$

where $\tilde{\mathbf{Y}}_{(2)}$ is equal to $\|_{k=S}^{E}\mathbf{V}_{(b)}\mathbf{F}_k^T$. To efficiently update the factor matrix $\tilde{\mathbf{V}}$, we re-express the above term using the block matrix multiplication as follows:

$$\begin{aligned}&\left(\|_{k=S}^{E}\mathbf{V}_{(b)}\mathbf{F}_k^T\right)\left(\vee_{k=S}^{E}\tilde{\mathbf{W}}(k,:) \odot \tilde{\mathbf{H}}\right)(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1} \\ &= \sum_{k=S}^{E}\mathbf{V}_{(b)}\mathbf{F}_k^T\left(\tilde{\mathbf{W}}(k,:) \odot \tilde{\mathbf{H}}\right)\left(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)^{-1}\end{aligned} \tag{4}$$

by which we obtain Equation (8) of the main paper. □

1.3 Proof of Lemma 3

Proof. The factor matrix $\tilde{\mathbf{W}}$ is updated by $\tilde{\mathbf{Y}}_{(3)}(\tilde{\mathbf{V}} \odot \tilde{\mathbf{H}})(\tilde{\mathbf{V}}^T\tilde{\mathbf{V}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1}$:

$$\begin{aligned}\tilde{\mathbf{W}} &\leftarrow \tilde{\mathbf{Y}}_{(3)}(\tilde{\mathbf{V}} \odot \tilde{\mathbf{H}})(\tilde{\mathbf{V}}^T\tilde{\mathbf{V}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1} \\ &= \left(\vee_{k=S}^{E} vec(\mathbf{F}_k\mathbf{V}_{(b)}^T)^T\right)\left(\tilde{\mathbf{V}} \odot \tilde{\mathbf{H}}\right)(\tilde{\mathbf{V}}^T\tilde{\mathbf{V}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1}\end{aligned} \tag{5}$$

where $vec(\mathbf{F}_k\mathbf{V}_{(b)}^T)$ is the vectorization of the matrix $\mathbf{F}_k\mathbf{V}_{(b)}^T$. By utilizing the property (i.e., $vec(\mathbf{AB}) = (\mathbf{B}^T \otimes \mathbf{I})vec(\mathbf{A})$) of the vectorization, we obtain the following equation:

$$\begin{aligned}&\left(\vee_{k=S}^{E} vec(\mathbf{F}_k\mathbf{V}_{(b)}^T)^T\right)\left(\tilde{\mathbf{V}} \odot \tilde{\mathbf{H}}\right)(\tilde{\mathbf{V}}^T\tilde{\mathbf{V}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1} \\ &= \left(\vee_{k=S}^{E} vec(\mathbf{F}_k)^T(\mathbf{V}_{(b)}^T \otimes \mathbf{I})\right)\left(\tilde{\mathbf{V}} \odot \tilde{\mathbf{H}}\right)(\tilde{\mathbf{V}}^T\tilde{\mathbf{V}} * \tilde{\mathbf{H}}^T\tilde{\mathbf{H}})^{-1}\end{aligned} \tag{6}$$

Since we can re-express $(\mathbf{V}_{(b)}^T \otimes \mathbf{I})(\tilde{\mathbf{V}} \odot \tilde{\mathbf{H}})$ as $(\mathbf{V}_{(b)}^T\tilde{\mathbf{V}} \odot \tilde{\mathbf{H}})$ using a mixed product property (i.e., $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \odot \mathbf{D}) = (\mathbf{AC}) \odot (\mathbf{BD})$), we obtain Equation (9) of the main paper from the right term above. □

1.4 Proof of Theorem 1

Proof. Using the Gram-Schmidt process on $\mathbf{Q}_k$, we can find a matrix $\mathbf{R}_k \in \mathbb{R}^{I_k \times (I_k - R)}$ that satisfies $\mathbf{T}_k^T\mathbf{T}_k = \mathbf{T}\mathbf{T}_k^T = \mathbf{I}$, where $\mathbf{T}_k = \begin{bmatrix}\mathbf{Q}_k & \mathbf{R}_k\end{bmatrix}$. In particular, $\mathbf{Q}_k^T\mathbf{Q}_k = \mathbf{I}$ and $\mathbf{R}_k^T\mathbf{Q}_k = \mathbf{O}$. Then, we derive the equation using the unitarily invariant property $\|\mathbf{T}_k^T\mathbf{B}\|_F^2 = \|\mathbf{B}\|_F^2$ and $\tilde{\mathbf{Q}}_k = \mathbf{Q}_k\mathbf{A}_k\mathbf{C}_k^T$ as follows:

$$\begin{aligned}&\sum_{k=S}^{E}\|\mathbf{Q}_k\mathbf{H}_{(b)}\mathbf{S}_k\mathbf{V}_{(b)}^T - \tilde{\mathbf{Q}}_k\tilde{\mathbf{H}}\tilde{\mathbf{S}}_k\tilde{\mathbf{V}}^T\|_F^2 \\ &= \sum_{k=S}^{E}\|\begin{bmatrix}\mathbf{Q}_k^T\\ \mathbf{R}_k^T\end{bmatrix}\mathbf{Q}_k\mathbf{H}_{(b)}\mathbf{S}_k\mathbf{V}_{(b)}^T - \begin{bmatrix}\mathbf{Q}_k^T\\ \mathbf{R}_k^T\end{bmatrix}\tilde{\mathbf{Q}}_k\tilde{\mathbf{H}}\tilde{\mathbf{S}}_k\tilde{\mathbf{V}}^T\|_F^2 \\ &= \sum_{k=S}^{E}\|\mathbf{H}_{(b)}\mathbf{S}_k\mathbf{V}_{(b)}^T - \mathbf{A}_k\mathbf{C}_k^T\tilde{\mathbf{H}}\tilde{\mathbf{S}}_k\tilde{\mathbf{V}}^T\|_F^2\end{aligned} \tag{7}$$

Finally, we obtain Equation (11) of the main paper by re-expressing the above term with the property of the trace, i.e., $\|\mathbf{A} - \mathbf{B}\|_F^2 = tr(\mathbf{A}\mathbf{A}^T) - 2tr(\mathbf{A}\mathbf{B}^T) + tr(\mathbf{B}\mathbf{B}^T)$. □

1.5 Proof of Theorem 2

Proof. We need to perform PARAFAC2 decomposition $\frac{K}{l_b}$ times since there are $\frac{K}{l_b}$ of block tensors. Therefore, the time complexity is $\mathcal{O}(\frac{K}{l_b}T)$. □

1.6 Proof of Theorem 3

Proof. There are four types of preprocessed data: $\mathbf{Q}_k$, $\mathbf{H}_{(b)}$, $\mathbf{W}$, and $\mathbf{V}_{(b)}$. The sizes of all the factor matrices $\mathbf{Q}_k$, $\mathbf{H}_{(b)}$, $\mathbf{W}$, and $\mathbf{V}_{(b)}$

are IKR, $\frac{KR^2}{l_b}$, KR, and $\frac{JKR}{l_b}$, respectively. Therefore, the size of preprocessed result is $\mathcal{O}(IKR + \frac{KR^2}{l_b} + KR + \frac{JKR}{l_b})$. □

1.7 Proof of Theorem 4

Proof. At each iteration, the main components to affect the time complexity are 1) $\mathbf{V}_{(b)}^T\mathbf{V}_{(b)}$ for all b, and 2) L matrix multiplications between matrices of the size $R \times R$. There are $\lceil \frac{L}{l_b} \rceil$ blocks in the query range, thus $\mathcal{O}(\frac{JLR^2}{l_b})$ is required for computing $\mathbf{V}_{(b)}^T\mathbf{V}_{(b)}$. In addition, $\mathcal{O}(LR^3)$ is required for L matrix multiplications between matrices of the size $R \times R$. After the iterations, we need to compute $\tilde{\mathbf{Q}}_k$ by multiplying $\mathbf{Q}_k$ with $\mathbf{A}_k$ and $\mathbf{C}_k^T$, which requires $\mathcal{O}(ILR^2)$. Therefore, the time complexity of the query phase is $\mathcal{O}(NLR^3 + N\frac{JLR^2}{l_b} + ILR^2)$. □

1.8 Proof of Theorem 5

Proof. A collection of $\mathbf{V}_{(b)}$ for all b requires a large space cost $\mathcal{O}(\frac{JLR}{l_b})$ in the query phase. In addition, there are L matrices of size $R \times R$, thereby they require $\mathcal{O}(LR^2)$. Finally, after the iterations, we need to use $\mathbf{Q}_k$ for obtaining $\tilde{\mathbf{Q}}_k$, which requires $\mathcal{O}(ILR)$ for $k = S, ..., E$. Therefore, the space cost that the query phase requires at each iteration is $\mathcal{O}(LR^2 + \frac{JLR}{l_b} + ILR)$. □

REFERENCES

[1] Milos Nikolic, Mohammed Elseidy, and Christoph Koch. 2014. LINVIEW: incremental view maintenance for complex analytical queries. In *International Conference on Management of Data, SIGMOD 2014, Snowbird, UT, USA, June 22-27, 2014*, Curtis E. Dyreson, Feifei Li, and M. Tamer Özsu (Eds.). ACM, 253–264.

[2] Sebastian Schelter. 2020. "Amnesia" - Machine Learning Models That Can Forget User Data Very Fast. In *10th Conference on Innovative Data Systems Research, CIDR 2020, Amsterdam, The Netherlands, January 12-15, 2020, Online Proceedings*. www.cidrdb.org.