Information Retrieval using T-CUR

Matrix to Tensor Transition

Suvendu Kar

Indian Institute of Science Bangalore

March 27, 2024

Contents

1. Introduction
2. Application
3. Advantages over SVD
4. Matrix CUR
5. Previous Methods vs Our Proposed CUR
6. Tensor CUR Decomposition
7. Mathematical Results and Observation
8. Conclusion

Introduction

Introduction

Matrix completion is a powerful tool for data analysis; tensor completion is its higher level analog. In this work, we formulate the three-way tensor's missing data recovery problem as a tensor completion problem. Our suggestion is a new technique for tensor completion that utilizes tensor-CUR decomposition to approximate missing data from small sample sizes. Computational experiments show that compared to other current methods, the suggested strategy performs better.**We provide a theoretical guarantee of how many samples are required for an exact recovery**

INDEX TERMS:

Data recovery, tensor completion, tensor-CUR decomposition

Application

Application

- Image Analysis[1,2]
- Recommender System[3,4]
- wireless spectrum map construction[5]
- seismology signal processing[6]
- computer vision[7]

In this work, we address the issue of missing data recovery of a 3-way tensor and define it as a tensor completeness problem, drawing inspiration from several effective applications of the tensor pattern.**But why CUR?What are the draw back in previously implemented methods for decomposition?**:CP[8−10] based decomposition requires to know (CP rank:The CP rank is defined by the minimum number of rank-one terms in CP decomposition) rank beforehand (calculating rank is NP complete[11]).Moreover, the best rank-K approximation to a tensor may not always exist in the CP approach. The tensor-SVD[12] algorithm is computationally expensive, since it requires computing the SVD decomposition of the approximate matrix at each iteration of the optimization.

Advantages over SVD

Advantages over SVD

Our suggested method, which computes the low rank approximation of a given tensor using the tensor's actual rows and columns, extends matrix-CUR decomposition to tensor. Because it simply has to resolve a typical regression problem, it is computationally efficient. Furthermore, our suggested method may effectively compute the low rank approximation of a given tensor utilizing the tensor's real rows and columns without requiring knowledge of the rank beforehand.

Matrix CUR

Matrix CUR[13,14]

RESULT1
Let the matrix A $\in F^{n1*n2}$ have rank(A)= r .Let I $\subset \{1{:}n1\}$ and J $\subset \{1{:}n2\}$ satisfying $|I| \geq r, |J| \geq r$. Let matrices C=A(:,J),R=(I,:),U=A(I,J),if rank(U)=rank(A),then A=$CU^{\dagger}R$.
THEOREM[15] **(Mahone and Drineas):**
CUR in O(n1*n2) time achieves $||A - CUR||_F \leq ||A - A_k||_F + \epsilon||A||_F$ with probability at least (1-δ), by picking O(klog($1/\delta$)/ϵ^2) columns , and O($\frac{k^2 log^3(1/\delta)}{\epsilon^6}$)
RESULT[16]:
Select c=O(klogk/ϵ^2) number of columns of Ausing column select algorithm[16], and select r=O(klogk/ϵ^2) rows of A using row select algorithm[16].Set U=$W^{\dagger}$, W being intersection of r rows and c columns.Then , we have $||A - CUR||_F \leq ||A - A_k||_F * (2 + \epsilon)$, with probability 98%.

Previous Methods vs Our Proposed CUR

CP vs CUR

Suppose we sample a 3-way tensor $A \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ at the set of indices in a set Ω. Let P_Ω denotes the sampling operator

$$P_\Omega : \mathbb{R}^{I_1 \times I_2 \times I_3} \to \mathbb{R}^{I_1 \times I_2 \times I_3},$$

which is defined by

$$P_\Omega(A)_{ijk} = \begin{cases} A_{ijk} & (i,j,k) \in \Omega \\ 0 & otherwise \end{cases}.$$

Jain and Oh [17] tried to complete the tensor by solving the following optimization problem

$$minimize_{\hat{A}, rank(\hat{A})=r} ||P_\Omega(A - \Sigma_{l=1}^{r} \sigma_l (a_l^{(1)} \circ a_l^{(2)} \circ a_l^{(3)}))||_F^2.$$

They showed that under certain standard assumptions, the proposed method can recover a three-mode $n \times n \times n$ dimensional rank-r tensor exactly from $O(n^{3/2} \cdot r^5 log^4 n)$ randomly sampled entries.But knowing r is NP complete.
Need more clarity?Look here!

Tucker vs CUR

Liu et al. [7] proposed a tensor completion algorithm based on minimizing tensor n-rank in Tucker decomposition format. It uses the matrix nuclear norm instead of matrix rank, and try to solve the convex problem as follows

$$\min_{\mathscr{X}} \sum_{i=1}^{n} \alpha_i \left\| X_{(i)} \right\|_*$$
$$\text{subject to } P_\Omega(\mathscr{X}) = P_\Omega(\mathscr{A}),$$

where $X_{(i)}$ is the mode- n matricization of $\mathscr{X}$ and $\alpha_{(i)}$ are prespecified constants satisfying $\alpha_{(i)} \geq 0, \sum_{i=1}^{n} \alpha_{(i)} = 1$.
Unlike the optimal dimensionality reduction in matrix PCA which can be obtained by truncating the SVD, there is no trivial multi-linear counterpart to dimensionality reduction for Tucker type. Alternatively, suitable choices of the truncation values of n-rank are not likely to be known a priori [18]. As a contrast, our proposed algorithm use the actual rows and columns of the tensor to efficiently compute the low rank approximation of a given tensor, without having to know the rank a priori.

SVD vs CUR

The tensor tubal rank is used in tensor-SVD, also referred to as tensor rank. Zhang et al. tried to complete the tensor by solving the following convex optimization problem

$$\min_{\mathscr{X}} \|\mathscr{X}\|_{TNN}$$
$$\text{subject to } P_\Omega(\mathscr{X}) = P_\Omega(\mathscr{A}),$$

where the tensor nuclear norm is taken as the convex relaxation of tensor tubal rank, $\| * \|_{TNN}$ is the tensor nuclear norm. Zhang et al. showed that one can perfectly recover a tensor of size $I_1 \times I_2 \times I_3$ with rank r under tensor-SVD as long as $O\left(rI_1I_3 \log\left(\left(I_1 + I_2\right) I_3\right)\right)$ samples are observed.

Tensor CUR Decomposition

Problem Formulation

Definition (Tensor CUR Decomposition (T-CUR)): For a 3-way tensor $\mathscr{A} \in R^{I_1 \times I_2 \times I_3}$, the tensor-CUR decomposition is given by

$$\mathscr{A} = \mathscr{C} * \mathscr{U} * \mathscr{R},$$

where $\mathscr{C}, \mathscr{U}$ and $\mathscr{R}$ are tensors of size $I_1 \times I_1 \times I_3, I_1 \times I_2 \times I_3$ and $I_2 \times I_2 \times I_3$ respectively.

The missing data recovery problem can be formulated as a tensor completion problem with the goal of finding its missing entries through the following optimization

$$\min_{\mathscr{U} \in R^{I_1 \times I_2 \times I_3}} \frac{1}{2} \left\| P_\Omega(\mathscr{Y} - \mathscr{C} * \mathscr{U} * \mathscr{R}) \right\|_F^2$$

$$\text{subject to } P_\Omega(\mathscr{Y}) = P_\Omega(\mathscr{A})$$

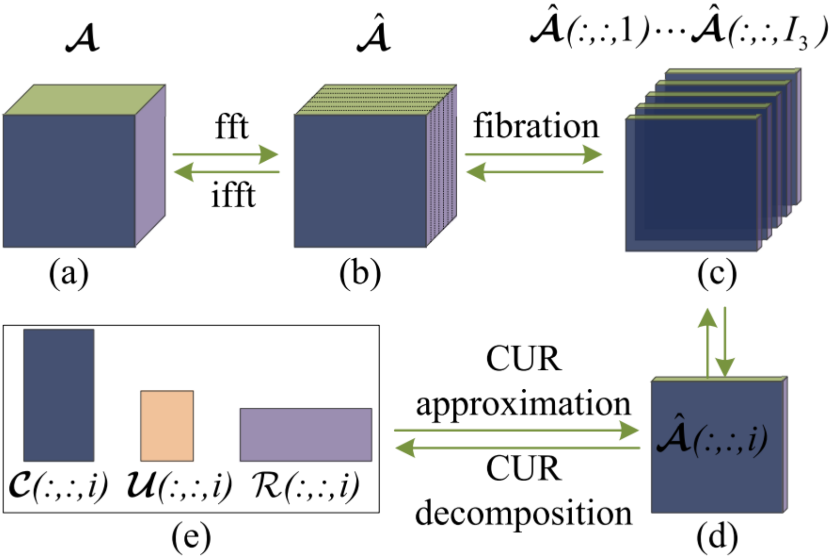
Used Algorithm:Really?

Algorithm : Tensor-CUR Approximation From Partially Observed Entries (T-CUR)
Require: Observation data $P_{\Omega}(\mathscr{A}) \in R^{I_1 \times I_2 \times I_3}$, vector of sample columns/rows number $d/l \in R^{1 \times I_3}$
1. $\hat{\mathscr{A}} = FFT\left(P_{\Omega}(\mathscr{A}), [], 3\right);$
2. for $i \leftarrow 1, \ldots, I_3$ do
3. $\left[\hat{\mathscr{C}}^{(i)}, \hat{\mathscr{U}}^{(i)}, \hat{\mathscr{R}}^{(i)}\right] = \text{M-CUR}\left(P_{\Omega}\left(\hat{\mathscr{A}}^{(i)}\right), d_i, l_i\right)$
4.end for
5.$\mathscr{C} = \mathrm{IFFT}(\hat{\mathscr{C}}, [], 3); \mathscr{U} = \mathrm{IFFT}(\hat{\mathscr{U}}, [], 3); \mathscr{R} = \mathrm{IFFT}\ (\ \hat{\mathscr{R}}, [], 3)$
6.Return a tensor cur approximation of $\mathscr{A}$: $\mathscr{Y} = \mathscr{C} * \mathscr{U} * \mathscr{R} \in R^{I_1 \times I_2 \times I_3}$

Actual Algorithm

```
function [C, U, R] = TENSOR_CUR(A, d1, d2)
    % Input: A - a tensor of size n1 x n2 x n3
    %        d1 - sample size for rows
    %        d2 - sample size for columns

    A_tilde = fft(A,[],3);

    % Sample row index I and column index J
    %I = randsample(size(A,1), d1);
    %J = randsample(size(A,2), d2);

    C = zeros([size(A,1), d2, size(A,3)]);
    R = zeros([d1, size(A,2), size(A,3)]);
    U = zeros([d2,d1,size(A_tilde ,3)]);

    for k= 1:size(A_tilde ,3)
        C(:,:,k) = A_tilde(:,1:d2,k);
        R(:,:,k) = A_tilde(1:d1,:,k);
        U(:,:,k) = MAT_PSEUDOINV(A_tilde(1:d1,1:d2,k));
    end

    C = ifft(C,[],3);
    R = ifft(R,[],3);
    U = ifft(U,[],3);
end
```

T-CUR



$\mathcal{A}$
$\hat{\mathcal{A}}$
$\hat{\mathcal{A}}(:,:,1)\cdots\hat{\mathcal{A}}(:,:,I_3)$
fft
ifft
fibration
(a)
(b)
(c)
CUR approximation
CUR decomposition
$\hat{\mathcal{A}}(:,:,i)$
$\mathcal{C}(:,:,i)$
$\mathcal{U}(:,:,i)$
$\mathcal{R}(:,:,i)$
(e)
(d)

Mathematical Results and Observation

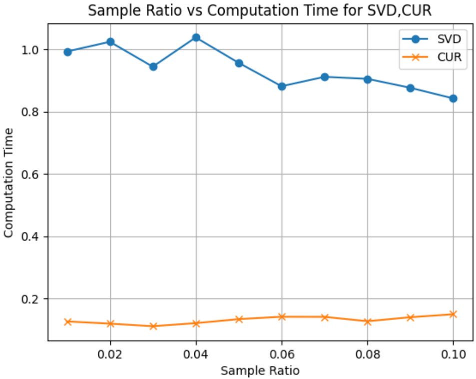
Mathematical Results

1)CUR decomposition of a real valued tensor always exists.(**why?**Link to existence of CUR in matrix)
2)Hey!!What accuracy metric was used?

Observation

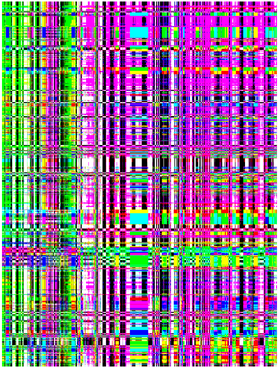
» For a certain Tensor SVD took time of 2.074059 sec, where CUR took time of 1.144329 sec, which indeed tells that CUR is comparatively less computation expensive.

» We created a 1000*1000*3 tensor with each frontal slice is diagonal as got the following:[svd took 10X more time for such small tensor size!!]



Observations

(a) Actual Image



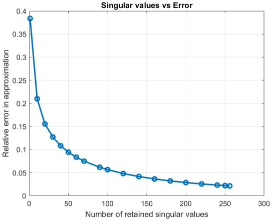
(b) CUR Recovery with 10 rows and columns

(c) SVD Recovery with 10 singular values

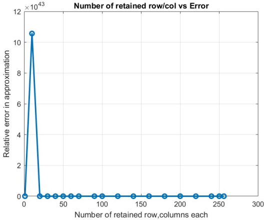
Figure: Is SVD better than CUR?

Observations

(a) Actual Image



(b) Error in SVD based recovery

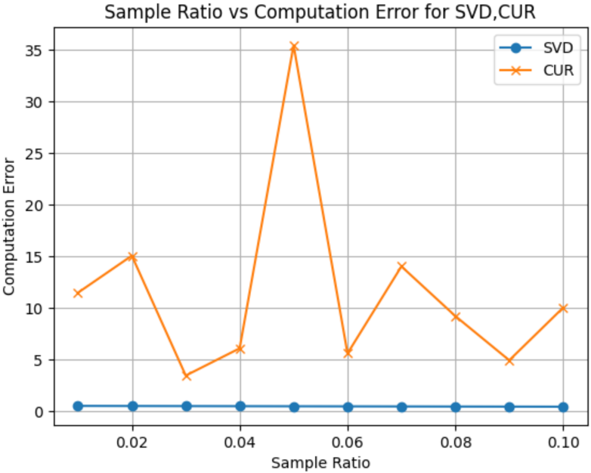


(c) Error in CUR based recovery.Clearly the maximal rank of the image is <20

Figure: Is SVD better than CUR?

Observation

For a fixed tensor A(saved for regeneration purpose) of size 1000X1000X3 with full maximal rank, we got the error as:



Conclusion

Conclusion

The missing data recovery problem of a 3-way tensor was defined as a tensor completion problem in this study.Our suggestion was a new tensor completion approach that could efficiently retrieve the absent data from small samples.It's a novel approach of turning the matrix-CUR into a three-way tensor. This means that one may now express a 3-way tensor as a product of other 3-way tensors.Using our own created tensors as test subjects, we discovered that although CUR is time-efficient, it is not always possible to find a smaller error (when rank of W, which is the intersection of C,R is less than true rank of tensor).
Why CUR is giving higher error?:Let's efficiently choose rows and columns(random sampling with clever utilization of sampling,ket's not fix row and column as same.
Which is a good scope of future experiment.TOWARDS END-TERM I WILL EXPLORE DIFFERENT SAMPLING TECHNIQUES INCLUDES TERM SAMPLING WHICH PROVIDES THEORETICAL GURANTEE OF MORE EFFICIENT APPROXIMATION, WHICH SURELY WILL BE A NICE RESEARCH WORK TOWARDS PUBLICATION.

Thank you!

Reference

1[13])J. A. Bengua, H. N. Phiem, H. D. Tuan, and M. N. Do, ‘‘Efficient tensor completion for color image and video recovery: Low-rank tensor train,'' IEEE Trans. Image Process., vol. 26, no. 5, pp. 2466–2479, May 2017.
2[14])] N. Li and B. Li, ‘‘Tensor completion for on-board compression of hyperspectral images,'' in Proc. IEEE Int. Conf. Image Process., Sep. 2010, pp. 517–520.
3[15]) A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, ‘‘Multiverse recommendation: N-dimensional tensor factorization for context-aware collaborative filtering,'' in Proc. ACM Conf. Recommender Syst. (RecSys), Barcelona, Spain, Sep. 2010, pp. 79–86.
4[16])W. W. Sun, J. Lu, H. Liu, and G. Cheng, ‘‘Provable sparse tensor decomposition,'' J. Roy. Stat. Soc., vol. 79, no. 3, pp. 899–916, 2016.
5[17])M. Tang, G. Ding, Q. Wu, Z. Xue, and T. A. Tsiftsis, ‘‘A joint tensor completion and prediction scheme for multi-dimensional spectrum map construction,'' IEEE Access, vol. 4, no. 99, pp. 8044–8052, 2016.
6[18])Y. Zhang, C. D. Silva, R. Kumar, and F. J. Herrmann, ‘‘Massive 3D seismic data compression and inversion with hierarchical tucker,'' in Proc. SEG, 2017, pp. 1347–1352.

Reference

7[19])J. Liu, P. Musialski, P. Wonka, and J. Ye, ‘‘Tensor completion for estimating missing values in visual data,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 1, pp. 208–220, Jan. 2013.

8[20])B. Ran, H. Tan, Y. Wu, and P. J. Jin, ‘‘Tensor based missing traffic data completion with spatial–temporal correlation,’’ Phys. A, Stat. Mech. Appl., vol. 446, pp. 54–63, Mar. 2016.

9[21])E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, ‘‘Scalable tensor factorizations for incomplete data,’’ Chemometrics Intell. Lab. Syst., vol. 106, no. 1, pp. 41–56, 2011.

10[22]) H. Tan, G. Feng, J. Feng, W. Wang, Y.-J. Zhang, and F. Li, ‘‘A tensorbased method for missing traffic data completion,’’ Transp. Res. C, Emerg. Technol., vol. 28, pp. 15–27, Mar. 2013.

11[23])J. Håstad, ‘‘Tensor rank is NP-complete,’’ J. Algorithms, vol. 11, no. 4, pp. 644–654, 1990.

12[25])Z. Zhang and S. Aeron, ‘‘Exact tensor completion using t-SVD,’’ IEEE Trans. Signal Process., vol. 65, no. 6, pp. 1511–1526, Mar. 2017.

Reference

13) Cai, H., Hamm, K., Huang, L., Li, J., Wang, T. (2021). Rapid robust principal component analysis: CUR accelerated inexact low rank estimation. IEEE Signal Process. Lett. 28:116–120. DOI: 10.1109/LSP.2020.3044130.
14) Hamm, K., Huang, L. (2020). Perspectives on CUR decompositions. Appl. Comput. Harmon. Anal. 48(3):1088–1099. DOI: 10.1016/j.acha.2019.08.006.
15)CUR Decomposition
16)Advanced CUR Decomposition
17[30])P. Jain and S. Oh, ''Provable tensor factorization with missing data,'' in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 1431–1439.
18[24])N. Hao, M. E. Kilmer, K. Braman, and R. C. Hoover, ''Facial recognition using tensor-tensor decompositions,'' SIAM J. Imag. Sci., vol. 6, no. 3, pp. 437–463, 2013.