IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-7, Special Issue-10, May 2019 E-ISSN: 2347-2693

A Review on Various Matrix Factorizaton Techniques

R. Mishra[1*], S. Choudhary[2]

Computer Science and Engineering Shri Ram Group of Institution Jabalpur, M.P., India

Corresponding Author: kanakmishra2011@gmail.com, Tel.: +91-7415422987

DOI: https://doi.org/10.26438/ijcse/v7si10.1315 | Available online at: www.ijcseonline.org

Abstract— In this work, we give the related work of fundamental matrix decomposition techniques. The primary strategy that we talk about is known as Eigen value decomposition, which breaks down the underlying matrix into an authoritative shape. The second strategy is nonnegative matrix factorization (NMF), which factorizes the underlying grid into two littler matrixes with the imperative that every component of the factorized matrix ought to be nonnegative. The third strategy is singular value decomposition (SVD) that utilizations particular estimations of the underlying network to factorize it. The last technique is CUR decomposition, which faces the issue of high thickness in factorized matrixes (an issue that is confronted when utilizing the SVD strategy). This work concludes with a description of other state-of-the-art matrix decomposition techniques

Keywords—Matrix Factorization, Non Negative Matrix Factorization, Singular Value Decomposition

I. INTRODUCTION

Numerous issues throughout our life are spoken to by matrices. Because of the high dimensionality of information in these issues, the initial matrix is generally factorized into at least two "littler" matrices. These matrices have the upside of littler measurements, bringing about diminished required storage room and less required time for handling them.

Consequently, they can be processed more productively by calculations than the initial matrix. There are numerous strategies [1-3] on the best way to break down a matrix and manage a high-dimensional informational collection. Principal component analysis (PCA) is an information mining system that replaces the high-dimensional unique information by its projection onto the most critical axes. This procedure maps directly the information to a lower dimensional matrix such that the difference of information in the low-dimensional portrayal is maximized. It is a straightforward strategy, which depends on eigen-values and the eigenvectors of a matrix.

Let A be a square (N×N) matrix with N linearly independent eigenvectors, $q_i, (i = 1,2,3,..., N)$. Then A can be factorized as:

$$A = Q\Lambda Q^{-1}, \quad (1)$$

where Q is the square (N×N) matrix whose ith column is the eigenvector q_i of A and Λ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues, i.e., $\Lambda_{ii} = \lambda_i$. Note that only diagonalizable matrices can be factorized in this way. For example, the defective matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ cannot be diagonalized.

In the following example we are going to show the importance of matrix factorization. Consider a situation where we have a mapping and users and items in the form of matrix. As the number of user increases then the size of the matrix also increases. However with the help of suitable matrix decomposition techniques we can divide the matrix into sub-matrix of smaller size and thus we can save the space for further processing. Fig. 1 shows an example of matrix decomposition.

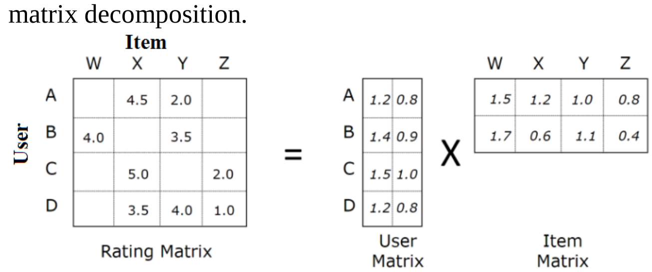

Fig. 1: an example of matrix decomposition.

II. THE MATRIX FACTORIZATION TECHNIQUES

A. Eigen Value Decomposition (EVD)

In this segment, we display a decomposition technique known as eigen-value decomposition. In linear algebra based math, the eigen-value decomposition strategy is the factorization of a matrix A into a canonical form. The eigen-values and the eigenvectors of A are utilized to speak to the matrix. To apply this technique to a matrix, the framework ought be square as well as diagonalizable [4-6]. **Fig. 2 shows an example of EVD.**

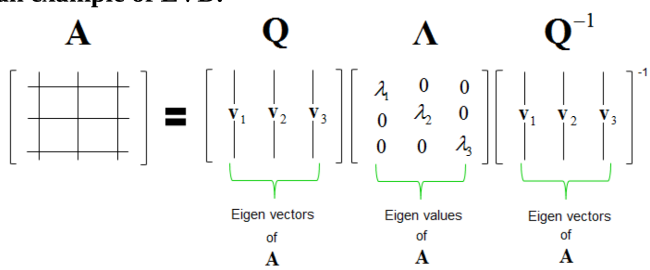

Fig. 2: an example of matrix EVD.

B. Non-Negative Matrix Factorization (NMF)
Another broadly known strategy in dimensionality decrease and information investigation is nonnegative matrix factorization (NMF). In this segment, we will examine how the NMF calculation functions and apply it to the training information of the running illustration. The NMF calculation factorizes a matrix A of every two frameworks U and V, with the property that each of the three matrix have no negative components. This non-negativity makes coming about grids more appropriate for the clustering of objects.

Let matrix V be the product of the matrices W and H,

$$V = W.H \tag{2}$$

Matrix multiplication can be implemented as computing the column vectors of V as linear combinations of the column vectors in W using coefficients supplied by columns of H. That is, each column of V can be computed as follows:

$$v_i = W.h_i \tag{3}$$

C. Singular Value Decomposition (SVD)
Singular value decomposition (SVD) is a critical linear algebra tool that we use to solve many mathematical problems. The SVD strategy is a factorization of a genuine or complex matrix [4]. In this segment, we exhibit the numerical plan of SVD and a portion of its varieties.

Suppose M is a m×n matrix whose entries come from the field K, which is either the field of real numbers or the field of complex numbers. Then there exists a factorization, called a 'singular value decomposition' of M, of the form:

$$M = U.\Sigma.V^* \tag{4}$$

where

- U is an m × m unitary matrix over K (if K=R), unitary matrices are orthogonal matrices),
- Σ is a diagonal m × n matrix with non-negative real numbers on the diagonal,
- V is an n × n unitary matrix over K, and V∗ is the conjugate transpose of V.

As a matter of first importance, we give SVD's association the eigen-value disintegration strategy. The connection amongst SVD and the eigen-value disintegration technique originates from an exceptional instance of the last mentioned one which will be examined along these lines. **Fig. 3 shows an example of SVD.**

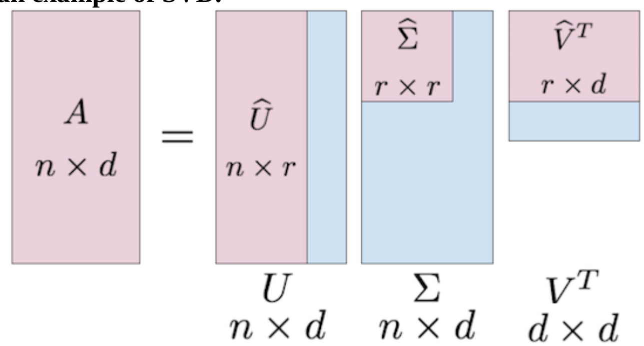

Fig. 3: an example of matrix SVD.

D. CUR Matrix Decomposition
In this section, we are going to present another matrix decomposition method, which is known as CUR matrix decomposition, because the initial matrix A is factorized to three matrices (C, U, and R). In high-dimensional data sets, several matrix decomposition methods, such as the SVD method, produce decomposed matrices which tend to be very dense, a fact that makes their processing a challenge in terms of efficiency.

The CUR matrix approximation is not unique and there are multiple algorithms for computing one. Hence we are not showing its mathematical equations here.

In contrast, the CUR decomposition method confronts this problem as it decomposes an original matrix into two sparse matrices C and R and only one dense matrix U, whose size is quite small. Moreover, CUR gives an exact decomposition no matter how many dimensions picked from the origin matrix (i.e., how big is parameter c), whereas in SVD, the parameter c should be at least equal to the rank of the origin matrix A [5-7]. **Fig. 4 shows an example of CUR matrix decomposition.**

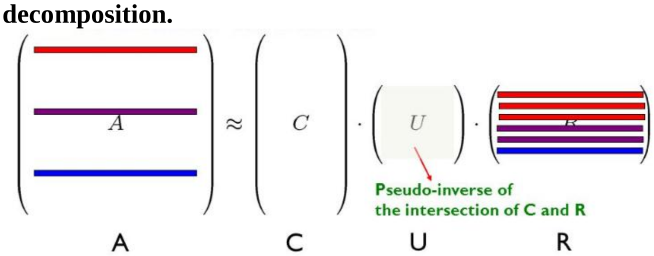

Fig. 4 shows an example of CUR matrix decomposition.

III. CONCLUSION AND FUTURE SCOPE

The matrix factorization technique is used in many party of computer science applications. The graph is most powerful data structure in computer science, and it is represented in the form of matrix. This makes the matrix as the most powerful data representation form. Thought for proposed use we are required to decompose the matrix into sub-matrix of reduced size. In this work we have reviewed few basic matrix decomposition techniques. Initially we have discussed about the need of matrix factorization. Later we have discussed four widely used matrix decomposition techniques. We have also discussed the advantages and disadvantages of these methods.

REFERENCES

[1] Pudil, P., Novoviˇ cová, J.: Novel methods for feature subset selection with respect to problem knowledge. In: Feature Extraction. Construction and Selection. Springer International Series in Engineering and Computer Science, vol. 453, pp. 101–116. Springer, US (1998).

[2] Anand Rajaraman and Jeffrey David Ullman: Mining of Massive Datasets. Cambridge University Press, New York, NY, USA (2011).

[3] Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science 290(5500), 2323–2326 (2000).

[4] Bensmail, H., Celeux, G.: Regularized gaussian discriminant analysis through eigenvalue decomposition. J. Am. Stat. Assoc. 91(436), 1743–1748 (1996).

[5] Drineas, P., Kannan, R., Mahoney, M.W.: Fast monte carlo algorithms for matrices III: computing a compressed approximate matrix decomposition. SIAM J. Comput. 36(1), 184–206 (2006).

[6] Mahoney, M.W., Maggioni, M., Drineas, P.: Tensor-cur decompositions for tensor-based data. SIAM J. Matrix Anal. Appl. 30(3), 957–987 (2008).

[7] Anand Rajaraman and Jeffrey David Ullman: Mining of Massive Datasets. Cambridge University Press, New York, NY, USA (2011).

Authors Profile

Miss. Roshni Mishra pursed Bachelor of Engineering from Computer science branch RGPV Bhopal in 2014 .She is currently pursuing Master of Technology (M.Tech) from Computer Science branch . This is her first paper to be published in reputed journals online. Her main research work focuses on Data Sharing Over Cloud Using Matrix Factorization Techniques .

Mrs Sapna Choudhary is currently working as Assistant Professor in Department of Computer Science , Head Of The Department M.Tech Computers Science field in Shri Ram Group of Institution. She has many years of teaching experience .
Email id : choudharysapnajain@gmail.com

