Compendium of Computer Science
ISSN Print: 3079-5362
ISSN Online: 3079-5370
https://doi.org/10.63313/CS.8008

Convex Nonnegative Matrix Factorization on Bipartite Graph

Pengfei Zhang
College of Computer Science and Technology, Qingdao University, Qingdao 266071, China
Email: 17667540383@126.com

How to cite this paper: Zhang, P. F. (2025). Convex Nonnegative Matrix Factorization on Bipartite Graph. Compendium of Computer Science, 1(1), 96–105. ISSN Print: 3079-5362, ISSN Online: 3079-5370.
https://doi.org/10.63313/CS.8008
Published: 2025-05-07

Abstract

In this paper, we propose a novel convex nonnegative matrix factorization (CNMF) method to learn a bipartite graph with exactly k connected components, where k is the number of clusters. The new bipartite graph learned in our model approximates the original graph but maintains an explicit cluster structure, from which we can immediately get the clustering results. Besides, inspired by self-expressive assumption of the data and subspace clustering, we added additional constraints to further recover global structural information of the data. Multiplicative updating rules are developed. The superiority of our method over the effectiveness and efficiency is demonstrated by extensive experiments.

Keywords

Nonnegative Matrix Factorization; Clustering; Bipartite Graph

1. Introduction

In recent years, non-negative matrix factorization (NMF) as an effective matrix decomposition method has attracted wide attention. The basic idea of NMF method is to decompose the original data matrix into two non-negative matrices, the basis matrix and the coefficient matrix, so that their product approximates the original data matrix, the column vector in the original data matrix can be interpreted as the weighted sum of all the columns in the basis matrix, and the items in the coefficient matrix correspond to the weight coefficient. This combination method conforms to the concept of ”part constitutes a whole” in the general thinking of human beings[1, 2], and the non-negative constraint is interpretable, and has the advantages of simple implementation and small storage space. Therefore, the study of non-negative matrix factorization has certain practical significance. In order to improve the efficiency and recognition rate of NMF, many scholars at home and abroad have conducted in-depth research on NMF algorithm. Liu etal. designed a constrained non-negative matrix factorization algorithm (CNMF)[3], which uses known label information to improve the efficiency of matrix factorization. In order to ensure the sparsity of data structure, many researchers put forward correlation models. In order to enhance and ensure the sparsity of data structures, Hoyer etal. combined the sparsity constraint l1 and l2 norm on the basis of the non-negative matrix decomposition model. The sparse Nonnegative moment matrix (SNMF) is proposed[4]. Ding etal., in order to generate independent clusters, applied orthogonal constraints to the basis matrix or low- dimensional representation matrix in the non-negative matrix decomposition model. Orthogonal Nonnegative Matrix Factorization (ONMF) is proposed[5]. Cai etal. first assumed the local invariance of the data, modeled directly on the manifold structure of the data, and constructed the nearest neighbor plot on the scattered points of the data. Graph regularized Nonnegative Matrix Factorization (GNMF)[6]. Li etal. used the low-rank decomposition algorithm to decompose the low-rank effective part of the original image data and obtain the feature representation of the data. Graph Regularized Nonnegative Low-Rank Matrix Factorization, (GNLNMF)[7]. Peng etal. propose a new NMF method with log-norm imposed on the factor matrices to enhance sparseness and robustness(RLS-NMF)[8]. Besides, To learn more discriminative representations, extend the NMF to learn robust representations. Li etal. proposed robust structured non-negative matrix factorization by explicit block diagonal structure and l2,p norm(RSNMF)[9]. Huang etal. propose a robust non-negative matrix factorization(RNMF) with structural regularization to explore the global and local structure of the data[10]. Peng etal. proposed a non-negative matrix factorization method with local similarity learning(KLS-NMF)[11], which learns local similarity and clustering in a mutually enhanced way, learning both the global and local structure of the data. Deng etal. proposed a graph regularized sparse non-negative matrix factorization(GSNMF) for clustering[12], in this model, l1 norm is added to the low-dimensional matrix to obtain clearer data to approximate the high-dimensional matrix, and the influence of noise introduced into the data reconstruction due to the noise sensitivity of square-loss method is solved. Kong etal.[13] proposed non-negative matrix factorization based on l2,1 norm, which replaced l2 in the standard NMF loss function to solve the noise problem, and the loss function based on the l2,1 norm has good robustness to noise and outliers[14].

Although these methods have improved efficiency, they are aimed at sparsity, low rank, and robustness, but do not consider a clearer clustering structure. Encouraged by bipartite graph partition[15] and ridge regression[16], we apply bipartite graph to non-negative matrix decomposition and use F-form to ensure sparsity. The new bipartite graph learned in our model approximates the original graph but maintains an explicit cluster structure, to achieve such an ideal structure of the new bipartite graph, we impose constraints on the rank of its normalized Laplacian matrix and derive algorithms to optimize the objective. We conduct several experiments to evaluate the effectiveness and robustness of our mode.

In summary, the key contributions of this letter are as follows:

(1) We propose a novel nonnegative matrix factorization method. Different from existing methods, we target to learn a new bipartite graph similarity matrix that has exact k connected components. Obviously such a new graph can be considered as the ideal graph for clustering task with providing clear clustering structure.
(2) In order to directly connects the problem of NMF with clustering, we impose an orthogonality constraint of coefficient matrix. Besides, we impose rank constraint to recover global structural information of the data.
(3) An efficient optimization strategy are designed and comprehensive theoretical analysis is provided to guarantee the convergence.
(4) Experimental results based on four datasets and multiple comparison methods have demonstrated the effectiveness and efficiency of the model.

2. Related work

2.1. NMF

Given the nonnegative data $X = [x_1, x_2, \dots, x_n] \in \mathcal{R}^{d\times n}$, d and n being the dimension and sample size, respectively, NMF is to factor the nonnegative matrix X into $U \in \mathcal{R}^{d\times k}$(basis vectors) and $V \in \mathcal{R}^{n\times k}$(coefficient matrix) with k being a proper value. The NMF problem can be formally represented as follows:

$$\min_{U\geq 0, V\geq 0} \left\| X - UV^T \right\|_F^2, \tag{2.1}$$

where $||\cdot||_F$ is the Frobenius norm. Usually, multiplicative updating rules are designed for the optimization of NMF problems.

3. Proposed method

Based on the original NMF, the CNMF limits the base of the NMF to a convex combination of data columns, resulting in the following results:

$$\min_{W\geq 0, V\geq 0} \left\| X - XWV^T \right\|_F^2, \tag{3.1}$$

with $X \in \mathcal{R}^{d\times n}$, $W \in \mathcal{R}^{n\times k}$, $V \in \mathcal{R}^{n\times k}$, and k being the number of clusters. Here, (3.1) is also known as the concept factorization[3]. By restricting $U = XW$, (3.1) has the advantage that it could interpret the columns of U as weighted sums of certain data points and these columns correspond to centroids[17]. It is natural to see that W_{ij} reveals the importance of basis U_j to x_i by the value of W_{ij}.
It is worth noted that (3.1) is closely related to subspace clustering[18,19]. It is observed that high-dimensional data usually reside in low-dimensional subspaces, and often has self-expressive assumptions to recover these subspaces. The self-expressiveness assumption is that the data X can be approximately by X ≈ XZ, where Z is a representation matrix. That is, we can treat WVT as a single matrix and denote it by Z for simplicity, inspired by the subspace clustering model, we can use the ridge regression model[19] to find a suitable representation of the data, which

can be represented as:

$$\min_{W\geq 0, V\geq 0}\left\|X - XWV^T\right\|_F^2 + \gamma\left\|WV^T\right\|_F^2, \tag{3.2}$$

where γ ≥ 0 is a balancing parameter. Furthermore, the squared Frobenius norm regularization is a relaxed penalization for the rank constraint, where the kernel norm for low-rank representations is the sum of singular values, while the square Frobenius norm is the sum of squares of singular values. As can be seen, rank constrained rendering (3.2) can recover the global structure information of the data[20,21]. However, we can see from the previous section that the given z does not have a very clear clustering structure, in order to make the obtained graph a clear clustering structure, we hope to construct a graph containing precisely k connected parts. So we define a new graph similarity matrix as:

$$S = \begin{bmatrix} & \left(WV^T\right)^T \\ \left(WV^T\right) & \end{bmatrix} \in \mathcal{R}^{(n+n)\times(n+n)} \tag{3.3}$$

Then we get the normalized Laplacian matrix $\tilde{L} = I - D^{-\frac{1}{2}}SD^{-\frac{1}{2}}$, where $d_i = \sum_{j=1}^{n+n} s_{ij}$ is the i-th element of the diagonal matrix D. According to the following theorem, the rank constraint of Laplacian matrix ensures S to be k connected.

Theorem1: The multiplicity k of the eigenvalue 0 of the normalized Laplacian matrix $\tilde{L}$ is equal to the number of connected components in the graph associated with S.

Therefore, the graph S has precisely k connected components with rank constraints on $\tilde{L}$. Since S is composed of Z, Z has the same amount of connected components as S. Besides, Theorem1 indicates that if $rank(\tilde{L}) = n + n - k$, the problem (3.2) can be rewritten as:

$$\begin{gathered}\min_{W,V}\left\|X - XWV^T\right\|_F^2 + \gamma\left\|WV^T\right\|_F^2 \\ \text{s.t. } W\geq 0, V\geq 0, \operatorname{rank}(\mathring{L}) = n+n-k.\end{gathered} \tag{3.4}$$

Define $\sigma_i(\tilde{L})$ as the i-th smallest eigenvalue of $\tilde{L}$. Then $\sigma_i(\tilde{\boldsymbol{L}}) \geq 0$ since $\sigma_i(\tilde{\boldsymbol{L}})$ is semi-definite. With λ to be large enough, (3.4) equals to the following problem:

$$\begin{gathered}\min_{W,V}\left\|X - XWV^T\right\|_F^2 + \gamma\left\|WV^T\right\|_F^2 + \lambda\sum_{i=1}^{k}\sigma_i(\mathring{L}) \\ \text{s.t. } W\geq 0, V\geq 0.\end{gathered} \tag{3.5}$$

When λ is large enough (note that $\sigma_i(\tilde{\boldsymbol{L}}) \geq 0$ for every i), the optimal solution S to the problem (3.5) will make the third term $\lambda\sum_{i=1}^{k}\sigma_i(\tilde{\boldsymbol{L}})$ to be zero, and thus the constraint $rank(\tilde{\boldsymbol{L}}) = n + n - k$ in the problem (3.4) would be satisfied. Besides, to allow for immediate interpretation of clustering from the coefficient matrix, we impose an orthogonality constraint of V. It should be noted that the constraint $V^TV = I_k$ directly connects the problem of NMF with clustering in that V can be regarded as relaxed cluster indicators. Eventually, the final model of our proposed method can be formulated as:

$$\min_{W,V,F} \left\| X - XWV^T \right\|_F^2 + \gamma \left\| WV^T \right\|_F^2 + \lambda Tr\left(F^T \mathring{L} F \right) \tag{3.6}$$

$$\text{s.t. } W \geq 0, V \geq 0, F^T F = I, V^T V = I.$$

In the next section, we will develop an efficient algorithm to optimize the above model.

4. Optimization

To solve the above optimization, all variables are updated alternatively, which means updating one variable with others being fixed. Specially, we first consider the optimization for F while keeping W and V fixed; then we update W and V with fixed F. We will present the detailed optimization strategy in rest of this section.

4.1. Optimize F by Fixing Other Variables

When fixed other variables, the optimization problem for F is as follows:

$$\min_{F^T F} Tr\left(F^T \mathring{L} F \right). \tag{4.1}$$

To solve (4.1), we need to compute the eigenvectors of $\tilde{L}$. In order to improve time efficiency, we compute the eigenvectors of WVT instead of S. Specifically, according to the previous section, F and D can be decomposed as:

$$F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}, D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}, \tag{4.2}$$

Then we can rewrite (4.1) as follows:

$$\max_{F_1^T F_1 + F_2^T F_2 = I} Tr\left(F_1^T D_1^{-\frac{1}{2}} WV^T D_2^{-\frac{1}{2}} F_2 \right), \tag{4.3}$$

(4.3) can be easily solved by the following theorem.

*Theorem*2: Suppose $M \in \mathcal{R}^{n\times m}$, $A \in \mathcal{R}^{n\times k}$, $B \in \mathcal{R}^{m\times k}$, we have the follow problem:

$$\max_{A^T A + B^T B = I_k} Tr\left(A^T MB \right), \tag{4.4}$$

The optimal solutions of the problem are $A = \sqrt{2}/2L$ and $B = \sqrt{2}/2R$, where L, R are the leading k left and right singular vectors of M, respectively.

After optimizing F, we optimize W, V in (3.6) with the optimal F. To solve the problem, wo noticed:

$$
\begin{aligned}
Tr\left(F^T \mathring{L} F\right) &= \frac{1}{2}\sum_{i=1}^{2n}\sum_{j=1}^{2n}\left\|\frac{f^i}{\sqrt{D(i,i)}}-\frac{f^j}{\sqrt{D(j,j)}}\right\|_2^2 s_{ij} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\left\|\frac{f_{(1)}^i}{\sqrt{D_{(1)}(i,i)}}-\frac{f_{(2)}^j}{\sqrt{D_{(2)}(j,j)}}\right\|_2^2 \left(wv^T\right)_{ji} \\
&= Tr\left(TWV^T\right).
\end{aligned} \tag{4.5}
$$

4.2. Optimize W, V by Fixing Other Variables

When fixed other variables, the optimization problem for W, V is as follows:

$$
\begin{aligned}
&\min_{W,V}\left\|X-XWV^T\right\|_F^2+\gamma\left\|WV^T\right\|_F^2+\lambda Tr\left(TWV^T\right) \\
&\text{s.t. } W\geq 0, V\geq 0
\end{aligned} \tag{4.6}
$$

For this problem, we provide the following alternating optimization for W and V in the following:

$$
W_{ik}\leftarrow W_{ik}\sqrt{\frac{\left(2X^T XV\right)_{ik}}{\lambda\left(T^T V\right)_{ik}+\left(2X^T XWV^T V\right)_{ik}+\gamma(2W)_{ik}}}. \tag{4.7}
$$

$$
V_{ik}\leftarrow V_{ik}\sqrt{\frac{\left(2X^T XW\right)_{ik}+\lambda\left(VV^T TW\right)_{ik}}{\left(2VV^T X^T XW\right)_{ik}+\lambda(TW)_{ik}}}. \tag{4.8}
$$

5. Experiments

In this section, we conduct extensive experiments to verify the effectiveness of the Figer-CF. We compare our method with eight state-of-the-art clustering methods, including the WNMF[22], ONMF[5], CNMF[3], KNMF[17], RMNMF[23], OPMC[24], MKKM-SR [25], and EWNMF[26], on four widely used mark data sets, including the Jaffe, PIX, Semeion, and COIL20. Three metrics, including the clustering accuracy (ACC), normalized mutual information (NMI), and purity (PUR) are used in our experiments for evaluation, whose details can be found in[11].

5.1. Experimental settings

As a common step of all methods, we apply the K-means method to the coefficient matrix to obtain the final clustering results. Since the K-means is nonconvex, we follow[11]and conduct 200 independent trials, among which the one with the minimum objective value is used as the final result. For each data set, the performance is recorded and reported in **Table 1-4**. Since the OPMC is a multi-view method, we use the original features and the extracted LBP features to construct multi-view features.

For the multi-kernel methods, the MKKM-SR, we follow their original parameter settings[25]. For all the other methods, we tune the balancing parameters within $Y = \{10^{-3}, 10^{-2}, \cdots, 10^{3}\}$. For all the single kernel-based methods, the RBF kernel is used and the radial parameter is searched within Y.

Table 1. Clustering performance on Jaffe

Jaffe	Evaluation metrics		
	ACC	NMI	PUR
WNMF	90.61	89.44	90.61
RMNMF	82.16	83.09	83.57
CNMF	69.95	70.65	74.18
KNMF	81.69	82.38	82.16
ONMF	80.28	83.80	81.22
MKKM-SR	84.98	86.59	84.98
OPMC	80.75	85.44	83.57
EWNMF	95.31	93.43	95.31
Our-Method	96.71	95.47	96.71

Table 2. Clustering performance on PIX

PIX	Evaluation metrics		
	ACC	NMI	PUR
WNMF	74.00	83.91	79.00
RMNMF	75.00	76.88	77.00
CNMF	80.00	82.97	83.00
KNMF	69.00	80.90	74.00
ONMF	69.00	80.06	76.00
MKKM-SR	79.00	85.47	82.00
OPMC	62.00	66.78	64.00
EWNMF	81.00	89.18	86.00
Our-Method	94.00	94.40	94.00

Table 3. Clustering performance on Semeion

Semeion	Evaluation metrics		
	ACC	NMI	PUR
WNMF	55.56	44.82	56.56
RMNMF	41.31	36.59	45.26
CNMF	45.20	37.96	45.20
KNMF	52.54	47.38	52.54
ONMF	50.78	42.75	52.92
MKKM-SR	41.75	38.28	44.88
OPMC	65.41	54.62	65.47
EWNMF	50.97	42.95	53.80
Our-Method	59.07	52.78	63.34

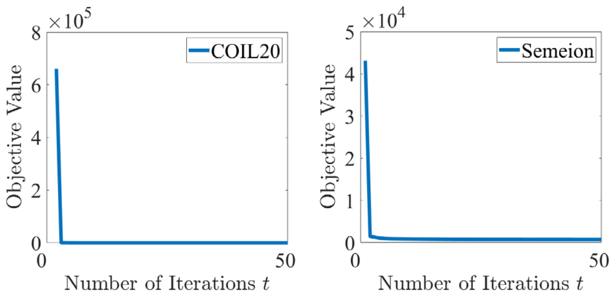
Table 4. Clustering performance on COIL20

COIL20	Evaluation metrics		
	ACC	NMI	PUR
WNMF	66.88	73.96	67.99
RMNMF	56.39	67.11	59.24
CNMF	52.64	65.33	56.46
KNMF	61.60	71.04	65.28
ONMF	57.01	71.34	59.38
MKKM-SR	63.06	74.44	65.90
OPMC	52.22	68.16	56.32
EWNMF	68.06	76.22	68.43
Our-Method	68.33	78.42	68.61

5.2. Convergence Study

To verify the convergence of the method proposed in this paper through experiments, the sequence curves of the objective function values on the Semeion and COIL20 datasets are respectively presented in **Figure 1**, with the fixed parameters and without loss of generality. It can be seen that the target value sequences of the two datasets usually converge within approximately 50 iterations, which confirms the effectiveness and efficiency of the optimization algorithm.

Figure 1. Convergence curve of the objective value sequence



Acknowledgements

This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grants 62276147, 62172246, and 62106063.

References

[1] A. Sotiras, S. M. Resnick, and C. Davatzikos, "Finding imaging patterns of structural co-variance via non-negative matrix factorization," NeuroImage, p. 1–16, Mar 2015. [Online]. Available: http://dx.doi.org/ 10.1016/j.neuroimage.2014.11.045

[2] D. Behnia, K. Ahangari, and S. R. Moeinossadat, "Modeling of shear wave ve-locity in limestone by soft computing methods," International Journal of Mining Science and Technology, vol. 27, no. 3, p. 423–430, May 2017. [Online]. Available: http://dx.doi.org/10.1016/j.ijmst.2017. 03.006

[3] H. Liu, Z. Wu, D. Cai, and T. S. Huang, "Constrained nonnegative matrix fac-torization for image representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 1299–1311, Jul 2012. [Online]. Available: http://dx.doi.org/10.1109/tpami.2011.217

[4] Y. Akashi and T. Okatani, “Separation of reflection components by sparse non-negative matrix factorization,” Computer Vision and Image Understanding, p. 77–85, May 2016. [Online]. Available: http://dx.doi.org/10.1016/j.cviu.2015.09.001

[5] C. Ding, T. Li, W. Peng, and H. Park, “Orthogonal nonnegative matrix t-factorizations for clustering,” in Proceedings of the 12th ACM SIGKDD interna-tional conference on Knowledge discovery and data mining, Aug 2006. [Online]. Available: http://dx.doi.org/10.1145/1150402.1150420

[6] D. Cai, X. He, J. Han, and T. Huang, “Graph regularized nonnegative matrix factorization for data representation.” IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 8, pp. 1548–60, 2011.

[7] X. Li, G. Cui, and Y. Dong, “Graph regularized non-negative low-rank matrix factorization for image clustering,” IEEE Transactions on Cybernetics, p. 3840–3853, Nov 2017. [Online]. Available: http://dx.doi.org/10.1109/tcyb.2016.2585355

[8] C. Peng, Y. Zhang, Y. Chen, Z. Kang, C. Chen, and Q. Cheng, “Log-based sparse nonnegative matrix factorization for data representation,” Apr 2022.

[9] Z. Li, J. Tang, and X. He, “Robust structured nonnegative matrix factorization for image representation,” IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 5, p. 1947–1960, May 2018. [Online]. Available: http://dx.doi.org/10.1109/tnnls.2017.2691725

[10] Q. Huang, X. Yin, S. Chen, Y. Wang, and B. Chen, “Robust nonnegative matrix factorization with structure regularization,” Neurocomputing, vol. 412, p. 72–90, Oct 2020. [Online]. Available: http://dx.doi.org/10. 1016/j.neucom.2020.06.049

[11] C. Peng, Z. Zhang, Z. Kang, C. Chen, and Q. Cheng, “Nonnegative matrix fac-torization with local similarity learning,” Information Sciences, p. 325–346, Jul 2021. [Online]. Available: http://dx.doi.org/ 10.1016/j.ins.2021.01.087

[12] P. Deng, H. Wang, T. Li, H. Zhao, and Y. Wu, “Graph regularized sparse non-negative ma-trix factorization for clustering,” in Developments of Artificial Intelligence Technologies in Computation and Robotics, Oct 2020. [Online]. Available: http://dx.doi.org/10.1142/9789811223334 0119

[13] D. Kong, C. Ding, and H. Huang, “Robust nonnegative matrix factorization using l21-norm,” in Proceedings of the 20th ACM international conference on In-formation and knowledge management, Oct 2011. [Online]. Available: http://dx.doi.org/10.1145/2063576. 2063676

[14] F. Nie, H. Huang, X. Cai, and C. Ding, “Efficient and robust feature selection via joint l2,1-norms minimization,” Neural Information Processing Sys-tems,Neural Information Processing Systems, Dec 2010.

[15] F. Nie, X. Wang, C. Deng, and H. Huang, “Learning a structured optimal bi-partite graph for co-clustering,” Neural Information Processing Systems,Neural Information Pro-cessing Systems, Jan 2017.

[16] C. Peng and Q. Cheng, “Discriminative ridge machine: A classifier for high-dimensional data or imbalanced data,” IEEE Transactions on Neural Net-works and Learning Systems, p. 2595–2609, Jun 2021. [Online]. Available: http://dx.doi.org/10.1109/tnnls.2020.3006877

[17] C. H. Q. Ding, T. Li, and M. I. Jordan, “Convex and semi-nonnegative matrix factorizations.” in Machine Learning & Knowledge Discovery in Databases, Euro-pean Conference, Ecml Pkdd, Bled, Slovenia, September, 2009.

[18] C. Peng, Z. Kang, H. Li, and Q. Cheng, “Subspace clustering using log-determinant rank approximation,” in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug 2015. [Online]. Available: http://dx.doi.org/10.1145/ 2783258.2783303

[19] C. Peng, Q. Zhang, Z. Kang, C. Chen, and Q. Cheng, “Kernel two-dimensional ridge regres-sion for subspace clustering,” Cornell University - arXiv,Cornell Uni-versity - arXiv, Nov 2020.

[20] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank repre-sentation," International Conference on Machine Learning,International Confer-ence on Machine Learning, Jun 2010.

[21] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 171–184, Jan 2013. [Online]. Available: http://dx.doi.org/10.1109/tpami.2012.88

[22] Yong-Deok Kim and Seungjin Choi. Weighted non-negative matrix factoriza-tion. In Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE Interna-tional Conference on, pages 1541–1544. IEEE, 2009.

[23] J. Huang, F. Nie, H. Huang, and C. Ding, "Robust manifold nonnegative matrix factorization," ACM Transactions on Knowledge Discovery from Data, p. 1–21, Jun 2014. [Online]. Available: http://dx.doi.org/10.1145/2601434

[24] J. Liu, X. Liu, Y. Yang, L. Liu, S. Wang, W. Liang, and J. Shi, "Onepass mul-ti-view clustering for large-scale data," in 2021 IEEE/CVF International Confer-ence on Computer Vision (ICCV), Oct 2021. [Online]. Available: http://dx.doi.org/10.1109/iccv48922.2021.01212

[25] J. Lu, Y. Lu, R. Wang, F. Nie, and X. Li, "Multiple kernel k-means clustering with simultaneous spectral rotation," in ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2022. [Online]. Available: http://dx.doi.org/10.1109/icassp43922.2022.9746905

[26] J. Wei, C. Tong, B. Wu, Q. He, S. Qi, Y. Yao, and Y. Teng, "An entropy weighted nonnegative matrix factorization algorithm for feature representation."