Double Manifolds Regularized Non-negative Matrix Factorization for Data Representation

Jipeng Guo, Shuai Yin, Yanfeng Sun and Yongli Hu

Beijing Key Laboratory of Multimedia and Intelligent Software Technology,
Beijing Institute of Artificial Intelligence, Faculty of Information Technology,
Beijing University of Technology, China

Abstract

Non-negative matrix factorization (NMF) is an important method in learning latent data representation. The local geometrical structure can make the learned representation more effectively and significantly improve the performance of NMF. However, most of existing graph-based learning methods are determined by a predefined similarity graph which may be not optimal for specific tasks. To solve the above problem, we propose the Double Manifolds Regularized NMF (DMR-NMF) model which jointly learns an adaptive affinity matrix with the non-negative matrix factorization. The learned affinity matrix can guide the NMF to fit the clustering task. Moreover, we develop the iterative updating optimization schemes for DMR-NMF, and provide the strict convergence proof of our optimization strategy. Empirical experiments on four different real-world data sets demonstrate the state-of-the-art performance of DMR-NMF in comparison with the other related algorithms.

Contribution

- Proposing a novel double manifolds regularized non-negative matrix factorization (DMR-NMF) algorithm for data representation and clustering;
- DMR-NMF utilizes a basic subspace clustering on the latent representation to obtain the low rank self-expressiveness affinity matrix. The DMR-NMF model jointly learns an adaptive affinity matrix with the non-negative matrix factorization, therefore, the ideal data structure under our assumption can be well uncovered, and the learned affinity matrix may better guide the matrix factorization to learn efficient data representation.; and
- DMR-NMF is formulated as an optimization problem with a well-defined objective function. An efficient iterative method is designed to solve the DMR-NMF problem. Both the theoretical analysis and empirical performance are provided to demonstrate the convergence of the optimization strategy;
- The experimental results on several real world data sets demonstrate the excellent performance of DMR-NMF.

Related Works

The **non-negative matrix factorization (NMF)** aims to find two non-negative matrices whose product provides a better approximation for the original feature matrix. Given a data matrix $\mathbf{X} \in \mathbb{R}^{d\times n}$, the objective function of the stand NMF model is defined as

$$\min_{\mathbf{U}\geq 0,\mathbf{V}\geq 0} \|\mathbf{X}-\mathbf{U}\mathbf{V}\|_F^2, \tag{1}$$

where $\mathbf{U}=[u_{ik}] \in \mathbb{R}^{d\times r}(r \leq d)$ is a basis matrix and $\mathbf{V}=[v_{kj}] \in \mathbb{R}^{r\times n}$ is the latent feature representation.

Graph Laplacian-based embedding effectively characterizes the similarity between data and has been widely used to preserve the local manifold structure. Given the similarity matrix $\mathbf{W}$, the **graph regularized NMF (GNMF)** model is formulated as below:

$$\min_{\mathbf{U}\geq 0,\mathbf{V}\geq 0} \|\mathbf{X}-\mathbf{U}\mathbf{V}\|_F^2 + \lambda \mathrm{tr}(\mathbf{V}\mathbf{L}\mathbf{V}^T), \tag{2}$$

where $\mathbf{L}=\mathbf{D}-\mathbf{W}$ is the graph Laplacian matrix, in which $\mathbf{D}$ is a diagonal matrix and $\mathbf{D}_{jj}=\sum_{l=1}^{n}\mathbf{W}_{jl}$.

Motivation

- The similarity matrix $\mathbf{W}$ is artificially defined according to raw feature in advance, which may be not accurate since the existence of noises. Thus, $\mathbf{W}$ in GNMF is not an optimal graph for characterizing the complex intrinsic structure of data.
- And, the global structure of data is not explored for GNMF. Those reduce the flexibility of NMF and heavily affect the performance of the algorithm.

Main Reference

[1] D. D. Lee and H. H. Seung. Learning the parts of objects by non-negative matrix factorization. Nature 1999.
[2] D. Cai, X. He, J. Han and T. S. Huang. Graph regularized nonnegative matrix factorization for data representation. PAMI 2010.
[3] G. Liu, Z. Lin, J. Sun, Y. Yu and Y. Ma. Robust Recovery of Subspace Structures by Low Rank Representation. PAMI 2013.
[4] W. Qian, B. Hong, D. Cai, X. He and X. Li. Non-negative matrix factorization with sinkhorn distance. In IJCAI 2016.

Double manifolds regularized non-negative matrix factorization

Low rank representation(LRR) [3] aims to seek low-rank affinity graph which can effectively reveal the global structures of data. And due to the complex data distribution, the single manifold structures (such as only global or only local) may be not sufficient to describe the underlying true structure. Thus, double manifolds regularized non-negative matrix factorization (DMR-NMF) minimizes the following objective function with the non-negative constraints $\mathbf{U}\geq 0$ and $\mathbf{V}\geq 0$:

$$\begin{aligned}&\min_{\mathbf{U},\mathbf{V}} \|\mathbf{X}-\mathbf{U}\mathbf{V}\|_F^2 + \lambda\|\mathbf{Z}\|_* + \gamma\|\mathbf{V}-\mathbf{V}\mathbf{Z}\|_F^2 + \beta \mathrm{tr}(\mathbf{V}\mathbf{L}\mathbf{V}^T),\\ &\text{s.t.}\ \mathbf{U}\geq 0, \mathbf{V}\geq 0.\end{aligned} \tag{3}$$

Main Experiments

Databases	K	ACC Kmeans	NMF [1]	GNMF [2]	SDNMF [4]	DMR-NMF
ORL	10	75.80	75.80	77.60	77.70	**80.00**
	20	63.70	66.70	68.20	68.50	**69.30**
	30	59.33	61.40	61.00	61.58	**63.17**
	40	57.50	57.30	58.95	**60.90**	59.75
UMIST	5	76.20	79.80	89.40	89.00	**92.40**
	10	66.70	70.00	80.00	79.10	**83.00**
	15	61.47	60.40	68.87	69.60	**72.27**
	20	54.65	56.95	66.75	64.60	**68.55**
Yale	7	76.36	74.29	78.44	76.10	**82.86**
	9	68.48	66.46	71.52	71.11	**73.94**
	11	60.99	61.16	63.14	61.36	**65.95**
	13	58.18	57.90	58.80	57.48	**60.35**
	15	55.88	54.52	56.36	57.58	**58.42**
COIL20	4	91.32	90.90	95.35	95.42	**98.71**
	8	85.14	85.66	91.70	91.15	**92.88**
	12	78.45	73.52	85.32	86.92	**88.23**
	20	68.56	66.90	76.25	78.54	**79.55**

Table 1: Clustering results of different methods by the measurement of ACC on four databases.

Databases	K	NMI Kmeans	NMF [1]	GNMF [2]	SDNMF [4]	DMR-NMF
ORL	10	73.78	74.36	74.69	72.34	**78.62**
	20	72.34	72.58	74.85	75.54	**76.08**
	30	70.62	71.32	72.73	71.06	**73.19**
	40	71.50	71.62	72.44	72.40	**73.62**
UMIST	5	68.42	70.22	84.77	83.54	**87.03**
	10	67.37	69.09	78.75	78.26	**80.90**
	15	66.70	65.27	73.78	74.58	**76.14**
	20	66.66	67.20	72.81	71.79	**74.90**
Yale	7	58.62	60.54	64.74	64.50	**68.47**
	9	54.97	54.02	60.70	55.60	**61.83**
	11	49.37	50.74	54.22	49.83	**55.31**
	13	49.12	48.74	49.56	51.04	**52.31**
	15	49.65	48.34	51.86	51.73	**52.04**
COIL20	4	84.07	80.94	92.01	92.01	**95.32**
	8	79.73	78.75	88.84	85.70	**90.09**
	12	76.68	72.09	81.28	84.73	**85.28**
	20	71.55	67.42	80.69	81.41	**81.43**

Table 2: Clustering results of different methods by the measurement of NMI on four databases.

guojipeng@emails.bjut.edu.cn