## ADMM, ALM and LADMM for solving NMF\_l20

To better demonstrate that our optimization scheme is necessary, we compare against other ways of implementing and optimizing a L2,0 constraints, including Alternating Direction Method of Multipliers (ADMM) [1,2], Augmented Lagrangian Multiplier (ALM) [3] and Linearized Alternating Direction Method of Multipliers (LADMM) [4]. We compare the PALM and maPALM algorithms for solving NMF-L20 with ADMM, ALM, and LADMM algorithms.

## Alternating Direction Method of Multipliers (ADMM) Method

We first show the row sparse NMF (NMF\_l20) model as follows

(1)

We use variable splitting in which a slack variable is introduced to get a new formulation.

(2)

To use the Alternating Direction Method of Multipliers (ADMM) algorithm, we construct the augmented Lagrangian function which is given by

(3)

where is the Lagrangian dual variable, and the ADMM method updates variables iteratively as the following

(4)

(5)

(6)

(7)

Because (4) and (5) are convex optimization problems, it is very clear how to update for and . I omit the detailed steps here. In addition, we can obtain the closed-form solution of Eq. (6),

(8)

For the definition of , see Eq. (18) in the main text. With all the subproblems solved, the procedure of the ADMM is summarized in Algorithm 1.

**Algorithm 1 ADMM for solving the NMF\_l20 model**

Input: Data matrix , (nonzero rows, i.e., the number of features)

Parameter (Refer to [1])

Cluster amount

Output: and .

1: Initialize () and

2: repeat

3: Update H using (4)

4: Update W using (5)

5: Update E using (8)

6: Update B using (7)

7: until convergence

8: Return and .

In Algorithm 1, regarding the value of , I refer to [1] and the idea of ADMM to solve non-convex optimization problems is shown in ref. [2].

## Augmented Lagrangian Multiplier (ALM) Method

Inspired by ref. [3], we adopt the ALM approach to solve the NMF\_l20 model. Similar to the derivation of the ADMM algorithm, we first obtain the augmented Lagrangian function

(9)

The ALM method updates variables iteratively as the following

(10)

(11)

(12)

(13)

ALM generally uses a sequence of penalty parameters , which is nondecreasing and possibly unbounded. The procedure of the ALM is summarized in Algorithm 2.

**Algorithm 2 ALM for solving the NMF\_l20 model**

Input: Data matrix , (nonzero rows, i.e., the number of features)

Cluster amount

Output: and .

1: Initialize () and

2: Initialize (Refer to [1])

3: repeat

4: Update H using (10)

5: Update W using (11)

6: Update E using (12)

7: Update B using (13)

8: Update

9: until convergence

10: Return and .

Note that the difference between ALM and ADMM algorithms is in step 8.

## Linearized ADMM (LADMM) Method

Due to the non-convexity of L20-norm in the NMF\_l20 model, we adopt the LADMM method to solve the NMF\_l20 model [4]. To use the LADMM method, we construct the augmented Lagrangian function as follows:

(14)

where is the Lagrangian dual variable. Compared with the ADMM method, the LADMM apply the Proximal Gradient Descent (PGD) method for updating and in each iteration. So, the LADMM method updates variables iteratively as the following

(15)

(16)

(17)

(18)

The procedure of the LADMM is summarized in Algorithm 3.

**Algorithm 3 LADMM for solving the NMF\_l20 model**

Input: Data matrix , (nonzero rows, i.e., the number of features)

Parameter (Refer to [1])

Cluster amount

Output: and .

1: Initialize () and

2: repeat

3: Update H using (15)

4: Update W using (16)

5: Update E using (17)

6: Update B using (18)

7: until convergence

8: Return and .

## Reference

1. Xu Y, Yin W, Wen Z, et al. An alternating direction algorithm for matrix completion with nonnegative factors[J]. Frontiers of Mathematics in China, 2012, 7(2): 365-384.
2. Wang Y, Yin W, Zeng J. Global convergence of ADMM in nonconvex nonsmooth optimization[J]. Journal of Scientific Computing, 2019, 78(1): 29-63.
3. Cai X, Nie F, Huang H. Exact top-k feature selection via l2,0-norm constraint[C]//Twenty-third international joint conference on artificial intelligence. 2013.
4. Liu Q, Shen X, Gu Y. Linearized ADMM for nonconvex nonsmooth optimization with convergence analysis[J]. IEEE Access, 2019, 7: 76131-76144.