Supplementary Materials

In this supplementary materials, we first prove a useful matrix inequality and then the convergence of Algorithm 1.

Important Matrix Inequalities

To prove Algorithm 1, we need the following matrix inequalities: (A) Inequality involving quadratic form of matrix \( G \) among 2 other matrices where we used the inequality \([\[1\]\])
\[
\text{Tr}[G^T PGQ] \leq \sum_{ik} (PG'Q)_{ik} \frac{G^2_{ik}}{G_{ik}} \tag{S1}
\]
where \( P, Q \in \mathbb{R}^{K \times K}_+ \), and \( P = P^T, Q = Q^T \). \( G, G' \in \mathbb{R}^{n \times K}_+ \). The inequality was proved in \([\[1\]\]).

Another useful inequality involving quartic form of matrix \( H \) and 2 other matrices
\[
\text{Tr}(HAH^T HBH^T) \leq \sum_{i=1}^n \sum_{k=1}^K \left( \frac{AH'HBH'^T + H'BAH'}{2} \right)_{ik} \frac{H^4_{ik}}{H^3_{ik}} \tag{S2}
\]
where \( A, B \in \mathbb{R}^{K \times K}_+, H, H' \in \mathbb{R}^{n \times K}_+ \) are nonnegative matrices; and \( A, B \) are symmetric: \( A = A^T, B = B^T \).

A slight variant of the above is
\[
\text{Tr}(H^T AHH^T BH) \leq \sum_{i=1}^n \sum_{k=1}^K \left( \frac{AH'HBH'^T + BH'AH'}{2} \right)_{ik} \frac{H^4_{ik}}{H^3_{ik}} \tag{S3}
\]
where \( A, B \in \mathbb{R}^{n \times n}_+ \).

This inequality is useful when analyzing objective functions involving 4-th order matrix polynomials, such as the NMF of Eq.(1). We will use this inequality to prove the convergence of the algorithm of Eq.(S2).

Proof of the inequality Eq.(S3)

Let \( H_{ik} = H'_{ik} \mu_{ik} \). The 2nd term in RHS of Eq.(S3) is
\[
\sum_{ik} (H'BH'T H' A)_{ik} \frac{H^4_{ik}}{H_{ik}^3} = \sum_{ijkrpq} H'_{ip}B_{pq}H'_{jq}H'_{jr}A_{kr}H'_{ik} \mu_{ik}^4 \tag{S4}
\]
Now, switching indexes: \( i \leftrightarrow j, p \leftrightarrow q, r \leftrightarrow k \), we obtain
\[
\sum_{ik} (H'BH'T H' A)_{ik} \frac{H^4_{ik}}{H_{ik}^3} = \sum_{ijkrpq} H'_{jq}B_{pq}H'_{ip}H'_{ik}A_{kr}H'_{jr} \mu_{jr}^4 \tag{S5}
\]
The 1st term in RHS of Eq.(S3) is
\[
\sum_{ik} (H'AH'T H' B)_{ik} \frac{H^4_{ik}}{H_{ik}^3} = \sum_{ijkrpq} H'_{jr}A_{kr}H'_{ik}H'_{ip}B_{pq}H'_{jq} \mu_{jq}^4 \tag{S6}
\]
Now, switching indexes: $i \iff j, p \iff q, r \iff k$, we obtain

$$\sum_{i \neq k} (H'_i A H^T i H' j B)_{ik} \frac{H'_{ik}}{H'_{jk}} = \sum_{i \neq j \neq k \neq p} H'_{ik} A_k H'_i H'_{jq} B_{qp} H'_{ip}$$  \hspace{1cm} (S7)

Carefully examination of the RHS of Eqs. (S4 - S7) show that they are identical except $\mu^4$ terms. Adding Eqs. (S4 - S7), we obtain that the RHS of Eq. (S3) is equal to

$$\sum_{i \neq j \neq k \neq p} H'_{ip} B_{pq} H'_{jq} A_{rk} H'_i H'_{ik} \mu_{ik} + \mu_{jr} + \mu_{jq} + \mu_{ip}$$  \hspace{1cm} (S8)

The LHS of Eq. (S3) is equal to

$$\sum_{i \neq j \neq k \neq p} H'_{ip} B_{pq} H'_{jq} A_{rk} H'_i H'_{ik} \mu_{ik} \mu_{jr} \mu_{jq} \mu_{ip}$$  \hspace{1cm} (S9)

Therefore, if we can establish

$$\mu_{ik} \mu_{jr} \mu_{jq} \mu_{ip} \leq \frac{\mu_{ik}^4 + \mu_{jr}^4 + \mu_{jq}^4 + \mu_{ip}^4}{4},$$  \hspace{1cm} (S10)

then the inequality Eq. (S3) holds. For any $a, b, c, d > 0$, we have

$$a^4 + b^4 \geq 2a^2 b^2, \quad c^4 + d^4 \geq 2c^2 d^2,$$

$$\left(ab\right)^2 + \left(cd\right)^2 \geq 2(ab)(cd).$$  \hspace{1cm} (S11)

Thus

$$a^4 + b^4 + c^4 + d^4 \geq 2\left(a^2 b^2 + c^2 d^2\right) \geq 4(ab)(cd).$$  \hspace{1cm} (S12)

This is Eq. (S10).

\textbf{Convergence Proof of Algorithm 1}

\textbf{Proof.} We use the auxiliary function approach\cite{15}. Function $Z(G, G')$ is an auxiliary function of $J(G)$ if it satisfies

$$Z(G, G') \geq J(G), \quad Z(G, G) = J(G).$$  \hspace{1cm} (S13)

Once an auxiliary function is found, we define

$$G^{t+1} = \arg \min_G Z(G, G(t))$$  \hspace{1cm} (S14)

Then, by construction, we have

$$J(G^{(t)}) = Z(G^{(t)}, G^{(t)}) \leq Z(G^{(t+1)}, G^{(t)}) \leq J(G^{(t+1)}).$$  \hspace{1cm} (S15)

This proves that $J(G^{(t)})$ is monotonically decreasing. The key steps in the remainder of the proof are: (1) Finding an appropriate auxiliary function of the objective function; (2) Finding the global maxima of the auxiliary function; (3) Showing that the update rule of Eq. (S4) gives the global minima of Eq. (S14).

\textbf{Auxiliary function}
The objective function Eq. (7) is

$$Tr[-2G^T(A - B)G + G^T(A - B)GG^T G] =$$

$$Tr[2G^T BG + G^T AGG^T G] - 2G^T AG - G^T BGG^T G]$$  \hspace{1cm} (S16)
where $A = (X^T X)_+$ $\geq 0$ and $B = (X^T X)_- \geq 0$.

Now, we construct an auxiliary function $Z(G, G')$ as the following

$$Z(G, G') = \text{RHS of Eq.}(S18) + \text{RHS of Eq.}(S20)$$

$$-\text{RHS of Eq.}(S21) - \text{RHS of Eq.}(S22).$$

This is done by introducing the upper bounds of the first two terms of Eq.(S16) and the lower bounds of the last two terms (ignoring the negative sign).

The upper bounds of the first terms of Eq.(S16) is

$$\text{Tr}[G^T BG] \leq \sum_{ik}(BG')_{ik}\frac{G^2_{ik}}{G_{ik}} \leq \sum_{ik}(BG')_{ik}\frac{G^4_{ik} + G'^4_{ik}}{2G_{ik}}$$

and (derived from $2gg' \leq g^2 + g'^2$)

$$G^2_{ik} \leq \frac{G^4_{ik} + G'^4_{ik}}{2G_{ik}^4}$$

Clearly, Eq.(S18) becomes equality when $G = G'$, satisfying Eq.(S13). The upper bound of 2nd term is

$$R \leq \sum_{ik} \frac{(AGG^T G)_{ik} + (GG^T AG)_{ik}}{2} G^3_{ik}$$

where $R = \text{Tr}[G^T AGG^T G]$, Eq.(S20) becomes equality when $G = G'$, satisfying Eq.(S13).

The lower bound of 3rd term is

$$\text{Tr}[2G^T BG] \leq 2 \sum_{ijk} G'_{ik} A_{ij} G'_{jk} (1 + \frac{G_{ik} G_{jk}}{G^2_{ik} G^2_{jk}})$$

where we use $z \geq 1 + \log z$, $z > 0$ and setting $z = \frac{G_{ik} G_{jk}}{G^2_{ik} G^2_{jk}}$. Eq.(S21) becomes equality when $G = G'$, satisfying Eq.(S13).

The lower-bound of 4th term is

$$R \leq \sum_{ijkl} G_{ik} A_{ij} G_{jl} G'_{lj} G'_{jk} (1 + \frac{G_{ik} G_{lj} G'_{lj} G'_{jk}}{G^2_{ik} G^2_{lj} G^2_{jk} G^2_{jk}})$$

where $R = \text{Tr}[G^T AGG^T G]$, Eq.(S22) becomes equality when $G = G'$, satisfying Eq.(S13). This completes the construction of the auxiliary function $Z(G, G')$.

**Finding global minima.**

Now we seek the global minima of $f(G) \equiv Z(G, G')$ while fixing $G'$. The gradient w.r.t. $G$ is [collecting RHS of Eqs.(S18,S20,S21,S22)]

\[
\frac{1}{2} \frac{\partial Z(G, G')}{\partial G_{ik}} = \frac{(BG')_{ik} G^3_{ik}}{G^2_{ik}} + \left(\frac{AGG^T G}{G^2_{ik}} + \frac{(G'G'T G')_{ik} + (G'G'T AG')_{ik}}{G^2_{ik}}\right) G^3_{ik} G^3_{ik} - \frac{G_{ik} (BG')_{ik}}{G^3_{ik}} - \frac{(BG'G'T G')_{ik} + (G'G'T BG')_{ik}}{G^2_{ik}}
\]

The Hessian (2nd derivative) are

\[
\frac{\partial^2 Z(Q, Q')}{\partial G_{ik} \partial G_{jl}} = \delta_{ij} \delta_{kl} D_{ik}
\]
where

\[
D_{ik} = 6\left(\frac{BG'}{G'_{ik}}\right) + 6\left[\frac{(AG'G'TG')_{ik} + (G'G'TAG')_{ik}}{G'_{ik}}\right] + 2\frac{G_{ik}(BG')_{ik}}{G'_{ik}} + 2\left(\frac{BG'G'TG'}{G'_{ik}} + (G'G'TBG')_{ik}\right) \tag{S25}
\]

Clearly the Hessian is diagonal with positive elements, and hence is semi-definite positive. Therefore \(f(Q) = Z(Q, Q')\) is a convex function. The global minima is obtained by setting the gradient Eq. (S23) to zero. We obtain

\[
G_{ik}^4 = G_{ik}^{4'} \left[\begin{array}{c}
(BG' + AG'G'TG' + G'G'TAG')_{ik} \\
(AG' + BG'G'TG' + G'G'TBG')_{ik}
\end{array}\right] \tag{S26}
\]

as the optimal solution to Eq.(S12) (note \(G^t \leftarrow G', G^{t+1} \leftarrow G\). Eq.(S26) which is identical to the update rule Eq.(S4). Thus we proved that under the update rule Eq.(S4) the objective function \(J(G)\) monotonically decrease.

References

1. Ding C, Li T, Jordan MI. Convex and semi-nonnegative matrix factorizations. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2010;24:45–55.