Uniqueness of solution of Minimum volume NMF

Andersen Ang

Mathématique et recherche opérationnelle
UMONS, Belgium

manshun.ang@umons.ac.be Homepage: angms.science

First draft : Feburary 9, 2020
Last update : February 9, 2020

Minimum volume NMF

Minimum volume regularized Exact (N)MF (minvol-NMF).
Given a matrix $\mathbf{X} \in \mathbb{R}^{m\times n}$ and a factorization rank $r \ll \min\{m,n\}$, find $\mathbf{W} \in \mathbb{R}^{m\times r}$ and $\mathbf{H} \in \mathbb{R}_+^{r\times n}$ by solving

$$
\begin{aligned}
\min_{\mathbf{W}} \quad & f(\mathbf{W}) := \det(\mathbf{W}^\top \mathbf{W}) && (1)\\
\text{s.t.} \quad & \mathbf{X} = \mathbf{W}\mathbf{H} && (2)\\
& \mathbf{H} \geq 0 && (3)\\
& \mathbf{H}^\top \mathbf{1}_r = \mathbf{1}_n && (4)
\end{aligned}
$$

Note that only $\mathbf{H}$ has to be nonnegative here.

This document : show the solution to minvol-NMF is unique under some technical assumptions.

Theorem on uniqueness of solution

Let $(\mathbf{W}, \mathbf{H})$ be the solution to the minvol-NMF.

Theorem (Theorem 1, Fu2015) If the followings are satisfied :

I) the rank of $\mathbf{X}$

$$\operatorname{rank}(\mathbf{X}) = r. \tag{5}$$

II) the *sufficiently scattered condition*

$$\text{The cone } \mathcal{C} = \{\mathbf{x} \in \mathbb{R}^r_+ \mid \langle \mathbf{1}_r, \mathbf{x} \rangle \geq \sqrt{r-1}\|\mathbf{x}\|_2\} \subseteq \operatorname{cone}(\mathbf{H}). \tag{6}$$

$$\begin{array}{l}\text{There does not exists any orthogonal matrix } \mathbf{Q} \text{ s.t.} \\ \operatorname{cone}(\mathbf{H}) \subseteq \operatorname{cone}(\mathbf{Q}), \text{except for permutation matrices.}\end{array} \tag{7}$$

then the solution $(\mathbf{W}, \mathbf{H})$ of minvol-NMF is unique (up to permutation).

Note that there is no scaling ambiguity due to (4).

Reference Fu et al., "Blind separation of quasi-stationary sources: Exploiting convex geometry in covariance domain", IEEE Trans. Signal Processing 63(9), 2015

Idea of the proof : proof by contradiction

- To prove a mathematical object is unique, a way is to use *proof by contradiction* : we first assume there are multiple objects exist, then show that this gives a contradiction, and thereby showing that the assumption is false, and hence showed the object is unique.

- For the minvol-NMF, we first assume that there are two solutions : $(\mathbf{W}_1, \mathbf{H}_1)$, $(\mathbf{W}_2, \mathbf{H}_2)$. Then we try to come up with a contradiction. The contradiction here is violation of the optimality — we assumed both $(\mathbf{W}_1, \mathbf{H}_1)$, $(\mathbf{W}_2, \mathbf{H}_2)$ solve and minimize the problem, thus they should give the same objective value in f. A contradiction in f will show the solution is unique.

Tools for the proof

- The matrix product $\mathbf{AB}$ gives a nonnegative matrix if $\text{cone}(\mathbf{A}^\top) \subseteq \text{cone}^*(\mathbf{B})$ (see here for the proof).

- For cone set, if $A \subseteq B$ then by duality $B^* \subseteq A^*$.

- The dual of the ice-cream cone $\mathcal{C}$ is

$$\mathcal{C}^* = \{\mathbf{x} \in \mathbb{R}^r \mid \langle \mathbf{1}_r, \mathbf{x} \rangle \geq \|\mathbf{x}\|_2\}.$$

- The second ssufficiently scattered condition (7) means that, if $\mathbf{H}$ satisfies (6) then

$$\text{cone}^*(\mathbf{H}) \cap \mathbf{bd}\, \mathcal{C}^* = \{\lambda \mathbf{e}_k \mid \lambda \geq 0, k = 1, 2, \ldots, r\}$$

  i.e., the intersection of the dual cone of $\mathbf{H}$ and the boundary of the dual cone of $\mathbf{C}$ is the set of standard basis vectors.

The proof ... 1/7

For the shake of contradiction, assume there are two sets of solution $(\mathbf{W}_1, \mathbf{H}_1)$, $(\mathbf{W}_2, \mathbf{H}_2)$ for the minvol-NMF.

The two solutions $(\mathbf{W}_1, \mathbf{H}_1)$, $(\mathbf{W}_2, \mathbf{H}_2)$ satisfy (1) to (7). More importantly, we have

$$f(\mathbf{W}_1) = \det(\mathbf{W}_1^\top \mathbf{W}_1) \overset{(1)}{=} \det(\mathbf{W}_2^\top \mathbf{W}_2) = f(\mathbf{W}_2). \tag{8}$$

The whole proof is to show (8) is false, thus a contradiction. To show this we first try to link up the relationship between $\mathbf{W}_1$, $\mathbf{W}_2$, and between $\mathbf{H}_1$, $\mathbf{H}_2$.

The proof ... 2/7

The solutions satisfy (2), we have $\mathbf{W}_1\mathbf{H}_1 = \mathbf{X} = \mathbf{W}_2\mathbf{H}_2$, i.e.,

$$\mathbf{W}_1\mathbf{H}_1 = \mathbf{W}_2\mathbf{H}_2. \tag{9}$$

The solutions satisfy (5), thus

$$\mathrm{rank}\,(\mathbf{W}_1) = \mathrm{rank}\,(\mathbf{H}_1) = \mathrm{rank}\,(\mathbf{W}_2) = \mathrm{rank}\,(\mathbf{H}_2) = r. \tag{10}$$

As $\mathbf{W}_1, \mathbf{W}_2$ have the same size, and (10) showed that they have the same rank, that means the two linear subspaces spanned by each $\mathbf{W}$ are linked by a rotation-and-scaling matrix $\mathbf{Q}$. In other words, there exists an invertible matrix $\mathbf{Q} \in \mathbb{R}^{r\times r}$ such that

$$\mathbf{W}_2 = \mathbf{W}_1\mathbf{Q} \text{ and } \mathbf{H}_2 = \mathbf{Q}\mathbf{H}_1 \tag{11}$$

We have now set up the relationship between $\mathbf{W}_1$, $\mathbf{W}_2$, and between $\mathbf{H}_1$, $\mathbf{H}_2$, which is based on the matrix $\mathbf{Q}$. Next we will use the model assumptions to show that $\mathbf{Q}$ is only a permutation matrix (thus there is no scaling degree of freedom in $\mathbf{Q}$).

The proof ... 3/7

By (10), $\mathbf{H}_1$ has rank $r \ll \min\{m, n\}$, so $\mathbf{H}_1$ has right inverse

$$\mathbf{H}_1 \mathbf{H}_1^{\dagger} = \mathbf{I}, \tag{12}$$

put this into right hand side of (11) gives

$$\begin{aligned} \mathbf{H}_2 \mathbf{H}_1^{\dagger} &= \mathbf{Q} \\ \mathbf{1}_r^{\top} \mathbf{H}_2 \mathbf{H}_1^{\dagger} &= \mathbf{1}_r^{\top} \mathbf{Q}. \end{aligned}$$

By (4), we have $\mathbf{1}_r^{\top} \mathbf{H}_1 = \mathbf{1}_r^{\top} \mathbf{H}_2 = \mathbf{1}_n^{\top}$, so

$$\mathbf{1}_r^{\top} \mathbf{H}_2 \mathbf{H}_1^{\dagger} \overset{(4)}{=} \mathbf{1}_n^{\top} \mathbf{H}_1^{\dagger} \overset{(4)}{=} \mathbf{1}_r^{\top} \underbrace{\mathbf{H}_1 \mathbf{H}_1^{\dagger}}_{\overset{(12)}{=} \mathbf{I}} = \mathbf{1}_r^{\top} \mathbf{Q}.$$

Now we have

$$\mathbf{1}_r^{\top} = \mathbf{1}_r^{\top} \mathbf{Q}, \tag{13}$$

so we get "some descriptions" on $\mathbf{Q}$. But we can go further, see next slide.

The proof ... 4/7

As both solutions satisfy (3), we have

$$\mathbf{H}_2 \overset{(11)}{=} \mathbf{Q}\mathbf{H}_1 \overset{(3)}{\geq} 0.$$

By the fact that $\mathbf{AB} \geq 0$ if $\mathrm{cone}(\mathbf{A}^\top) \subseteq \mathrm{cone}^*(\mathbf{B})$, then the above gives

$$\mathrm{cone}(\mathbf{Q}^\top) \subseteq \mathrm{cone}^*(\mathbf{H}_1). \tag{14}$$

As $\mathbf{H}_1$ satisfies (6), $\mathcal{C} \subseteq \mathrm{cone}(\mathbf{H}_1)$, by duality, $\mathrm{cone}^*(\mathbf{H}_1) \subseteq \mathcal{C}^*$, thus $\mathrm{cone}(\mathbf{Q}^\top) \subseteq \mathcal{C}^*$. Based on the definition of $\mathcal{C}^*$, we have

$$\langle \mathbf{Q}(j,:), \mathbf{1}_r \rangle \geq \|\mathbf{Q}(j,:)\|_2 \text{ for } j = 1, 2, \ldots, r \tag{15}$$

so we get "more descriptions" on $\mathbf{Q}$. We can now bound $|\det(\mathbf{Q})|$, see next slide.

The proof ... 5/7

A series of mathematical trick

$$
\begin{aligned}
|\det(\mathbf{Q})| &= |\det(\mathbf{Q}^\top)| \\
&\le \prod_{j=1}^{r} \|\mathbf{Q}(j,:)\|_2 && \text{Hardmard inequality} \\
&\le \prod_{j=1}^{r} \langle \mathbf{Q}(j,:), \mathbf{1}_r \rangle && \text{by (15)} \\
&\le \left(\frac{\sum_{j=1}^{r} \langle \mathbf{Q}(j,:), \mathbf{1}_r \rangle}{r} \right)^r && \text{AM-GM inequality} \\
&= \left(\frac{\mathbf{1}_r^\top \mathbf{Q} \mathbf{1}_r}{r} \right)^r \\
&= 1 && \text{by (13)}
\end{aligned}
$$

Hence we have

$$|\det(\mathbf{Q})| \le 1.$$

We have two cases, $|\det(\mathbf{Q})| = 1$ and $|\det(\mathbf{Q})| < 1$.

The proof ... 6/7 (The case $|\det(\mathbf{Q})| = 1$)

If $|\det(\mathbf{Q})| = 1$, all the inequalities in the last page are equalities. Hence

$$\|\mathbf{Q}(j,:)\|_2 = \langle \mathbf{Q}(j,:), \mathbf{1}_r \rangle.$$

This means $\mathbf{Q}(j,:)^\top \in \mathbf{bd}\, \mathcal{C}^*$, based on the definition of $\mathcal{C}^*$.

Recall from (14), $\mathrm{cone}(\mathbf{Q}^\top) \subseteq \mathrm{cone}^*(\mathbf{H}_1)$, hence we now have

$$\mathbf{Q}(j,:)^\top \in \mathbf{bd}\, \mathcal{C}^* \cap \mathrm{cone}^*(\mathbf{H}_1)$$

As $\mathbf{H}_1$ satisfies (7), so $\mathbf{bd}\, \mathcal{C}^* \cap \mathrm{cone}^*(\mathbf{H}_1)$ are the unit vectors so $|\det \mathbf{Q}| = 1$ meaning $\mathbf{Q}$ can only be a permutation matrix.

The proof ... 7/7 (The case $|\det(\mathbf{Q})| < 1$)

If $|\det(\mathbf{Q})| < 1$,

$$
\begin{aligned}
f(\mathbf{W}_2) &= \det(\mathbf{W}_2^\top \mathbf{W}_2) \\
&\overset{(11)}{=} \det\left((\mathbf{W}_1\mathbf{Q})^\top \mathbf{W}_1\mathbf{Q}\right) \\
&= |\det(\mathbf{Q})|^2 \det(\mathbf{W}_1^\top \mathbf{W}_1) \\
&< \det(\mathbf{W}_1^\top \mathbf{W}_1) = f(\mathbf{W}_1).
\end{aligned}
$$

This contradicts with (8), hence the solution of minvol-NMF is unique. □

End of document