A SHORT PROOF THAT NMF IS NP-HARD

YAROSLAV SHITOV

Abstract. We give a short combinatorial proof that the nonnegative matrix factorization is an NP-hard problem. Moreover, we prove that NMF remains NP-hard when restricted to 01-matrices, answering a recent question of Moitra.

The (exact) nonnegative matrix factorization is the following problem. Given an integer \( k \) and a matrix \( A \) with nonnegative entries, do there exist \( k \) nonnegative rank-one matrices that sum to \( A \)? The smallest \( k \) for which this is possible is called the nonnegative rank of \( A \) and denoted by \( \text{rank}^+(A) \). We give a short combinatorial proof of a seminal result of Vavasis [6] stating that NMF is NP-hard. Moreover, we prove that NMF remains hard when restricted to Boolean matrices, answering a recent question of Moitra [4].

Theorem 1. It is NP-hard to decide whether \( \text{rank}^+(A) \leq k \), given an integer \( k \) and a matrix \( A \) with entries in \( \{0, 1\} \).

Recall that a (directed) graph \( G \) is a finite set of vertices \( V \) and edges \( E \subset V \times V \). We assume that \( G \) has no loops, that is, \( (v, v) \notin E \) for all \( v \in V \). An independent set in \( G \) is a subset \( U \subset V \) such that \( (u_1, u_2) \notin E \) for all \( u_1, u_2 \in U \). The chromatic number of \( G \), denoted by \( c(G) \), is the smallest \( c \) such that \( V \) is a union of \( c \) independent sets. The following is a classical NP-complete problem [3].

Problem 2. Given a graph \( G \) and an integer \( C \). Is \( c(G) \leq C \)?

To construct a reduction from Problem 2 to NMF, we define the matrix \( \mathcal{N} = \mathcal{N}(G) \) with \( 5|V| \) rows and columns indexed by the set \( V \cup V^1 \cup V^2 \cup V^3 \cup V^4 \), which is the union of five copies of \( V \). For any \( v \in V \), we define the entry \( \mathcal{N}(v|v) \) as 1 and enumerate the vertices in \( V \setminus \{v\} \) as \( u_1, \ldots, u_m \); we set the submatrix \( \mathcal{N}(v^1, v^2, v^3, v^4, v|v^1, v^2, v^3, v^4, u_1, \ldots, u_m) \) equal to

\[
\begin{pmatrix}
1 & 1 & 0 & 0 & 1 & \ldots & 1 \\
0 & 1 & 1 & 0 & 1 & \ldots & 1 \\
0 & 0 & 1 & 1 & 1 & \ldots & 1 \\
1 & 0 & 0 & 1 & 1 & \ldots & 1 \\
1 & 0 & 0 & 0 & x_1 & \ldots & x_m
\end{pmatrix},
\]

where \( x_i = 0 \) if \( (v, u_i) \in E \) and \( x_i = 1 \) otherwise. The entries of \( \mathcal{N} \) that are not yet specified are equal to 0.

We denote by \( \mathcal{N} \) the upper left \( 4 \times 4 \) submatrix of (0.1); one has \( \text{rank}^+(\mathcal{N}) = 4 \). Since every column of (0.1) is a linear combination of the first four columns taken with nonnegative coefficients, the nonnegative rank of (0.1) equals four.

2010 Mathematics Subject Classification. 15A23.
Key words and phrases. Nonnegative matrix factorization, NP-hard problem.
Observation 3. Let $M$ be a nonnegative rank-one matrix such that $M \leq N(G)$. If $M(v|v) \neq 0$ for some $v \in V$, then $M(u|u^j) = 0$ for all $u \in V$ and $i, j \in \{1, 2, 3, 4\}$.

Proof. By the construction, the entry $N(v|u^j)$ can be nonzero only if $u = v$, but in this case we have $N(u|v) = 0$. Since $M \leq N$, the entries $M(v|u^j)$ and $M(u|v)$ cannot be positive simultaneously, and the same holds for $M(v|v)$ and $M(u|u^j)$ because $M$ is rank-one. \hfill \Box

Proposition 4. We have $\text{rank}_+(N(G)) = 4|V(G)| + c(G)$.

Proof. Let $U_1, \ldots, U_c$ be a partition of $V$ into disjoint independent sets of $G$. Let $H_i$ be the matrix such that $H_i(\alpha|\beta) = 1$ if $\alpha, \beta \in U_i$ and $H_i(\alpha|\beta) = 0$ otherwise. We see that $N - H_1 - \ldots - H_c$ is a nonnegative matrix whose nonzero entries are contained in $|V|$ disjoint submatrices of the form $(0, 1)$. Since $\text{rank}_+(H_i) = 1$, we get $\text{rank}_+(N) \leq 4|V| + c$.

Now let $M_1, \ldots, M_r$ be nonnegative rank-one matrices that sum to $N$. Since the set $C_j = \{v \in V : M_j(v|v) \neq 0\}$ is independent for every $j$, this set is non-empty for at least $c(G)$ values of $j$. Observation 3 shows that, for these $j$, the submatrices $M_j(V^1 \cup V^2 \cup V^3 \cup V^4)$ are zero. It remains to note that $N(V^1 \cup V^2 \cup V^3 \cup V^4)$ has nonnegative rank $4|V|$ because it is the block-diagonal matrix with $|V|$ blocks equal to $N$. \hfill \Box

Now we see that $(G, C) \rightarrow (N(G), 4|V(G)| + C)$ is a polynomial reduction from Problem 2 to NMF. Since $N(G)$ is Boolean, the proof of Theorem 1 is complete.

Many interesting problems regarding the complexity of NMF remain open. Let us recall a remarkable result [1] providing a polynomial time algorithm for NMF with fixed nonnegative rank. However, it is not known whether such algorithm exists if we fix the conventional rank instead of nonnegative rank [2].

Despite having proved that NMF is NP-hard, we do not know anything about completeness of this problem. We note that the entries of rank-one matrices in the optimal factorization may not be rational functions of entries of the initial matrix, see [3]. Is NMF (restricted to rational matrices) NP-complete or \exists R-complete?

References

[1] S. Arora, R. Ge, R. Kannan, A. Moitra, Computing a nonnegative matrix factorization — probably, Proceedings of the Annual Symposium on Theory of Computing (2012) 145–162.
[2] N. Gillis, F. Glineur, On the geometric interpretation of the nonnegative rank, Linear Algebra Appl. 437 (2012) 2685–2712.
[3] R. Karp, Reducibility Among Combinatorial Problems, Proceedings of the Symposium on the Complexity of Computer Computations (1972) 85–103.
[4] A. Moitra, Nonnegative Matrix Factorization: Algorithms, Complexity and Applications. A presentation at the International Symposium on Symbolic and Algebraic Computation (2015). Available via [http://www.issac-conference.org/2015/Slides/Moitra.pdf](http://www.issac-conference.org/2015/Slides/Moitra.pdf)
[5] Y. Shitov, Nonnegative rank depends on the field, preprint (2015) [arXiv:1505.01893](http://arxiv.org/abs/1505.01893)
[6] S. Vavasis, On the complexity of nonnegative matrix factorization, SIAM J. Optimization 20 (2009) 1364–1377.

National Research University Higher School of Economics, 20 Myasnitskaya Ulitsa, Moscow 101000, Russia
E-mail address: yaroslav-shitov@yandex.ru