Constraint-Induced Symmetric Nonnegative Matrix Factorization for Accurate Community Detection: Supplementary File

Zhigang Liu[1,2], Xin Luo[3,*], and Zidong Wang[4]

[1]*College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China*
[2]*Chongqing Key Laboratory of Big Data and Intelligent Computing, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, 400714, China*
[3]*College of Computer and Information Science, Southwest University, Chongqing 400715, China*
[4]*Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom*

1. Introduction

This is the supplementary file for the paper entitled "*Constraint-Induced Symmetric Nonnegative Matrix Factorization for Accurate Community Detection*". In this file, we provide the convergence proof of the proposed model with the learning scheme regarding ***Updating U***.

2. Preliminaries

2.1. The value of $\mathcal{J}(U, X)$ in (8) is non-increasing when updating U by the learning rule (13a)

Following the proof sketch in the manuscript, we start with considering an arbitrary entry u_{ik} in U, and adopting $F(u_{ik})$ to denote the corresponding component from $\mathcal{J}(U, X)$ that is relevant to u_{ik} only. Thus, its first-order and second-order derivatives with respect to u_{ik} is computed as:

$$F'(u_{ik}) = \left(-AX + (1+\mu)UX^{\mathrm{T}}X - \mu XU^{\mathrm{T}}X\right)_{ik}, \tag{S1}$$

$$F''(u_{ik}) = (1+\mu)\left(X^{\mathrm{T}}X\right)_{kk} - \mu x_{ik}^2. \tag{S2}$$

We also proportionally omit all the constant coefficient in (S1) and (S2), which will not impact the proof. Next, we design a rational auxiliary function as:

***Lemma* S1.** The following $\Gamma(u_{ik}, u_{ik}^{(t)})$ is an auxiliary function for $F(u_{ik})$

$$\Gamma\left(u_{ik}, u_{ik}^{(t)}\right) = F\left(u_{ik}^{(t)}\right) + F'\left(u_{ik}^{(t)}\right)\left(u_{ik} - u_{ik}^{(t)}\right) + \frac{\left((1+\mu)UX^{\mathrm{T}}X\right)_{ik}}{2u_{ik}^{(t)}}\left(u_{ik} - u_{ik}^{(t)}\right)^2. \tag{S3}$$

Proof. Note that $\Gamma(u_{ik}, u_{ik})=F(u_{ik})$ evidently holds according to (S3). Thus, we need to prove that $\Gamma(u_{ik}, u_{ik}^{(t)})\geq F(u_{ik})$ also holds. To achieve this objective, we expand $F(u_{ik})$ into its second-order Taylor series at the state point $u_{ik}^{(t)}$ as:

$$F(u_{ik}) = F\left(u_{ik}^{(t)}\right) + F'\left(u_{ik}^{(t)}\right)\left(u_{ik} - u_{ik}^{(t)}\right) + \frac{1}{2}\left((1+\mu)\left(X^{\mathrm{T}}X\right)_{kk} - \mu x_{ik}^2\right)\left(u_{ik} - u_{ik}^{(t)}\right)^2. \tag{S4}$$

By combining (S3) and (S4), we see that the desired condition of $\Gamma(u_{ik}, u_{ik}^{(t)})\geq F(u_{ik})$ is equivalent to the following condition:

$$\frac{\left((1+\mu)UX^{\mathrm{T}}X\right)_{ik}}{u_{ik}^{(t)}} \geq (1+\mu)\left(X^{\mathrm{T}}X\right)_{kk} - \mu x_{ik}^2. \tag{S5}$$

Since $u_{ik}^{(t)}\geq 0$[1], (S5) is transformed into:

$$\left((1+\mu)UX^{\mathrm{T}}X\right)_{ik} \geq (1+\mu)u_{ik}\left(X^{\mathrm{T}}X\right)_{kk} - \mu x_{ik}^2 u_{ik}^{(t)}. \tag{S6}$$

Note that it is easy to prove that

[1] Please note that if $u_{ik}^{(t)}$ goes to zero after the t-th iteration, then it will keep zero constantly with the learning scheme (13a) in the follow-up iterations, which will not affect the convergence of the learning algorithm. Therefore, without loss of generality, we only take into account the case of $u_{ik}^{(t)}>0$ in our proof.

$$\left(UX^{\mathrm{T}}X\right)_{ik}=\sum_{p=1}^{K}\left(u_{ip}\left(X^{\mathrm{T}}X\right)_{pk}\right)=\sum_{p=1,\,p\neq k}^{K}\left(u_{ip}\left(X^{\mathrm{T}}X\right)_{pk}\right)+u_{ik}\left(X^{\mathrm{T}}X\right)_{kk}\geq u_{ik}\left(X^{\mathrm{T}}X\right)_{kk}. \tag{S7}$$

Besides, since $u_{ik}\geq0$, $x_{ik}\geq0$ and $\mu>0$, we obviously have that $\mu x_{ik}^2 x_{ik}\geq0$. By substituting this into (S7), we conclude that (S6) holds, resulting in the establishment of (S5). Hence, we have proven that $\Gamma(u_{ik}, u_{ik}^{(t)})\geq F(u_{ik})$, which makes $\Gamma(u_{ik}, u_{ik}^{(t)})$ be an auxiliary function for $F(u_{ik})$.□

With *Lemmas* 1 in the manuscript and *Lemma* S1, we further obtain the following theorem:

***Theorem* S1.** The value of $\mathcal{J}(U, X)$ in (8) in the manuscript keeps non-increasing when updating U by the learning scheme (13a).

Proof. It is equivalent to proving that the learning scheme (13a) is identical to that in (14) in the manuscript with the designed auxiliary function (S3). Hence, we replace the auxiliary function Γ in (14) with the one in (S3), and achieve:

$$\begin{aligned}
&u_{ik}^{(t+1)}=\arg\min_{u_{ik}}\Gamma\left(u_{ik},u_{ik}^{(t)}\right)\\
&\Rightarrow\Gamma'\left(u_{ik},u_{ik}^{(t)}\right)=F'\left(u_{ik}^{(t)}\right)+\frac{\left((1+\mu)UX^{\mathrm{T}}X\right)_{ik}}{u_{ik}^{(t)}}\left(u_{ik}-u_{ik}^{(t)}\right)=0\\
&\Rightarrow u_{ik}=u_{ik}^{(t)}-u_{ik}^{(t)}\frac{F'\left(u_{ik}^{(t)}\right)}{\left((1+\mu)UX^{\mathrm{T}}X\right)_{ik}}\\
&\Rightarrow u_{ik}^{(t+1)}=u_{ik}^{(t)}\frac{\left(AX+\mu XU^{\mathrm{T}}X\right)_{ik}}{\left((1+\mu)UX^{\mathrm{T}}X\right)_{ik}}.
\end{aligned} \tag{S8}$$

Note that the updating rule in (S8) is identical to the one in (13a). Similarly, $\forall i\in\{1, 2, \ldots, n\}$ and $k\in\{1, 2, \ldots, K\}$, $F(u_{ik})$ always keeps non-increasing under the learning scheme (13a) as X is alternatively fixed. Hence, we have shown that $\mathcal{J}(U, X)$ in (8) keeps non-increasing when updating U by the learning rule (13a), and *Theorem* S1 stands.■

2.2. Proof of Theorem 2 Related to U

We start with making $u_{ik}^{(*)}=\lim_{t\to+\infty}u_{ik}^{(t)}$, and the following KKT conditions are desired:

$$\begin{aligned}
&\forall i\in\{1,2,\ldots,n\},k\in\{1,2,\ldots,K\}:\\
&(\mathrm{a})\ \left.\frac{\partial\mathcal{L}(U,X)}{\partial u_{ik}}\right|_{u_{ik}=u_{ik}^{(*)}}=\left(-AX+(1+\mu)U^{(*)}X^{\mathrm{T}}X-\mu X\left(U^{(*)}\right)^{\mathrm{T}}X-\Phi^{(*)}\right)_{ik}=0,\\
&(\mathrm{b})\ \phi_{ik}^{(*)}u_{ik}^{(*)}=0,\\
&(\mathrm{c})\ u_{ik}^{(*)}\geq0,\\
&(\mathrm{d})\ \phi_{ik}^{(*)}\geq0.
\end{aligned} \tag{S9}$$

Based on the deduction in (8)-(13), we can see that condition (a) constantly holds with (13a). Then, we have:

$$\phi_{ik}^{(*)}=\left(-AX+(1+\mu)U^{(*)}X^{\mathrm{T}}X-\mu X\left(U^{(*)}\right)^{\mathrm{T}}X\right)_{ik}. \tag{S10}$$

To analyze conditions (b)-(d), we present an auxiliary variable, i.e.,

$$\tau_{ik}^{(t)}=\frac{\left(AX+\mu X\left(U^{(t)}\right)^{\mathrm{T}}X\right)_{ik}}{\left((1+\mu)U^{(t)}X^{\mathrm{T}}X\right)_{ik}}. \tag{S11}$$

Thus, the learning rule for u_{ik} is simplified as:

$$u_{ik}^{(t+1)}=u_{ik}^{(t)}\frac{\left(AX+\mu X\left(U^{(t)}\right)^{\mathrm{T}}X\right)_{ik}}{\left((1+\mu)U^{(t)}X^{\mathrm{T}}X\right)_{ik}}=u_{ik}^{(t)}\tau_{ik}^{(t)}. \tag{S12}$$

In our context, entries in A, U and X are all equal or greater than zero, which yields the following bound of $\tau_{ik}^{(t)}$ as:

$$0\leq\tau_{ik}^{(*)}=\lim_{t\to+\infty}\tau_{ik}^{(t)}=\frac{\left(AX+\mu X\left(U^{(*)}\right)^{\mathrm{T}}X\right)_{ik}}{\left((1+\mu)U^{(*)}X^{\mathrm{T}}X\right)_{ik}}<+\infty. \tag{S13}$$

Based on (27) in the manuscript and (S13), we obtain:

$$\lim_{t\to+\infty}\left|u_{ik}^{(t+1)}-u_{ik}^{(t)}\right|=0\Rightarrow u_{ik}^{(*)}\tau_{ik}^{(*)}-u_{ik}^{(*)}=0. \tag{S14}$$

It should be pointed that with the learning rule (13a), we have that $u_{ik}^{(*)}$. Thus, we discuss the following two cases.

(a) $\boldsymbol{u_{ik}^{(*)}}$>0. By combining (S10), (S13) and (S14), we have the following inferences:

$$\begin{aligned}
&u_{ik}^{(*)}\tau_{ik}^{(*)}-u_{ik}^{(*)}=0,u_{ik}^{(*)}>0\Rightarrow\tau_{ik}^{(*)}=1,\\
&\Rightarrow\left(AX+\mu X\left(U^{(*)}\right)^{\mathrm{T}}X\right)_{ik}=\left((1+\mu)U^{(*)}X^{\mathrm{T}}X\right)_{ik},\\
&\Rightarrow\phi_{ik}^{(*)}=\left((1+\mu)U^{(*)}X^{\mathrm{T}}X\right)_{ik}-\left(AX+\mu X\left(U^{(*)}\right)^{\mathrm{T}}X\right)_{ik}=0,\Rightarrow\phi_{ik}^{(*)}u_{ik}^{(*)}=0.
\end{aligned} \tag{S15}$$

Form (S15), conditions (b), (c) and (d) in (S9) are all fulfilled. Naturally, we have proven that the KKT conditions in (S9) constantly hold when $u_{ik}^{(*)}$>0.

(b) $\boldsymbol{u_{ik}^{(*)}}$=0. In this case, conditions (b) and (c) are constantly fulfilled. So we need to prove condition (d) in (S9) is also fulfilled. To do this, we firstly expand $u_{ik}^{(*)}$ into the following form:

$$u_{ik}^{(*)}=u_{ik}^{(0)}\lim_{t\to+\infty}\prod_{q=1}^{t}\tau_{ik}^{(q)}. \tag{S16}$$

As discussed in part a) of the convergence proof, we consider $u_{ik}^{(0)}$>0 without loss of generality. Note that if make $u_{ik}^{(0)}$=0, it will keep zero constantly with the learning rule (13a) in the follow-up iterations, which leads to poor representation to *A*. Hence, with $u_{ik}^{(0)}$>0, we have the following inferences:

$$u_{ik}^{(0)}>0,u_{ik}^{(*)}=u_{ik}^{(0)}\lim_{t\to+\infty}\prod_{q=1}^{t}\tau_{ik}^{(q)}=0\Rightarrow\lim_{t\to+\infty}\prod_{q=1}^{t}\tau_{ik}^{(q)}=0. \tag{S17}$$

Furthermore, part of *A*'s entries should be one to describe a meaningful network, making the numerator of $\tau_{ik}^{(t)}$ as shown in (S11) can be greater than zero. Based on (S17), we further achieve:

$$\begin{aligned}
&\tau_{ik}^{(q)}>0,\lim_{t\to+\infty}\prod_{q=1}^{t}\tau_{ik}^{(q)}=0\Rightarrow\lim_{t\to+\infty}\tau_{ik}^{(t)}=\tau_{ik}^{(*)}<1,\\
&\Rightarrow\left((1+\mu)U^{(*)}X^{\mathrm{T}}X\right)_{ik}-\left(AX+\mu X\left(U^{(*)}\right)^{\mathrm{T}}X\right)_{ik}=\phi_{ik}^{(*)}>0.
\end{aligned} \tag{S18}$$

Note that (S18) indicates that condition (d) in (S9) holds. Hence, we have evidently proven that the KKT conditions in (S9) are all fulfilled when $u_{ik}^{(*)}$=0.■