A sharp NMF result with applications in network modeling

Jiashun Jin

Statistics Department
Carnegie Mellon University

October 22, 2022

An NMF problem

NMF: Non-negative Matrix Factorization

Let Ω be an (entry-wise) non-negative matrix satisfying

$$\underbrace{\Omega}_{n\times n} = \underbrace{Y}_{n\times K} \underbrace{J_{K,m}}_{K\times K} \underbrace{Y'}_{K\times n}, \qquad K = \mathrm{rank}(\Omega),$$

where $J_{K,m} = \mathrm{diag}(\underbrace{1,\ldots,1}_{K-m}, \underbrace{-1,\ldots,-1}_{m})$

Problem. Is there an orthogonal matrix Q such that both YQ and $Q'J_{K,m}Q$ are (entry-wise) non-negative? If so, then

$$\Omega = YJ_{K,m}Y' = \underbrace{YQ}_{\text{non-negative}} \underbrace{Q'J_{K,m}Q}_{\text{non-negative}} \underbrace{(YQ)'}_{\text{non-negative}}$$

A sharp NMF result

$$\text{Write} \qquad Y = \begin{bmatrix} y_1' \\ y_2' \\ \ldots \\ y_n' \end{bmatrix}$$

- When $m > K/2$, no such Q exists (so either the NMF problem is not solvable or we need a different approach)
- When $m \leq K/2$ and

$$\sum_{k=1}^{K-1} y_i^2(k+1) \leq y_i^2(1)/(K-1),$$

the desired Q exists so the NMF problem is solvable

See the paper for why the result is sharp and also for results in more general settings

Social networks

Data: $n \times n$ adjacency matrix A (symmetric)

$$A(i,j) = \begin{cases} 1, & \text{an edge between nodes } i \ \& \ j, \\ 0, & \text{otherwise}, \end{cases} \qquad 1 \leq i \neq j \leq n$$

Assume there are K perceivable "communities"

$$\mathcal{C}_1, \ldots, \mathcal{C}_K,$$

For a rank-K matrix Ω, the upper triangle entries of A are independent Bernoulli:

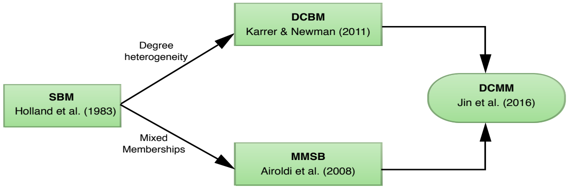
$$A = \underbrace{\Omega - \text{diag}(\Omega)}_{\text{"signal"}} + \underbrace{W}_{\text{"noise"}}, \quad \Omega_{ij} = \mathbb{P}(A(i,j) = 1)$$

The DCMM model for networks

$$\Omega = \Theta \Pi P \Pi' \Theta, \quad \Theta = \mathrm{diag}(\theta_1, \ldots, \theta_n), \quad \Pi = [\pi_1, \ldots, \pi_n]'$$

- $\theta_i > 0$: degree parameters for author i
- $\pi_i \in \mathbb{R}^K$: membership for node i:

  $\pi_i(k)$ = weight he/she has in community k, $\quad 1 \leq k \leq K$
- $P \in \mathbb{R}^{K,K}$: baseline connectivity between different communities (symmetric, nonnegative, unit diagonals)



An example

$$A = \Omega - \text{diag}(\Omega) + W], \quad \text{where} \quad \Omega = \Theta \Pi P \Pi' \Theta$$

$$\begin{bmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{bmatrix} \begin{bmatrix} 1 & 0 \\ .5 & .5 \\ \vdots & \vdots \\ .1 & .9 \end{bmatrix} \begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix} \begin{bmatrix} 1 & .5 \ldots & .1 \\ 0 & .5 \ldots & .9 \end{bmatrix} \begin{bmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{bmatrix}$$

Network modeling: a problem

- A rank-K model for networks is broad and only assumes $\mathrm{rank}(\Omega) = K$. It includes DCMM and many others (e.g., Random dot product model) as special cases
- However, we prefer to use a DCMM model, for it is practically more interpretable (note that (Θ, Π, P) all have practical meanings)

Problem. When a rank-K model can be rewritten as a DCMM model?

A rank-K model is almost a DCMM model

$$\Omega = \sum_{k=1}^{n} \lambda_k \xi_k \xi_k' : \qquad SVD$$

A rank-K model can be rewritten as a DCMM model if

- At least half of $\{\lambda_1, \ldots, \lambda_K\}$ are negative (it is possible that the condition can be removed)
- For each $1 \leq i \leq K$,

$$r_i' \,\mathrm{diag}(|\lambda_2|, \ldots, |\lambda_K|)\, r_i \leq |\lambda_1|/(K-1),$$

where $r_i(k) = \xi_{k+1}(i)/\xi_1(i)$, $1 \leq k \leq K-1$

For many networks, these conditions hold as

$$\max_{1 \leq i \leq n} \{\|r_i\|\} \leq C, \qquad \max_{2 \leq k \leq K} \{|\lambda_k|/|\lambda_1|\} \to 0,$$

Take home message

- Presented a sharp NMF result
- Applied it to network modeling and showed that we only need mild regularity conditions to rewrite a rank-K network model as a DCMM model

Jin, J (2022). A sharp NMF result with applications in network modeling