T-product factorization based method for matrix and tensor completion problems

Xinzhen Zhang (TJU)

Joint work with Quan Yu

Central China Normal University, 2022.06.06

Contents

Contents

Matrix completion

Research background

Low rank matrix completion:

$$\min_{X \in \mathbb{R}^{m \times n}} \operatorname{rank}(X), \quad \text{s.t.} \quad P_\Omega(X - M) = 0. \tag{1}$$

Problem (1) is NP-hard to solve.

- Relaxation method: nuclear norm, Schatten p-norm, truncated nuclear norm, etc.——the SVD of the matrix needs to be calculated, which is computationally very expensive.
- Matrix factorization: $X = PQ^T$——the rank r of the matrix is pre-estimated.

Tensor completion

Research background

Low rank tensor completion problem:

$$\min_{\mathcal{X}\in\mathbb{R}^{n_1\times\cdots\times n_m}} \operatorname{rank}(\mathcal{X}), \quad \text{s.t.} \quad P_\Omega(\mathcal{X}-\mathcal{M})=0. \qquad (2)$$

There are various definitions of tensor rank: CP rank, Tucker rank, Tubal rank, etc.

- CP rank

$$\operatorname{rank}_{cp}(\mathcal{X}) = \min\left\{ r \mid \mathcal{X} = \sum_{i=1}^{r} a_1^{(i)} \otimes a_2^{(i)} \otimes \cdots \otimes a_m^{(i)} \right\} \quad (3)$$

——Computing the CP rank is NP-hard.

- Tucker rank

$$\operatorname{rank}_{tc}(\mathcal{X}) = \left(\operatorname{rank}\left(X_{(1)}\right), \cdots, \operatorname{rank}\left(X_{(m)}\right)\right) \quad (4)$$

——Unfolding a tensor would destroy the original multi-way structure of the data.

- Tubal rank

$$\operatorname{rank}_{t}(\mathcal{X}) = \max\left\{ \operatorname{rank}\left(\bar{X}^{(1)}\right), \cdots, \operatorname{rank}\left(\bar{X}^{(n_3)}\right) \right\} \quad (5)$$

where $\bar{X}^{(i)} = \bar{\mathcal{X}}(:,:,i)$, $\bar{\mathcal{X}} = fft(\mathcal{X}, [], 3)$.
——Fourier transform is performed only for a third order tensors.

Contents

Contents

Model:

$$\min_{X \in \mathbb{R}^{n_1 \times h}} \ \mathrm{rank}(X), \quad \text{s.t.} \quad P_{\tilde{\Omega}}(X - M) = 0. \tag{6}$$

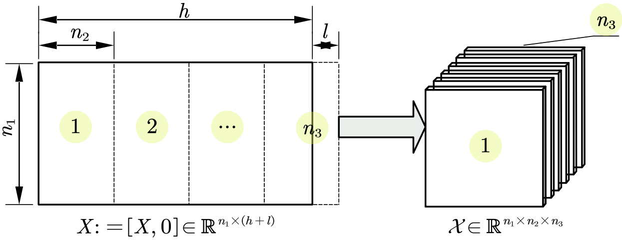

Figure 1: Reshaping the matrix X into the tensor $\mathcal{X}$

Theorem 1

Suppose that matrix $X \in \mathbb{R}^{n_1 \times h}$ and tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ obtained by reshaping matrix X with Figure 1. Then

$$\begin{aligned} &\operatorname{rank}_t(\mathcal{X}) \leq \operatorname{rank}(X) \leq n_3 \operatorname{rank}_t(\mathcal{X}), \\ &\operatorname{rank}(X) \leq \left\|\operatorname{rank}_m(\mathcal{X})\right\|_1 \leq n_3 \operatorname{rank}(X). \end{aligned} \tag{7}$$

Based on Theorem 1, we consider the following tensor completion problem for solving the matrix completion problem.

$$\min_{\mathcal{X}\in\mathbb{R}^{n_1\times n_2\times n_3}} \operatorname{rank}_t(\mathcal{X}), \quad \text{s.t.} \quad P_\Omega(\mathcal{X}-\mathcal{M})=0. \tag{8}$$

We consider the following tensor factorization model[1] to solve (8).

$$\min_{\mathcal{X},\mathcal{P},\mathcal{Q}} \frac{1}{2}\|\mathcal{P}*\mathcal{Q}-\mathcal{X}\|_F^2, \quad \text{s.t.} \quad P_\Omega(\mathcal{X}-\mathcal{M})=0. \tag{9}$$

[1] Pan Zhou et al. "Tensor Factorization for Low-Rank Tensor Completion". In: *IEEE Transactions on Image Processing* 27.3 (Mar. 2018), pp. 1152–1163.

$$\mathcal{X} = \underset{P_\Omega(\mathcal{X}-\mathcal{M})=0}{\operatorname{argmin}} \frac{1}{2}\|\mathcal{P} * \mathcal{Q} - \mathcal{X}\|_F^2 = P_{\Omega^c}(\mathcal{P} * \mathcal{Q}) + P_\Omega(\mathcal{M}). \tag{10}$$

$$\hat{P}^{(k)} = \begin{cases} \bar{X}^{(k)}\left(\hat{Q}^{(k)}\right)^*\left(\hat{Q}^{(k)}\left(\hat{Q}^{(k)}\right)^*\right)^\dagger, \; k = 1, \ldots, \left\lceil \frac{n_3+1}{2} \right\rceil, \\ conj\left(\hat{P}^{(n_3-k+2)}\right), \; k = \left\lceil \frac{n_3+1}{2} \right\rceil + 1, \ldots, n_3, \end{cases} \tag{11}$$

$$\hat{Q}^{(k)} = \begin{cases} \left(\left(\hat{P}^{(k)}\right)^*\hat{P}^{(k)}\right)^\dagger\left(\hat{P}^{(k)}\right)^*\bar{X}^{(k)}, \; k = 1, \ldots, \left\lceil \frac{n_3+1}{2} \right\rceil, \\ conj\left(\hat{Q}^{(n_3-k+2)}\right), \; k = \left\lceil \frac{n_3+1}{2} \right\rceil + 1, \ldots, n_3. \end{cases} \tag{12}$$

Algorithm 1: Matrix Completion Algorithm (TCTF-M)

Input: The matrix (tensor) data $M \in \mathbb{R}^{n_1 \times h}$ ($\mathcal{M} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$), the observed set $\tilde{\Omega}\,(\Omega)$ and t_0.

Input: $\mathcal{X}^0$, $\hat{P}^0$, $\hat{Q}^0$ and the multi-rank $\boldsymbol{r}^0_{\mathcal{X}} \in \mathbb{R}^{n_3}$.

While not converge do

1. Fix $\hat{Q}^t$ and $\mathcal{X}^t$ to update $\hat{P}^{t+1}$ by (11).

2. If $t \leq t_0$ then

Fix $\hat{P}^{t+1}$ and $\hat{Q}^t$ to compute $\mathcal{X}^t$ by (10).

3. Fix $\hat{P}^{t+1}$ and $\mathcal{X}^t$ to update $\hat{Q}^{t+1}$ by (12).

4. Adopt the rank decreasing scheme to adjust $\boldsymbol{r}^t_{\mathcal{X}}$, adjust the sizes of $\hat{P}^{t+1}$, $\hat{Q}^{t+1}$.

5. Fix $\hat{P}^{t+1}$ and $\hat{Q}^{t+1}$ to compute $\mathcal{X}^{t+1}$ by (10).

6. Check the stop criterion: $\left\|\mathcal{X}^{t+1} - \mathcal{X}^t\right\|_F / \left\|\mathcal{X}^t\right\|_F < \varepsilon$.

7. $t \leftarrow t + 1$.

end while

Output: $\mathcal{X}^{t+1}$.

Contents

Model:

$$\min_{\mathcal{X}\in\mathbb{R}^{n_1\times n_2\times n_3}} \operatorname{rank}_t(\mathcal{X}), \quad \text{s.t.} \quad P_\Omega(\mathcal{X}-\mathcal{M})=0. \tag{13}$$

Lemma 1

For a tensor $\mathcal{X}\in\mathbb{R}^{n_1\times n_2\times n_3}$, it holds

$$\operatorname{rank}_t(\mathcal{X}) \leq \operatorname{rank}\left(X_{(i)}\right) \leq n_3 \operatorname{rank}_t(\mathcal{X}), \quad i=1,2. \tag{14}$$

Compared to Tucker rank, tubal rank does not involve the low rank structure information of the mode-3 unfolding matrix from Lemma 1. Hence, we define an improved tensor rank as follows:

$$\operatorname{rank}_{ttr}(\mathcal{X}) = \left(\operatorname{rank}_t(\mathcal{X}), \operatorname{rank}(X_{(3)})\right). \tag{15}$$

We change (15) into double tubal rank:

$$\operatorname{rank}_{dt}(\mathcal{X}) = \left(\operatorname{rank}_t(\mathcal{X}), \operatorname{rank}_t(\tilde{\mathcal{X}})\right), \tag{16}$$

where $\tilde{\mathcal{X}} \in \mathbb{R}^{n_3 \times p \times q}$ $(pq = n_1 n_2)$ satisfying $\tilde{X}_{(1)} = X_{(3)}$.

Lemma 2 (The relationship between double tubal rank and 3-tubal rank.)

For a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, we have

$$\operatorname{rank}_t(\tilde{\mathcal{X}})/n_2 \leq \operatorname{rank}_t(\mathcal{X}_{(13)}) \leq q \operatorname{rank}_t(\tilde{\mathcal{X}}),$$
$$\operatorname{rank}_t(\tilde{\mathcal{X}})/n_1 \leq \operatorname{rank}_t(\mathcal{X}_{(23)}) \leq q \operatorname{rank}_t(\tilde{\mathcal{X}}).$$

In particular, when $\tilde{\mathcal{X}} \in \mathbb{R}^{n_3 \times n_1 \times n_2}$, $\operatorname{rank}_t(\tilde{\mathcal{X}}) = \operatorname{rank}_t(\mathcal{X}_{(13)})$.

The low double tubal rank tensor completion problem can be modeled as

$$\min_{\mathcal{X}\in\mathbb{R}^{n_1\times n_2\times n_3}} \ \mathrm{rank}_{dt}(\mathcal{X}), \quad \text{s.t.} \quad P_\Omega(\mathcal{X}-\mathcal{M})=0. \tag{17}$$

To keep things simple, we consider the follow problem:

$$\min_{\mathcal{X}} \ \gamma_1 \, \mathrm{rank}_t(\mathcal{X}) + \gamma_2 \, \mathrm{rank}_t(\tilde{\mathcal{X}}), \quad \text{s.t.} \quad P_\Omega\left(\mathcal{X}-\mathcal{M}\right)=0. \tag{18}$$

Clearly, (18) reduces to the classical low tubal rank tensor completion model when $\gamma_1 = 1$ and $\gamma_2 = 0$.

We consider the following tensor factorization model

$$\min \ \frac{\gamma_1}{2}\|\mathcal{P}*\mathcal{Q}-\mathcal{X}\|_F^2+\frac{\gamma_2}{2}\left\|\mathcal{U}*\mathcal{V}-\tilde{\mathcal{X}}\right\|_F^2, \quad \text{s.t.} \quad P_\Omega(\mathcal{X}-\mathcal{M})=0. \tag{19}$$

Motivated by the reweighted strategies[2] and the supergradient concepts[3], problem (19) can be derived

$$\begin{aligned}&\min \frac{1}{2}\rho\left(\|\mathcal{P}*\mathcal{Q}-\mathcal{X}\|_F^2\right)+\frac{1}{2}\rho\left(\left\|\mathcal{U}*\mathcal{V}-\tilde{\mathcal{X}}\right\|_F^2\right)\\ &\text{s.t.} \quad P_\Omega(\mathcal{X}-\mathcal{M})=0.\end{aligned} \tag{20}$$

[2] Canyi Lu et al. "Generalized Nonconvex Nonsmooth Low-Rank Minimization". In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2014. DOI: 10.1109/cvpr.2014.526.

[3] KC Border. *The Supergradient of a Concave Function*. 2001. URL: https://healy.econ.ohio-state.edu/kcb/Notes/Supergrad.pdf.

Assumption 1

The function $\rho(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ is a proper, concave, lower semicontinuous function on $[0, +\infty)$, and there exists $a, b > 0$ such that $\partial\rho(t) \subset [a, b]$ for any $t \in [0, +\infty)$.

Remark

Since $\rho(\cdot)$ is concave on $[0, +\infty)$, by the definition of the supergradient, for any s and t, we have

$$\rho(t) \leq \rho(s) + w_s(t - s), \ \forall\, w_s \in \partial\rho(s).$$

Now, we are ready to update $\mathcal{X}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{U}$, $\mathcal{V}$. First of all, by Assumption 1, we can update $\mathcal{X}$ by

$$\begin{aligned}\mathcal{X} &= \underset{P_\Omega(\mathcal{X}-\mathcal{M})=0}{\operatorname{argmin}} \frac{\gamma_1}{2}\|\mathcal{P}*\mathcal{Q}-\mathcal{X}\|_F^2 + \frac{\gamma_2}{2}\left\|\mathcal{U}*\mathcal{V}-\tilde{\mathcal{X}}\right\|_F^2 \\ &= \underset{P_\Omega(\mathcal{X}-\mathcal{M})=0}{\operatorname{argmin}} \frac{\gamma_1}{2}\|\mathcal{P}*\mathcal{Q}-\mathcal{X}\|_F^2 + \frac{\gamma_2}{2}\left\|fold_3\left[(\mathcal{U}*\mathcal{V})_{(1)}\right]-\mathcal{X}\right\|_F^2 \\ &= \frac{1}{\gamma_1+\gamma_2}P_{\Omega^c}\left(\gamma_1\mathcal{P}*\mathcal{Q}+\gamma_2 fold_3\left[(\mathcal{U}*\mathcal{V})_{(1)}\right]\right)+P_\Omega(\mathcal{M}).\end{aligned} \tag{21}$$

After updating $\mathcal{X}$, we need to compute the weighting γ_1, γ_2 by

$$\gamma_1 \in \partial\rho\left(\|\mathcal{P}*\mathcal{Q}-\mathcal{X}\|_F^2\right), \quad \gamma_2 \in \partial\rho\left(\left\|\mathcal{U}*\mathcal{V}-\tilde{\mathcal{X}}\right\|_F^2\right). \tag{22}$$

Furthermore, due to $\rho(\cdot)$ is a monotonically increasing function, $\mathcal{P}$ and $\mathcal{Q}$ can be updated by solving the following problem

$$\underset{\mathcal{P},\mathcal{Q}}{\operatorname{argmin}} \ \frac{1}{2}\|\mathcal{P} * \mathcal{Q} - \mathcal{X}\|_F^2 . \tag{23}$$

Clearly, $\mathcal{P}$ and $\mathcal{Q}$ can be updated by (11) and (12) respectively.

Similarly, we can update $\hat{U}$ and $\hat{V}$ as follows:

$$\hat{U}^{(k)} = \begin{cases} \bar{\bar{X}}^{(k)}\left(\hat{V}^{(k)}\right)^{*}\left(\hat{V}^{(k)}\left(\hat{V}^{(k)}\right)^{*}\right)^{\dagger}, k = 1, \ldots, \left\lceil \frac{q+1}{2} \right\rceil, \\ conj\left(\hat{U}^{(q-k+2)}\right), k = \left\lceil \frac{q+1}{2} \right\rceil + 1, \ldots, q, \end{cases} \tag{24}$$

$$\hat{V}^{(k)} = \begin{cases} \left(\left(\hat{U}^{(k)}\right)^{*}\hat{U}^{(k)}\right)^{\dagger}\left(\hat{U}^{(k)}\right)^{*}\bar{\bar{X}}^{(k)}, k = 1, \ldots, \left\lceil \frac{q+1}{2} \right\rceil, \\ conj\left(\hat{V}^{(q-k+2)}\right), k = \left\lceil \frac{q+1}{2} \right\rceil + 1, \ldots, q. \end{cases} \tag{25}$$

Algorithm 2: Double Tubal Rank Tensor Completion (DTRTC)

Input: The tensor data $\mathcal{M} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the observed set Ω, t_0. $\rho(x)$.

Input: $\mathcal{X}^0$, $\hat{P}^0$, $\hat{Q}^0$, $\hat{U}^0$, $\hat{V}^0$. The initialized rank $\boldsymbol{r}^0_{\mathcal{X}} \in \mathbb{R}^{n_3}$ and $\boldsymbol{r}^0_{\tilde{\mathcal{X}}} \in \mathbb{R}^q$.

Parameters γ_1^0, γ_2^0. **While not converge do**

1. Fix $\hat{Q}^t$ and $\mathcal{X}^t$ to update $\hat{P}^{t+1}$ by (11).

2. If $t \leq t_0$ then
Fix $\hat{P}^{t+1}$ and $\hat{Q}^t$ to compute $\mathcal{X}^t$ by (21).

3. Fix $\hat{P}^{t+1}$ and $\mathcal{X}^t$ to update $\hat{Q}^{t+1}$ by (12).

4. If $t \leq t_0$ then
Fix $\hat{P}^{t+1}$ and $\hat{Q}^{t+1}$ to compute $\mathcal{X}^t$ by (21).

5. Fix $\hat{V}^t$ and $\mathcal{X}^t$ to update $\hat{U}^{t+1}$ by (24).

6. If $t \leq t_0$ then
Fix $\hat{U}^{t+1}$ and $\hat{V}^t$ to compute $\mathcal{X}^t$ by (21).

7. Fix $\hat{U}^{t+1}$ and $\mathcal{X}^t$ to update $\hat{V}^{t+1}$ by (25).

8. Adopt the rank decreasing scheme to adjust $\boldsymbol{r}^t_{\mathcal{X}}$ and $\boldsymbol{r}^t_{\tilde{\mathcal{X}}}$, adjust the sizes of $\hat{P}^{t+1}$, $\hat{Q}^{t+1}$, $\hat{U}^{t+1}$ and $\hat{V}^{t+1}$.

9. Fix $\hat{P}^{t+1}$, $\hat{Q}^{t+1}$, $\hat{U}^{t+1}$, $\hat{V}^{t+1}$ to compute $\mathcal{X}^{t+1}$ by (21).

10. Compute γ_1^{t+1}, γ_2^{t+1} by (22).

11. Check the stop criterion: $\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F / \|\mathcal{X}^t\|_F < \varepsilon$.

12. $t \leftarrow t + 1$.

end while

Theorem 2

Assume that the sequence $\{\mathcal{P}^t, \mathcal{Q}^t, \mathcal{U}^t, \mathcal{V}^t, \mathcal{X}^t\}$ generated by Algorithm 2 is bounded, Then it satisfies the following properties:

(1) f^t is monotonically decreasing. Actually, it satisfies the following inequality:

$$f^t - f^{t+1} \geq \frac{\gamma_1^t}{2n_3} \left\| \hat{P}^{t+1}\hat{Q}^{t+1} - \hat{P}^t\hat{Q}^t \right\|_F^2 + \frac{\gamma_2^t}{2q} \left\| \hat{U}^{t+1}\hat{V}^{t+1} - \hat{U}^t\hat{V}^t \right\|_F^2 + \frac{1}{2} \left\| \mathcal{X}^{t+1} - \mathcal{X}^t \right\|_F^2 \geq 0.$$

(2) Any accumulation point $(\mathcal{P}_\star, \mathcal{Q}_\star, \mathcal{U}_\star, \mathcal{V}_\star, \mathcal{X}_\star)$ of the sequence $\{\mathcal{P}^t, \mathcal{Q}^t, \mathcal{U}^t, \mathcal{V}^t, \mathcal{X}^t\}$ is a KKT point of problem (19).

Contents

Grayscale Image Inpainting

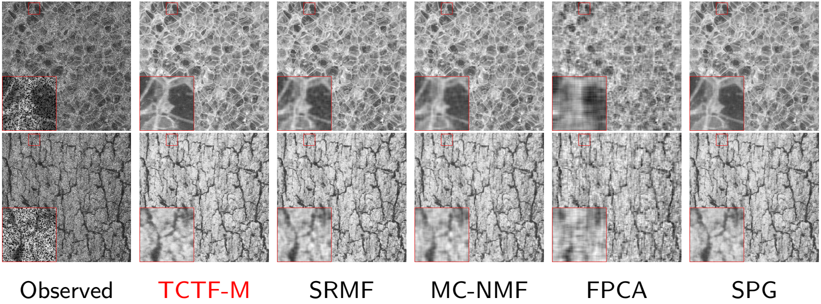

Figure 2: Examples of grayscale image inpainting. From top to bottom, the results are for “Plastic” and “Bark”, respectively

Table 1: Grayscale image inpainting performance comparison

Image	Methods	PSNR	SSIM	FSIM	Time
Plastic	TCTF-M	**30.762**	**0.872**	**0.995**	**0.796**
	SRMF	27.148	0.708	0.973	18.867
	MC-NMF	26.512	0.673	0.964	3.070
	FPCA	20.855	0.397	0.833	41.008
	SPG	29.709	0.841	0.984	15.725
Bark	TCTF-M	**29.590**	**0.890**	**0.996**	**0.765**
	SRMF	25.651	0.727	0.975	18.524
	MC-NMF	24.413	0.663	0.960	3.497
	FPCA	19.219	0.400	0.847	40.227
	SPG	29.306	0.881	0.990	17.890
Wash	TCTF-M	**24.207**	**0.816**	**0.996**	**3.157**
	SRMF	19.383	0.364	0.965	108.759
	MC-NMF	19.013	0.312	0.946	12.736
	FPCA	17.210	0.200	0.825	268.344
	SPG	24.046	0.783	0.990	184.925

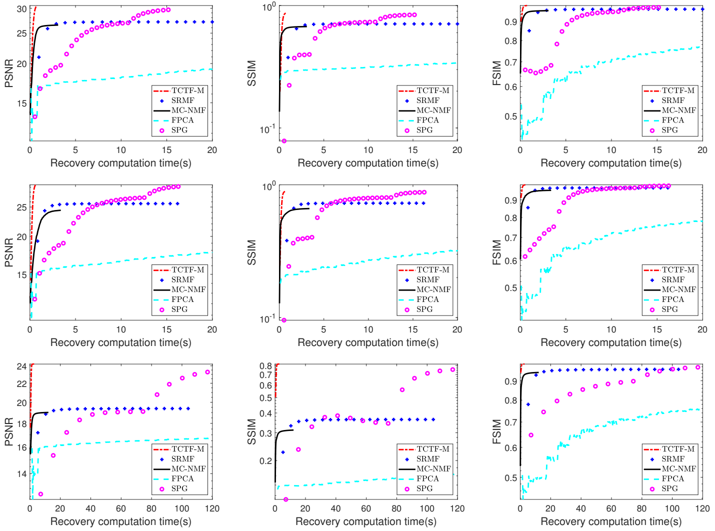

Figure 3: Grayscale image inpainting results. From top to bottom, the results are for “Plastic”, “Bark” and “Wash”, respectively

High Altitude Aerial Image Inpainting

	Methods	$SR = 40\%$				$SR = 50\%$			
		PSNR	SSIM	FSIM	Time	PSNR	SSIM	FSIM	Time
San Francisco	DTRTC	**29.997**	**0.832**	0.978	**5.73**	**31.897**	**0.884**	0.988	**4.76**
	WSTNN	29.938	0.806	**0.982**	244.55	31.807	0.858	**0.991**	181.32
	TCTF	27.159	0.752	0.915	9.99	28.907	0.802	0.969	10.42
	TNN	28.839	0.774	0.972	167.20	30.301	0.830	0.984	149.61
	NCPC	26.177	0.693	0.897	36.97	27.240	0.751	0.928	27.85
	NTD	25.586	0.703	0.878	11.06	26.776	0.754	0.918	10.58
Wash	DTRTC	**22.122**	**0.694**	**0.991**	39.16	**23.339**	**0.770**	**0.995**	**26.27**
	WSTNN	13.698	0.372	0.910	1473.13	16.488	0.485	0.960	1464.97
	TCTF	19.560	0.540	0.883	52.78	20.581	0.623	0.929	52.74
	TNN	21.727	0.644	0.980	1299.84	23.144	0.732	0.990	1303.74
	NCPC	19.310	0.528	0.871	118.27	20.298	0.608	0.919	121.55
	NTD	18.910	0.518	0.811	**33.84**	19.655	0.589	0.87	34.959

Contents

Matrix completion

We established a relationship between matrix rank and tensor tubal rank. Based on the relationship, we modeled the matrix completion problem as a third order tensor completion problem and proposed a two-stage tensor factorization based algorithm, which made a drastic reduction on the dimension of data and hence cut down on the running time.

Tensor completion

We introduced double tubal rank. Compared to tubal rank, 3-tubal rank and tensor fibered rank, double tubal rank can not only fully exploit the low rank structures of the tensor but also avoid the low rank structures redundancy. Based on this rank, a reweighted tensor factorization algorithm was proposed.

Thank you!