Tensor Principal Component Analysis via Convex Optimization

柏园超、王铁楠、乔晗

Outline

- Tensor Principal Component Analysis Model
- Nuclear Norm Penalty and SDP Relaxation Methods with Numerical Results
- ADMM Solution and Experiments

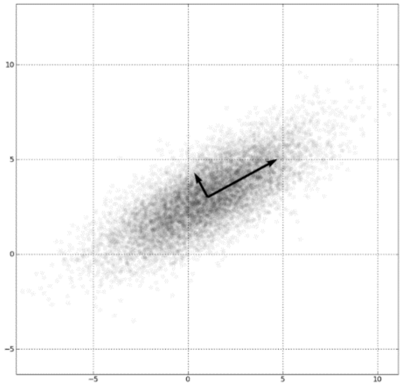
Tensor Principal Component Analysis Model

- Introduction of Principal Component Analysis(PCA)
- Tensor PCA Definition
- Reformulate Tensor PCA as Matrix Optimization

Introduction of PCA

- $X \in \; ^{n \times p}$ is a matrix which consists of <u>*n* sample</u> of <u>*p* variables</u>.
- When *p* is very large and most of *p* variables are similar for *n* samples, we can do <u>PCA</u> to generate the <u>Principal Component</u> and reduce the dimension of samples.
- <u>Principal Component</u> is $y_i = Xl_i$,where $l_i \in \; ^{p}$ is the weights of *p* variables. y_i satisfies that $\mathrm{Var}(y_i)$ is as much as possible and $Cov(y_i, y_j)$ equals 0.

	V 1	V 2
Sample1	……	……
…….	……	……



Introduction of PCA

$$Var(y_i) = Var(Xl_i) = l_i^T \sum l_i$$

$$Cov(y_i, y_j) = Cov(Xl_i, Xl_j) = l_i^T \sum l_j$$

$$\sum = (\sigma_{ij})_{p \times p} = E\left[\left(X - E(X)\right)^T \left(X - E(X)\right)\right]$$

- $\sum$ is covariance matrix of *p* variables.
- Find first PC:

$$\max \quad l_i^T \sum l_i$$
$$s.t. \quad \|l_i\| = 1$$

- l_i^* is eigenvector, $Opt = \lambda_1$.

Introduction of PCA

- When $B = A^T A$ is a symmetric matrix.
- $D = X^T BX$, $A = U \sum V^T$ → $D_{i,i} = \sum_{i,i}^2$, $V = X$

  diagonalize SVD Eigenvalue Singular value

$$\max \quad x^T Bx \qquad s.t. \quad \|x\| = 1$$

→

$$\max \quad x^T Ay \qquad s.t. \quad \|x\| = 1, \|y\| = 1$$

→

$$\min \quad \left\| A - \sigma xy^T \right\| \qquad s.t. \quad \|x\| = 1, \|y\| = 1$$

- PCA problem is equal to find maximal singular value problem!

Tensor Principal Component Analysis Model

- Introduction of Principal Component Analysis(PCA)
- Tensor PCA Definition
- Reformulate Tensor PCA as Matrix Optimization

Tensor PCA Definition

- Definition of Tensor PCA

$$\min \left\| A - \lambda x^1 \otimes x^2 \otimes \cdots \otimes x^m \right\|$$

$$s.t. \quad \lambda \in R, \left\| x^i \right\| = 1, i = 1, 2, \ldots, m$$

- A is a tensor, which is a <u>high dimensional array</u> of real data.
- A = $\left(A_{i_1 i_2 \ldots i_m} \right)_{n_1 \times n_2 \times \ldots \times n_m}$

Tensor PCA Definition

- We often encounter multidimensional data, such as images, video, range data and medical data such as CT and MRI.
- Color Image(3D), Color Video(4D)
- It is more reasonable to treat the multidimensional data as a tensor instead of unfolding it into a matrix.
- Images obtained by tensor PCA technique have higher quality than that by matrix PCA.

H. Wang and N. Ahuja. **Compact representation of multidimensional data using tensor rank-one decomposition.**
In Proceedings of the Pattern Recognition, 17th International Conference on ICPR, 2004.

Tensor PCA Definition

Tensor

$$\min \left\| A - \lambda x^1 \otimes x^2 \otimes \cdots \otimes x^m \right\|$$

$$s.t. \quad \lambda \in R, \left\| x^i \right\| = 1, i = 1, 2, \ldots, m$$

$$A = \sum_{i=1}^{r} \lambda_i \left(x^1 \otimes x^2 \otimes \cdots \otimes x^m \right)$$

CP-rank=r

$$\max\ Ag\left(x^1 \otimes x^2 \otimes \cdots \otimes x^m \right)$$

$$s.t. \quad \left\| x^i \right\| = 1, i = 1, 2, \ldots, m$$

Matrix

$$\min \quad \left\| A - \sigma x y^T \right\|$$

$$s.t. \quad \|x\| = 1, \|y\| = 1$$

$$A = \sum_{i=1}^{r} \sigma_i x_i y_i^T$$

rank=r

$$\max \quad x^T A y = Ag\left(x y^T \right)$$

$$s.t. \quad \|x\| = 1, \|y\| = 1$$

Tensor PCA Definition

Tensor

```
>> perms([1,1,2])

ans =

     2     1     1
     2     1     1
     1     2     1
     1     1     2
     1     1     2
     1     2     1
```

$$F_{i_1 i_2 \dots i_m} = F_{\pi(i_1 i_2 \dots i_m)} \in S^{n^m}$$

Super-symmetric

$$\max \; F \bullet (x \otimes x \otimes \cdots \otimes x)$$
$$s.t. \quad \|x\| = 1$$

$$\max \; F(x, x, \cdots, x)$$
$$s.t. \quad \|x\| = 1$$

Matrix

$$X^T = X \in S^{n^2}$$

$$\max \quad x^T A x = A \bullet (x x^T)$$
$$s.t. \quad \|x\| = 1$$

Tensor PCA Definition

Tensor

$$\max\ F(\overbrace{x, x, \cdots, x}^{2d})$$
$$s.t.\quad \|x\| = 1$$

Find maximal Z-eigenvalue

$$\max\ F \bullet X$$
$$s.t.\quad \sum_{k \in K(n,d)} \frac{d!}{\Pi_{j=1}^{n} k_j!} X_{1^{2k_1} 2^{2k_2} \ldots n^{2k_n}} = 1 = \|x\|^{2d} = Tr(X)$$
$$X \in S^{n^{2d}}, \text{cp}-rank(X) = 1$$
$$K(n,d) = \left\{ k = (k_1, \ldots, k_n) \in Z^n \mid \sum_{j=1}^{n} k_j = d \right\}$$

Matrix

$$\max\quad x^T A x = A \bullet (xx^T)$$
$$s.t.\quad \|x\| = 1$$

Find maximal eigenvalue

$$\max\quad A \bullet X = A \bullet (xx^T)$$
$$s.t.\quad Tr(X) = Tr(xx^T) = \|x\|^2 = 1$$
$$rank(X) = 1, X \in S^{n^2}$$

Tensor Principal Component Analysis Model

- Introduction of Principal Component Analysis(PCA)
- Tensor PCA Definition
- Reformulate Tensor PCA as Matrix Optimization

Reformulate Tensor PCA as Matrix Optimization

- Hard to solve the problem directly in tensor form

$$
\begin{aligned}
&\max\ F g X \\
&s.t.\quad Tr(X) = \sum_{k \in K(n,d)} \frac{d!}{\Pi_{j=1}^{n} k_j!} X_{1^{2k_1} 2^{2k_2} \ldots n^{2k_n}} = 1 \\
&\qquad\ X \in S^{n^{2d}}, CP - rank(X) = 1
\end{aligned}
$$

$$
K(n,d) = \left\{ k = (k_1, \ldots, k_n) \in Z^n \mid \sum_{j=1}^{n} k_j = d \right\}
$$

Reformulate Tensor PCA as Matrix Optimization

- Matrix rearrangement:

$$M(F)_{kl} := F_{i_1 \dots i_d i_{d+1} \dots i_{2d}}, 1 \le i_1, \dots, i_d, i_{d+1}, \dots, i_{2d} \le n$$

$$k = \sum_{j=1}^{d} (i_j - 1) n^{d-j} + 1, \; l = \sum_{j=d+1}^{2d} (i_j - 1) n^{2d-j} + 1$$

- Vectorization:

$$\mathrm{V}(F)_k := F_{i_1 \dots i_m},$$

$$k = \sum_{j=1}^{m} (i_j - 1) n^{m-j} + 1, 1 \le i_1, \dots, i_m \le n$$

Reformulate Tensor PCA as Matrix Optimization

- Suppose $A \in R^{n^d}$. Then the following two statements are equivalent:

  (i) $A \in S^{n^d}$ and $rank(A) = 1$

  (ii) $A \otimes A \in S^{n^{2d}}$

- Suppose $X_t \in S^{n^{2d}}$ and $X = M(X_t) \in R^{n^d \times n^d}$. Then we have

$$CP-rank(X_t) \Leftrightarrow rank(X) = 1$$

Reformulate Tensor PCA as Matrix Optimization

- Reformulate Tensor PCA as Matrix Optimization

$$\max F_t g X_t$$

$$s.t. \quad Tr(X_t) = \sum_{k \in K(n,d)} \frac{d!}{\Pi_{j=1}^{n} k_j !} (X_t)_{1^{2k_1} 2^{2k_2} \ldots n^{2k_n}} = 1$$

$$X_t \in S^{n^{2d}}, CP - rank(X_t) = 1$$

$$K(n,d) = \left\{ k = (k_1, \ldots, k_n) \in Z^n \mid \sum_{j=1}^{n} k_j = d \right\}$$

$$\max Tr(FX)$$

$$s.t. \quad Tr(X) = 1, M^{-1}(X) \in S^{n^{2d}}$$

$$X \in S^{n^d \times n^d}, rank(X) = 1$$

$$F g X = \|FX\|_F^2 = Tr(FX)$$

Outline

- Tensor Principal Component Analysis Model
- Nuclear Norm Penalty and SDP Relaxation Methods with Numerical Results
- ADMM Solution and Experiments

Nuclear Norm Penalty Approach

- Initial Problem:

$$\begin{aligned} max \quad & Tr(FX) \\ s.t. \quad & Tr(X) = 1, M^{-1}(X) \in S^{n^{2d}}, \\ & X \in S^{n^{2d} \times n^{2d}}, rank(X) = 1 \end{aligned}$$

- Problem $\begin{aligned} min \quad & rank(X) \\ s.t. \quad & C(X) = b \end{aligned}$ is equivalent to $\begin{aligned} min \quad & \|X\|_* \\ s.t. \quad & C(X) = b \end{aligned}$ with high probability under certain randomness hypothesis.

Nuclear Norm Penalty Approach

- Add nuclear norm term of X to initial problem to penalize high rank:

$$\begin{aligned} max \quad & Tr(FX) - \rho ||X||_* \\ s.t. \quad & Tr(X) = 1, M^{-1}(X) \in S^{n^{2d}}, \\ & X \in S^{n^{2d} \times n^{2d}} \end{aligned}$$

If the optimal solution of the above problem (denoted by X^*) is rank-one, then $||X^*||_* = Tr(X^*) = 1$

Semidefinite Programming Relaxation

- Replace the rank-one constraint by SDP constraint:

$$
\begin{aligned}
max \quad & Tr(FX) \\
s.t. \quad & Tr(X) = 1, \\
& M^{-1}(X) \in S^{n^{2d}}, X \geq 0
\end{aligned}
$$

Numerical Results

- To solve a super-symmetric tensor PCA problem, three steps are needed:
- 1、Create the super-symmetric tensor.
- 2、Unfold the tensor into a square matrix.
- 3、Solve the corresponding SDP or NNP convex problem.

Results for Two Examples

- Example 1:

$$\mathcal{F}_{1111} = 0.2883, \quad \mathcal{F}_{1112} = -0.0031, \quad \mathcal{F}_{1113} = 0.1973, \quad \mathcal{F}_{1122} = -0.2485, \quad \mathcal{F}_{1123} = -0.2939,$$
$$\mathcal{F}_{1133} = 0.3847, \quad \mathcal{F}_{1222} = 0.2972, \quad \mathcal{F}_{1223} = 0.1862, \quad \mathcal{F}_{1233} = 0.0919, \quad \mathcal{F}_{1333} = -0.3619,$$
$$\mathcal{F}_{2222} = 0.1241, \quad \mathcal{F}_{2223} = -0.3420, \quad \mathcal{F}_{2233} = 0.2127, \quad \mathcal{F}_{2333} = 0.2727, \quad \mathcal{F}_{3333} = -0.3054.$$

- Example2

$$\begin{aligned} & 0.74694x_1^4 - 0.435103x_1^3x_2 + 0.454945x_1^2x_2^2 + 0.0657818x_1x_2^3 + x_2^4 \\ + & 0.37089x_1^3x_3 - 0.29883x_1^2x_2x_3 - 0.795157x_1x_2^2x_3 + 0.139751x_2^3x_3 + 1.24733x_1^2x_3^2 \\ + & 0.714359x_1x_2x_3^2 + 0.316264x_2^2x_3^2 - 0.397391x_1x_3^3 - 0.405544x_2x_3^3 + 0.794869x_3^4. \end{aligned}$$

- Solution difference: $\|X_{publication} - X\|_F / max(\|X_{publication}\|_F, 1)$

	Example1	Example2
Sol.Dif	2.5056E-04	1.3822E-04

Frequency of rank-one solution

- The tensors are randomly generated with order 2d and dimension n, the entries are uniformly distributed between [-2, 2]
- Computed by CVX 2.1 (SeDuMi 1.34)

SDP			NNP (rho = 10)			NNP (rho = 100)		
n	rank1	CPU	n	rank1	CPU	n	rank1	CPU
3	100	0.2376	3	100	0.3758	3	100	0.3555
4	100	0.4907	4	80	1.0314	4	100	0.897
5	100	1.3044	5	7	2.711	5	100	2.3344
6	100	3.0026	6	0	3.1877	6	100	3.0793
7	100	5.2464	7	0	8.617	7	100	8.1737
8	100	6.2614	8	0	26.469	8	100	25.866
9	100	10.3896	9	0	80.822	9	100	81.6773

Comparison of NNP and SDP

	Inst	Sol Dif	Time NNP	Time SDP
n = 4	1	3.94E-05	1.13	0.53
	2	3.12E-05	0.93	0.51
	3	2.76E-05	0.94	0.57
	4	2.43E-05	0.95	0.52
	5	4.23E-05	1.05	0.65
	6	6.19E-05	1.01	0.54
	7	2.37E-05	0.94	0.52
	8	5.43E-05	0.94	0.54
	9	6.74E-06	0.99	0.53
	10	3.76E-05	0.93	0.53
n = 5	1	2.33E-05	2.21	1.13
	2	5.84E-05	2.36	1.13
	3	1.24E-05	2.52	1.11
	4	2.03E-05	2.36	1.12
	5	1.57E-06	2.48	1.12
	6	4.81E-06	2.49	1.11
	7	1.32E-05	2.49	1.13
	8	4.94E-05	2.21	1.12
	9	4.24E-06	2.51	1.32
	10	5.21E-06	2.74	1.12
n = 6	1	3.92E-05	3.49	2.19
	2	1.20E-05	3.03	2.45
	3	2.54E-05	3.3	2.31
	4	2.36E-05	3.24	2.39
	5	7.62E-05	3.45	2.27
	6	2.06E-05	3.47	2.49
	7	2.41E-05	3.76	2.19
	8	1.97E-05	2.97	2.24
	9	5.28E-05	3.59	2.18
	10	5.85E-05	3.62	2.17

Comparison of CVX and ADMM

	Inst	Sol Dif	ADMM.Iter
n = 2	1	1.36E-05	243
	2	3.51E-05	262
	3	0.000113833	244
	4	3.79E-05	243
	5	2.68E-05	239
	6	9.27E-05	249
	7	7.40E-05	236
	8	2.26E-05	233
	9	1.56E-06	238
	10	7.36E-06	241
n = 3	1	3.03E-05	500
	2	0.000146228	470
	3	9.63E-05	474
	4	0.000151246	471
	5	8.77E-06	479
	6	5.14E-05	468
	7	9.73E-05	475
	8	5.66E-05	468
	9	0.000141543	518
	10	0.000187851	473
n = 4	1	5.04E-05	734
	2	1.92E-05	716
	3	2.61E-05	733
	4	3.03E-05	711
	5	9.31E-06	744
	6	5.90E-05	718
	7	5.70E-05	717
	8	7.96E-06	740
	9	6.81E-06	734
	10	1.87E-05	711

Outline

- Tensor Principal Component Analysis Model
- Nuclear Norm Penalty and SDP Relaxation Methods with Numerical Results
- ADMM Solution and Experiments

ADMM标准形式

$$\begin{array}{ll} \min_{x\in\mathbf{R}^n, y\in\mathbf{R}^p} & f(x)+g(y) \\ \text{s.t.} & Ax+By=b \\ & x\in\mathcal{C},\ \ y\in\mathcal{D}, \end{array} \tag{21}$$

$$\left\{\begin{array}{lll} x^{k+1} & := & \operatorname{argmin}_{x\in\mathcal{C}}\ \mathcal{L}_\mu(x,y^k;\lambda^k) \\ y^{k+1} & := & \operatorname{argmin}_{y\in\mathcal{D}}\ \mathcal{L}_\mu(x^{k+1},y;\lambda^k) \\ \lambda^{k+1} & := & \lambda^k-(Ax^{k+1}+By^{k+1}-b)/\mu, \end{array}\right. \tag{22}$$

$$\mathcal{L}_\mu(x,y;\lambda) := f(x)+g(y)-\langle\lambda, Ax+By-b\rangle+\frac{1}{2\mu}\|Ax+By-b\|^2$$

子问题–投影

$$\begin{cases} X^{k+1} & := & \text{argmin}_{X\in\mathcal{C}} -\text{Tr}\,(FY^k) + \rho\|Y^k\|_* - \langle \Lambda^k, X - Y^k\rangle + \frac{1}{2\mu}\|X - Y^k\|_F^2 \\ Y^{k+1} & := & \text{argmin} \;\; -\text{Tr}\,(FY) + \rho\|Y\|_* - \langle \Lambda^k, X^{k+1} - Y\rangle + \frac{1}{2\mu}\|X^{k+1} - Y\|_F^2 \\ \Lambda^{k+1} & := & \Lambda^k - (X^{k+1} - Y^{k+1})/\mu, \end{cases} \tag{24}$$

$$X^{k+1} := \underset{X\in\mathcal{C}}{\text{argmin}}\, \frac{1}{2}\|X - (Y^k + \mu\Lambda^k)\|_F^2.$$

$$\begin{array}{ll} \min & \|X - Z\|_F^2 \\ \text{s.t.} & \text{Tr}\,(X) = 1, \\ & \boldsymbol{M}^{-1}(X) \in \mathbf{S}^{n^{2d}}. \end{array} \tag{27}$$

ADMM for Nuclear Norm Penalty Problem

原问题

$$\begin{aligned} \max \quad & \operatorname{Tr}(FX) - \rho\|X\|_* \\ \text{s.t.} \quad & \operatorname{Tr}(X) = 1,\ \boldsymbol{M}^{-1}(X) \in \mathbf{S}^{n^{2d}}, \\ & X \in \mathbf{S}^{n^d \times n^d}, \end{aligned} \tag{15}$$

变形

$$\begin{aligned} \min \quad & -\operatorname{Tr}(FY) + \rho\|Y\|_* \\ \text{s.t.} \quad & X - Y = 0, \\ & X \in \mathcal{C}, \end{aligned} \tag{23}$$

$$\mathcal{C} := \{X \in \mathbf{S}^{n^d \times n^d} \mid \operatorname{Tr}(X) = 1,\ \boldsymbol{M}^{-1}(X) \in \mathbf{S}^{n^{2d}}\}.$$

子问题-核范数

$$\begin{cases} X^{k+1} & := & \operatorname{argmin}_{X\in\mathcal{C}} -\operatorname{Tr}(FY^k) + \rho\|Y^k\|_* - \langle\Lambda^k, X - Y^k\rangle + \frac{1}{2\mu}\|X - Y^k\|_F^2 \\ Y^{k+1} & := & \operatorname{argmin} \ -\operatorname{Tr}(FY) + \rho\|Y\|_* - \langle\Lambda^k, X^{k+1} - Y\rangle + \frac{1}{2\mu}\|X^{k+1} - Y\|_F^2 \\ \Lambda^{k+1} & := & \Lambda^k - (X^{k+1} - Y^{k+1})/\mu, \end{cases} \quad (24)$$

matrix shrinkage operation

$$Y^{k+1} := \underset{Y}{\operatorname{argmin}} \quad \mu\rho\|Y\|_* + \frac{1}{2}\|Y - (X^{k+1} - \mu(\Lambda^k - F))\|_F^2. \quad (26)$$

$$Y^{k+1} := U\operatorname{Diag}(\max\{\sigma - \mu\rho, 0\})V^\top,$$

where $U\operatorname{Diag}(\sigma)V^\top$ is the singular value decomposition of matrix $X^{k+1} - \mu(\Lambda^k - F)$.

子问题-投影

$$\begin{aligned} \min \quad & \|X - Z\|_F^2 \\ \text{s.t.} \quad & \mathrm{Tr}\,(X) = 1, \\ & \boldsymbol{M}^{-1}(X) \in \mathbf{S}^{n^{2d}}. \end{aligned}$$

$$\begin{aligned} \min \quad & \|\mathcal{X} - \mathcal{Z}\|_F^2 \\ \text{s.t.} \quad & \sum_{k \in \mathbb{K}(n,d)} \frac{d!}{\prod_{j=1}^{n} k_j!} \mathcal{X}_{1^{2k_1} 2^{2k_2} \cdots n^{2k_n}} = 1, \\ & \mathcal{X} \in \mathbf{S}^{n^{2d}}, \end{aligned}$$

$$\mathbf{I} = \left\{ (i_1 \cdots i_{2d}) \in \pi(1^{2k_1} \cdots n^{2k_n}) \mid k = (k_1, \cdots, k_n) \in \mathbb{K}(n,d) \right\}.$$

$$\begin{cases} 2\left(|\pi(i_1 \cdots i_{2d})|\, \mathcal{X}_{i_1 \cdots i_{2d}} - \sum\limits_{j_1 \cdots j_{2d} \in \pi(i_1 \cdots i_{2d})} \mathcal{Z}_{j_1 \cdots j_{2d}} \right) = 0, & \text{if } (i_1 \cdots i_{2d}) \notin \mathbf{I}, \\ 2\left(\frac{(2d)!}{\prod_{j=1}^{n} (2k_j)!} \mathcal{X}_{1^{2k_1} \cdots n^{2k_n}} - \sum\limits_{j_1 \cdots j_{2d} \in \pi(1^{2k_1} \cdots n^{2k_n})} \mathcal{Z}_{j_1 \cdots j_{2d}} \right) - \lambda \frac{(d)!}{\prod_{j=1}^{n} (k_j)!} = 0, & \text{otherwise.} \end{cases}$$

$$\begin{cases} \mathcal{X}^*_{i_1 \cdots i_{2d}} = \hat{\mathcal{Z}}_{i_1 \cdots i_{2d}}, & \text{if } (i_1 \cdots i_{2d}) \notin \mathbf{I}, \\ \mathcal{X}^*_{1^{2k_1} \cdots n^{2k_n}} = \frac{\lambda}{2}\, \alpha(k,d) + \hat{\mathcal{Z}}_{1^{2k_1} \cdots n^{2k_n}}, & \text{otherwise}. \end{cases} \tag{29}$$

$$\lambda = 2\left(1 - \sum_{k \in \mathbb{K}(n,d)} \frac{(d)!}{\prod_{j=1}^{n} (k_j)!} \hat{\mathcal{Z}}_{1^{2k_1} \cdots n^{2k_n}} \right) \Bigg/ \sum_{k \in \mathbb{K}(n,d)} \frac{(d)!}{\prod_{j=1}^{n} (k_j)!} \alpha(k,d),$$

ADMM for SDP Relaxation

$$
\begin{array}{llll}
(SDR) & \max & \mathrm{Tr}\,(FX) & \\
& \text{s.t.} & \mathrm{Tr}\,(X) = 1, & (16) \\
& & \boldsymbol{M}^{-1}(X) \in \mathbf{S}^{n^{2d}},\ X \succeq 0. &
\end{array}
$$

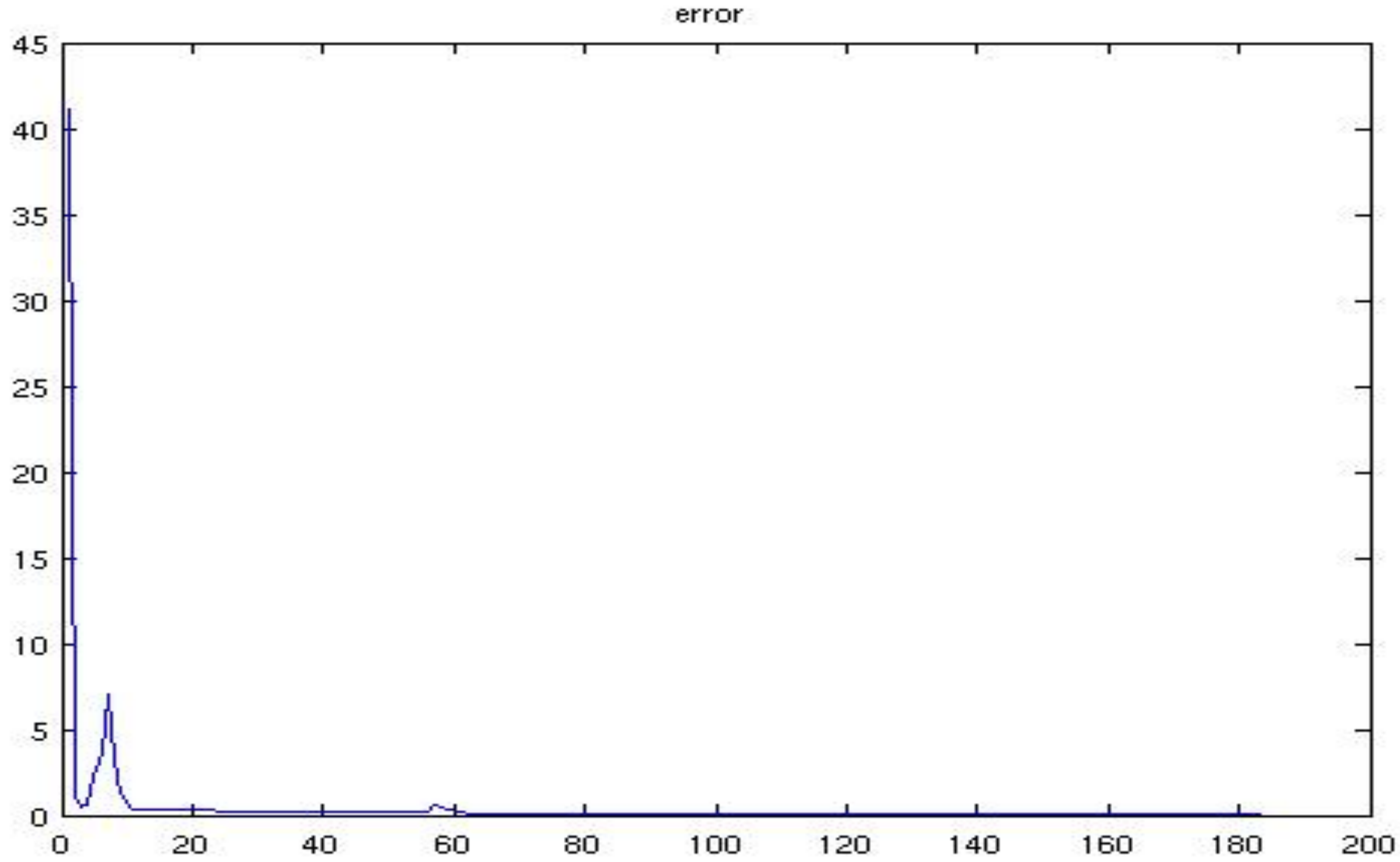
$$
\begin{array}{llll}
\min & -\mathrm{Tr}\,(FY) & & \\
\text{s.t.} & \mathrm{Tr}\,(X) = 1, & \boldsymbol{M}^{-1}(X) \in \mathbf{S}^{n^{2d}} & (30) \\
& X - Y = 0, & Y \succeq 0. &
\end{array}
$$

$$
\left\{
\begin{array}{lll}
X^{k+1} & := & \mathrm{argmin}_{X \in \mathcal{C}} -\mathrm{Tr}\,(FY^k) - \langle \Lambda^k, X - Y^k \rangle + \frac{1}{2\mu}\|X - Y^k\|_F^2 \\
Y^{k+1} & := & \mathrm{argmin}_{Y \succeq 0} -\mathrm{Tr}\,(FY) - \langle \Lambda^k, X^{k+1} - Y \rangle + \frac{1}{2\mu}\|X^{k+1} - Y\|_F^2 \\
\Lambda^{k+1} & := & \Lambda^k - (X^{k+1} - Y^{k+1})/\mu,
\end{array}
\right. \quad (31)
$$

$$Y^{k+1} := U\mathrm{Diag}\,(\max\{\sigma, 0\})U^\top$$

参考课本

收敛速度



$$\frac{\|X^k - X^{k-1}\|_F}{\|X^{k-1}\|_F} + \|X^k - Y^k\|_F \leq 10^{-6}$$