MULTIPLE k-MEANS CLUSTERING BASED LOCALLY LOW-RANK APPROACH TO NONLINEAR MATRIX COMPLETION

Katsumi Konishi

Hosei university

Tokyo, Japan

konishi@hosei.ac.jp

Tomoki Shise

Tokyo university of science

Tokyo, Japan

4418514@ed.tus.ac.jp

Ryohei Sasaki

Tokyo university of science

Tokyo, Japan

clustereddata@gmail.com

Toshihiro Furukawa

Tokyo university of science

Tokyo, Japan

tofurukawa@rs.tus.ac.jp

Abstract—This paper deals with nonlinear matrix completion problem, which is a problem of restoring missing entries in a given matrix, where its column vectors belong to a low dimensional manifold. Assuming that a low dimensional manifold can be approximated locally as a low dimensional linear subspace, this paper proposes a new locally low-rank approach. In this approach iteratively solves low-rank matrix completion problems for submatrices generated by using the k-means clustering for several values of k and restores missing entries. Numerical examples show that the proposed algorithm achieves better performance than other algorithms.

Index Terms—matrix completion, matrix rank minimization, nuclear norm minimization, compressed sensing

I. INTRODUCTION

This paper deals with matrix completion problem, which is a problem of restoring missing entries in a given matrix. Matrix completion problem has various applications in the field of signal processing such as a image impaitng, signal restoration, collaborative filtering, parameter estimation of linear system [1] and restoration of input and output signals [2]. Several methods of estimating missing entries have been studied, and most of them solve matrix completion problems by assuming that column or row vectors of a matrix belong to a low dimensional linear space and formulating them as matrix rank minimization problems [3]–[5]. To achieve high recovery performance, some algorithms assume that column vectors belong to multiple low dimensional linear subspaces. Based on this assumption, an algorithm using subspace clustering has been proposed in [6], and an algebraic variety approach has been proposed in [7]. However, observed signals in real applications do not always belong to linear subspaces, and restoration accuracy sometimes becomes worse for these applications.

To achieve higher recovery performance for real applications, this paper deals with a nonlinear matrix completion problem where column vectors of a matrix are assumed to belong to a low dimensional manifold. Based on the assumption that a low dimensional manifold can be approximated locally as a low dimensional linear subspace, the authors have already proposed a locally low-rank approach to this problem in [8]. This approach iteratively solves low-rank

matrix completion problems for submatrices consisting of neighbor column vectors of the matrix and achieves higher recovery performance. However, it takes a lot of computing time to restore missing entries because submatrices are generated for all column vectors using their neighbors and because each matrix completion problem is solved repeatedly until converge. To provide a faster algorithm based on the locally low-rank approach, this paper introduces the k-means clustering and proposes a new algorithm which solves low-rank matrix completion problems for submatrices generated by using the k-means clustering for several values of k. Numerical examples show that the proposed algorithm solves nonlinear matrix completion problems better than a standard low-rank approach and takes less computing time than the locally low-rank approach.

II. MATRIX COMPLETION AND LOW-RANK APPROACH

This section gives a standard linear matrix completion problem and a low-rank approach proposed in [5].

The matrix completion problem is a problem of estimating that missing entries of a matrix $X \in \mathbb{R}^{M \times N}$. We usually assume that its column or row vectors belong to a low dimensional linear subspace, that is, X is a low-rank matrix, and the problem is formulated as the following matrix rank minimization problem,

$$\begin{aligned} \text{Minimize} \quad & \textbf{rank}(X) \\ \text{subject to} \quad & P_\Omega(X) = P_\Omega(X_{true}) \end{aligned}, \quad (1)$$

where Ω denotes given index set, $P_\Omega : \mathbb{R}^{M \times N} \to \mathbb{R}^{M \times N}$ denotes a linear operator that projects all entries except subscripts included in the set Ω to 0, and X_{true} denotes a true matrix to be recovered. Although this problem is generally NP-hard, several algorithms have been proposed to obtain an approximate solution. This paper introduces a truncated nuclear norm minimization approach, which relaxes the objective function of (1) by the truncated nuclear norm of X as follows,

$$\begin{aligned} \text{Minimize} \quad & \|X\|_{*,r} \\ \text{subject to} \quad & P_\Omega(X) = P_\Omega(X_{true}) \end{aligned}, \quad (2)$$

where $\|X\|_{*,r}$ denotes the truncated nuclear norm defined by

$$\|X\|_{*,r} = \sum_{i=r+1}^{M} \sigma_i, \qquad (3)$$

with respect to the ith greatest singular value σ_i of X. This problem can be solved by iterative partial matrix shrinkage algorithm (IPMS) [5], which iterates the following update schemes until converge,

Step 1 $Z \leftarrow \mathcal{T}_{r,\lambda}(X)$
Step 2 $X \leftarrow P_{\Omega^c}(Z) + P_{\Omega}(X_{true})$

where $\mathcal{T}_{r,\lambda}(X)$ denotes a partial soft thresholding operator, which shrinks ith greatest singular values by λ, that is, replaces σ_i by $\max(0, \sigma_i - \lambda)$ for $i = r+1, r+2, \ldots, M$. If $r = 0$, (2) is equal to a standard nuclear norm relaxation problem, and the above scheme corresponds with a soft shrinkage algorithm. If the rank of X_{true} is unknown, this update scheme achieves the best performance by letting $r = \text{rank}(X_{true})$. However it is usually unknown, the IPMS algorithm updates r by estimating the rank as follows,

$$r \leftarrow \underset{i}{\arg\max}\, \sigma_i \quad \text{subject to } \sigma_i \geq \alpha\sigma_1,$$

where $\alpha \in (0,1)$ is a given constant. The details are written in [5]. Though this paper uses the IPMS algorithm for locally low-rank approaches due to its recovery performance, any low-rank approaches can be applied.

III. Nonlinear matrix completion and Locally Low-Rank Approach

This section introduces nonlinear matrix completion problems and a locally low-rank approach, which is originally proposed in [8] and modified to simplify notations.

In this paper, we focus on the following nonlinear matrix completion problem,

$$\begin{aligned}
\text{Minimize} \quad &\text{rank}\left([\phi(\boldsymbol{x}_1)\ \phi(\boldsymbol{x}_2)\ \ldots\ \phi(\boldsymbol{x}_N)]\right)\\
\text{subject to} \quad &P_{\Omega}(X) = P_{\Omega}(X_{true})
\end{aligned} \qquad (4)$$

where $\boldsymbol{x}_i$ denotes the ith column vector of matrix X, and $\phi : \mathbb{R}^n \to \mathbb{R}^p$ denotes a nonlinear mapping. If ϕ is known and has its inverse function after converting $\boldsymbol{x}_i$ to $\phi(\boldsymbol{x}_i)$, it can be solved as a matrix rank minimization problem. However, in real applications, ϕ is not always known, and this paper deals with the unknown case.

We assume here that $\phi(\boldsymbol{x}_1), \ldots, \phi(\boldsymbol{x}_N)$ belong to a low dimensional manifold which can be approximate locally as a low dimensional linear subspace, and consider a local matrix completion problem as follows,

$$\begin{aligned}
\text{Minimize} \quad &\sum_{i=1}^{N} \|XD^{(i)}\|_{*,r}\\
\text{subject to} \quad &P_{\Omega}(X) = P_{\Omega}(X_{true})
\end{aligned}, \qquad (5)$$

where $D^{(i)} \in \mathbb{R}^{N \times N}$ denotes diagonal matrix whose jth

---

Algorithm 1 Locally low-rank algorithm.

Input: $X^{(0)}, m, \delta, \alpha_{min}, \eta_\alpha, \epsilon, t_{max}$
1: $X \leftarrow X^{(0)}$; $t \leftarrow 0$; $\alpha \leftarrow 1$
2: **repeat**
3:    $t \leftarrow t + 1$
4:    $X^{old} \leftarrow X$
5:    $\alpha \leftarrow \max(\alpha/\eta_\alpha, \alpha_{min})$
6:    Set D to be $N \times N$ zero matrix
7:    **for** $i = 1$ to N **do**
8:       **for** $j = 1$ to N **do**
9:          **if** x_j is m nearest column vector of x_i **then**
10:            $D_{jj}^{(i)} \leftarrow 1$
11:          **else**
12:            $D_{jj}^{(i)} \leftarrow 0$
13:          **end if**
14:       **end for**
15:       $[U, \sigma_1, \sigma_2, \cdots, \sigma_M, V] \leftarrow \text{SVD}(XD^{(i)})$
16:       $\hat{r} \leftarrow \underset{r}{\arg\min}\, \sigma_r$ s.t. $\sigma_r \geq \alpha\sigma_1$
17:       $\lambda \leftarrow \delta\sigma_{\hat{r}}$
18:       $Z^{(i)} \leftarrow \mathcal{T}_{\hat{r},\lambda}(XD^{(i)})$
19:       $Y \leftarrow D^{(i)}Z^{(i)} + X(I - D^{(i)})$
20:       $X \leftarrow P_{\Omega^c}(Y) + P_{\Omega}(X^{(0)})$
21:    **end for**
22: **until** $\|X - X^{old}\|_F / \|X\|_F < \epsilon$ or $t_{max} < t$
Output: X

---

diagonal element $D_{jj}^{(i)}$ is defined by

$$D_{jj}^{(i)} = \begin{cases} 1 & \text{if } j = i\\ 1 & \text{if } \boldsymbol{x}_j \text{ is the } m\text{-nearest neighbor of } \boldsymbol{x}_i\\ 0 & \text{otherwise} \end{cases}.$$

The nearest neighbor is provided w.r.t. Euclidean distance, and m is a given constant. In order to provide an update scheme, (5) is relaxed as follows,

$$\begin{aligned}
\text{Minimize} \quad &\sum_{i=1}^{N} \tfrac{1}{2}\|XD^{(i)} - Z^{(i)}\|_F^2 + \lambda\|Z^{(i)}\|_{*,r}\\
\text{subject to} \quad &P_{\Omega}(X) = P_{\Omega}(X_{true})
\end{aligned}. \qquad (6)$$

Then the following update scheme is provided to obtain a solution of the above problem, where I denote the identity matrix,

Step 1.$Z^{(i)} \leftarrow \mathcal{T}_{r,\lambda}\left(XD^{(i)}\right)$.
Step 2.$Y \leftarrow D^{(i)}Z^{(i)} + X(I - D^{(i)})$.
Step 3.$X \leftarrow P_{\Omega^c}(Y) + P_{\Omega}(X_{true})$.

Based on the above update scheme and the IPMS algorithm, the locally low-rank algorithm is obtained as shown in Algorithm 1, where X is partially shrunk according to submatrix complete problems (6) after constructing $D^{(i)}$. This algorithm requires N singular value decompositions (SVDs) in one iteration and requires significant computing time.

IV. Multiple k-means clustering based Algorithm

Because the locally low-rank algorithm solves submatrix completion problems (6) for neighbor column vectors of all
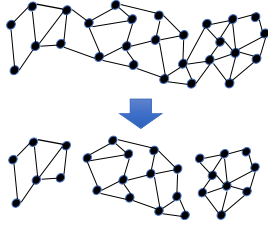
Fig. 1. k submatrix completion problems of k clusters. Black dots and edges denote x_j and their neighbor connectivity.

Fig. 2. A multiple k-means clustering based locally low-rank algorithm.

x_i, it takes high computing cost. To reduce the number of submatrix completion problems, this paper proposes a new algorithm as illustrated in Fig.1, which applies the k-means clustering method to x_i w.r.t. Euclidean distance and gives k submatrix completion problems of k clusters. We define diagonal matrix $D^{(i)}$ whose jth diagonal entries $D_{jj}^{(i)}$ is defined by

$$D_{jj}^{(i)} = \begin{cases} 1 & \text{if } x_j \text{ is a member of the } i\text{th clustering} \\ 0 & \text{otherwise} \end{cases}$$

and then consider a relaxed submatrix complete problem as follows,

$$\begin{aligned} \text{Minimize} \quad & \sum_{i=1}^{k} \tfrac{1}{2}\|XD^{(i)} - Z^{(i)}\|_F^2 + \lambda\|Z^{(i)}\|_{*,r} \\ \text{subject to} \quad & P_\Omega(X) = P_\Omega(X_{true}) \end{aligned} \quad (7)$$

However, this algorithm corresponds with a low-rank approach based on the assumption that column vectors belong to multiple low dimensional linear subspaces, that is, a manifold is approximated by piece-wise linear subspaces, and the recovery accuracy decreases. To achieve high recovery accuracy, this paper proposes multiple k-means clustering based algorithm as illustrated in Fig.2, which uses k-means clustering for $k \in \{k_1, k_2, \ldots, k_K\}$, where k_i denotes the number of members in the ith cluster. Let us define diagonal matrix $D^{(l,i)}$ whose jth diagonal entries $D_{jj}^{(l,i)}$ is defined by

$$D_{jj}^{(l,i)} = \begin{cases} 1 & \text{if } x_j \text{ is a member of the } i\text{th cluster} \\ & \text{of the } l\text{th clustring } (k = k_l) \\ 0 & \text{otherwise} \end{cases}$$

and then consider a relaxed submatrix complete problem as follows,

$$\begin{aligned} \text{Minimize} \quad & \sum_{l=1}^{K}\sum_{i=1}^{k_l} \tfrac{1}{2}\|XD^{(l,i)} - Z^{(l,i)}\|_F^2 + \lambda\|Z^{(l,i)}\|_{*,r} \\ \text{subject to} \quad & P_\Omega(X) = P_\Omega(X_{true}) \end{aligned} \quad (8)$$

A solution of this problem is obtained by the same update scheme for (6) replacing $D^{(i)}$ with $D^{(l,i)}$. Finally this paper proposes a multiple k-means clustering based locally low-rank algorithm as shown in Algorithm 2, where X is partially shrunk according to submatrix complete problems (8) after applying k_l-means clustering for $l \in \{1, 2, \ldots, K\}$ and constructing $\{D^{(l,i)}\}_{l=1}^{K}$. This algorithm requires $\sum_{i=1}^{K} k_i$
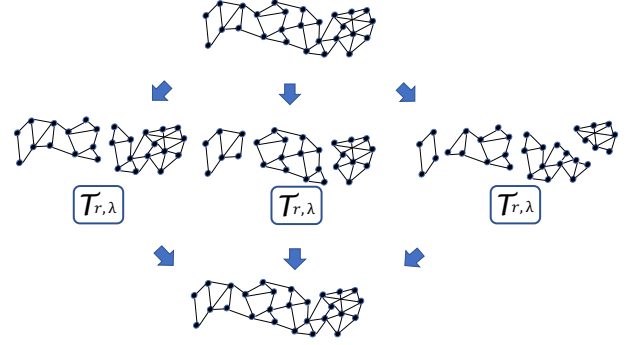

Algorithm 2 Multiple k-means clustering based locally low-rank algorithm.

Input: $X^{(0)}, \{k_l\}_{l=1}^{K}, \delta, \alpha_{min}, \eta_\alpha, \epsilon, t_{max}$
1: $X \leftarrow X^{(0)}$; $\leftarrow 0$; $\alpha \leftarrow 1$
2: **repeat**
3:     $t \leftarrow t+1$
4:     $X^{old} \leftarrow X$
5:     $\alpha \leftarrow \max(\alpha/\eta_\alpha, \alpha_{min})$
6:     **for** $l = 1$ to K **do**
7:         Apply the k_l-means clustering to $\{x_i\}_{i=1}^{N}$, obtain k_l clusters, and construct $D^{(l,i)}$.
8:         **for** $i = 1$ to k_l **do**
9:             $[U, \sigma_1, \sigma_2, \cdots, \sigma_M, V] \leftarrow \text{SVD}(XD^{(l,i)})$
10:           $\hat{r} \leftarrow \arg\min_r \sigma_r$ s.t. $\sigma_r \geq \alpha\sigma_1$
11:           $\lambda \leftarrow \delta\sigma_{\hat{r}}$
12:           $Z^{(l,i)} \leftarrow \mathcal{T}_{\hat{r},\lambda}(XD^{(l,i)})$
13:           $Y \leftarrow Z^{(l,i)}D^{(l,i)} + X(I - D^{(l,i)})$
14:           $X \leftarrow P_{\Omega^c}(Y) + P_\Omega(X^{(0)})$
15:         **end for**
16:     **end for**
17: **until** $\|X - X^{old}\|_F/\|X\|_F < \epsilon$ or $t_{max} < t$
Output: X

---

singular value decompositions (SVDs) in one iteration.

V. NUMERICAL EXAMPLES

This section presents numerical examples to show the efficiency of the proposed multiple k-means clustering based locally low-rank algorithm (Algorithm 2) comparing with the IPMS algorithm, the locally low-rank algorithm (Algorithm 1) and the variety-based matrix completion algorithm (VMC) [7]. In all examples, X restored by the IPMS algorithm was utilized as the initial value $X^{(0)}$ in Algorithm 1 and Algorithm 2, and we use the parameters in Algorithm 1 and Algorithm 2 as $\delta = 0.01$, $\alpha_{min} = 1.0 \times 10^{-4}$, $\eta_\alpha = 1.0093$, $\epsilon = 1.0 \times 10^{-2}$ and $t_{max} = 1000$, which achieve the best performance.

First, the IPMS algorithm, the VMC algorithm, Algorithm 1 and Algorithm 2 were compared using the nonlinear matrix

TABLE I
COMPARISON OF SNR [DB].

$f(x)$	(9)	(10)	(11)
IPMS	25.540	20.523	24.035
VMC($d = 2$)	42.267	41.354	38.634
VMC($d = 3$)	42.793	38.564	43.132
Algorithm 1	55.714	46.938	47.365
Algorithm 2	46.532	42.385	42.622

TABLE II
COMPARISON OF COMPUTING TIME [SEC].

$f(x)$	(9)	(10)	(11)
VMC($d = 2$)	11594.7	10578.7	11482.5
VMC($d = 3$)	11597.2	10741.4	11596.3
Algorithm 1	2177.1	2018.3	2162.2
Algorithm 2	94.3	89.9	92.1

completion problem (4) with ϕ defined by

$$\phi(\boldsymbol{x}) = [f(x_1) \ f(x_2) \ \ldots \ f(x_n)]^T.$$

We examined three kinds of f as follows,

$$
\begin{aligned}
f(x) &= 0.467x + 0.0472x^2 + 0.4727x^3 + 0.2148x^4 \\
&\quad + 0.1977x^5 - 0.2839x^6 + 0.4763x^7, &(9)\\
f(x) &= \sin 2\pi x, &(10)\\
f(x) &= \mathrm{sgn}(x)\frac{|x|}{0.1 + |x|}, &(11)
\end{aligned}
$$

where

$$
\mathrm{sgn}(x) = \left\{ \begin{array}{ll} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{array} \right. .
$$

We generated 50×5000 rank-2 matrix X_{lin}, obtain X by converting each column vector of X_{lin} using f and applied three algorithms under the condition that randomly selected 30% of entries in X are known. In Algorithm 2, we use $K = 3$ and $(k_1, k_2, k_3) = (20, 30, 40)$. Table I and II show the results of the average of 10 trials. The relative error is evaluated by SNR[dB] calculated as

$$\mathrm{SNR} = -20 \log_{10} \|X_{restore} - X_{true}\|_F / \|X_{true}\|_F,$$

where $X_{restore}$ is a matrix restored by the algorithms, and X_{true} is a matrix with complete entries. As can be seen, the IPMS algorithm has much worse performance than Algorithm 1 and Algorithm 2 since it restores a matrix assuming that column vectors belong to a linear subspace. We can see that the proposed algorithm is about 20 times faster and has almost the same recovery performance comparing with Algorithm 1.

Next, Algorithm 2 was applied to a signal restoration problem with a subspace clustering in oder to show its practical performance in the field of subspace clustering. This paper converted subspace signal restoration problems with a subspace clustering into matrix completion problems by generating $N \times M$ data matrices whose columns vector belong to s linear subspaces of dimension 2. We examined algorithms for $(N, M, s) = (50, 100, 5)$ and $(100, 200, 10)$ and used $K = 3$ and $(k_1, k_2, k_3) = (10, 20, 30)$. Table III show the results of the average of 5 trials. As can be seen, the VMC algorithm has much worse performance than Algorithm 1 and Algorithm 2.

VI. CONCLUSION

This paper deals with nonlinear matrix completion problem, which is a problem of restoring missing entries in a given matrix, where its column vectors belong to a low dimensional manifold. Although the locally low-rank algorithm has a good recovery performance for this problem, it requires a lot of computing time. To reduce computational cost, this paper proposes a new locally low-rank approach, which iteratively solves low-rank submatrix completion problems generated by using the k-means clustering for several values of k. Numerical examples show that the proposed algorithm is about 20 times faster and has almost the same recovery performance comparing with the original locally low-rank algorithm. An example of application to image inpaiting also shows its practical performance in the field of image processing. Because the convergence of the proposed algorithm is not guaranteed, further analysis of the algorithm would be required.

ACKNOWLEDGMENT

REFERENCES

[1] D. Scobee and L. Ratliff and R. Dong and H. Ohlsson and M. Verhaegen and S. S. Sastry, "Nuclear norm minimization for blind subspace identification (N2BSID)," IEEE 54th Annual Conference on Decision and Control (CDC), pp. 2127–2132, December 2015.

[2] Z. Liu and A. Hansson and L. Vandenberghe, "Nuclear norm system identification with missing inputs and outputs," vol. 62, pp.605–612, August 2013.

[3] J. F. Cai and E. J. Candès and Z. Shen, "A singular value thresholding algorithm for matrix completion," SIAM J. Optimiz., vol. 20, pp. 1956–1982, January 2010.

[4] K. Mohan and M. Fazel, "Iterative Reweighted Algorithms for Matrix Rank Minimization," The Journal of Machine Learning Research, vol. 13, pp. 3441–3473, November 2012.

[5] K. Konishi and K. Uruma and T. Takahashi and T. Furukawa, "Iterative Partial Matrix Shrinkage Algorithm for Matrix Rank Minimization," Signal Processing, vol. 100, pp. 124–131, July 2014.

[6] Brian Eriksson and Laura Balzano and Robert Nowak, "High-Rank Matrix Completion," Proc. International Conference on Artificial Intelligence and Statistics (AISTATS-12), pp. 373–381, April 2012.

[7] Greg Ongie and Rebecca Willett and Robert D. Nowak and Laura Balzano, "Algebraic Variety Models for High-Rank Matrix Completion," Proc. International Conference on Machine Learning, pp. 2691–2700, 2017.

[8] R. Sasaki and K. Konishi and T. Takahashi and T. Furukawa, "Low-Rank and Locally Linear Embedding Approach to Image Inpainting," Proc. IEEE International Conference on Visual Communications and Image Processing, 2018.

TABLE III

COMPARISON OF SNR WITH SIGNAL RESTORATION PROBLEM WITH SUBSPACE CLUSTERING [dB].

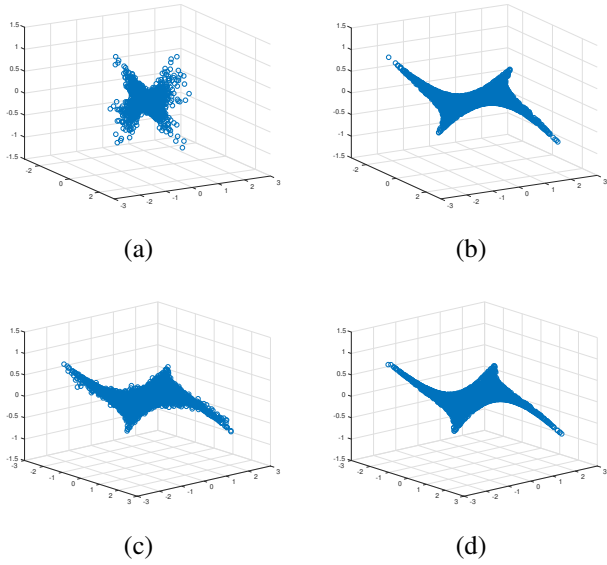| Problem			Algorithm			
N	M	s	IPMS	VMC($d=2$)	VMC($d=3$)	Algorithm 2
50	100	5	17.32	28.81	12.70	31.96
100	200	10	26.68	33.58	10.42	48.78

(a)

(b)

(c)

(d)

Fig. 3. 3D projection of column vectors (a) with missing entries, (b) with complete entries, (c) restored by the nuclear norm minimization, (d) restored by the proposed algorithm in equation (9).
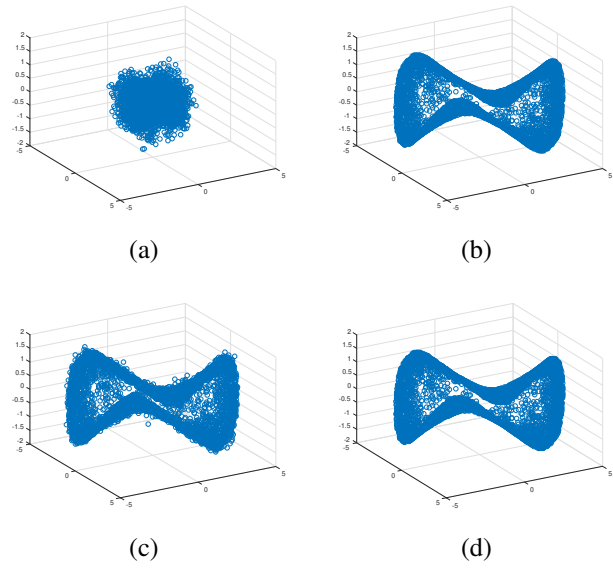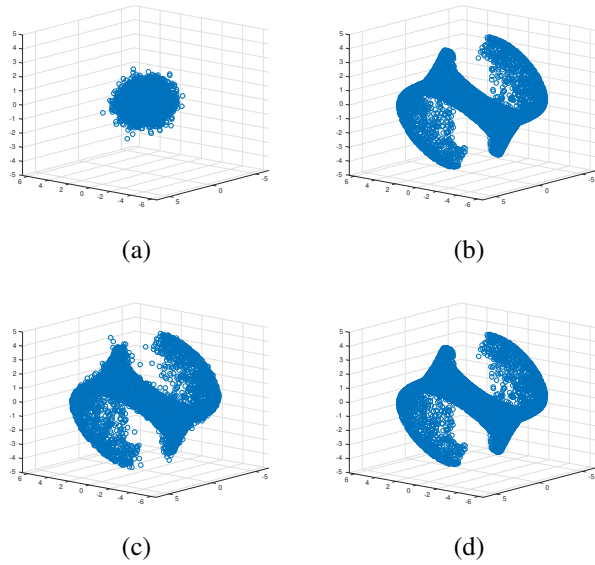

(a)

(b)

(c)

(d)

Fig. 4. 3D projection of column vectors (a) with missing entries, (b) with complete entries, (c) restored by the nuclear norm minimization, (d) restored by the proposed algorithm in equation (10).

(a)

(b)

(c)

(d)

Fig. 5. 3D projection of column vectors (a) with missing entries, (b) with complete entries, (c) restored by the nuclear norm minimization, (d) restored by the proposed algorithm in equation (11).