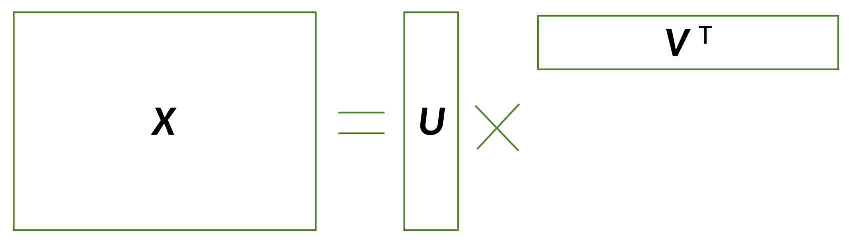
Online High-Rank Matrix Completion

Jicong Fan   Madeleine Udell
Cornell University

Low-rank matrix completion (LRMC)



(a) Low-rank matrix

$$\begin{bmatrix} 0.86 & ? & ? & 0.92 \\ ? & -11.99 & -3.98 & ? \\ -3.60 & ? & 4.90 & ? \\ ? & -5.64 & ? & ? \end{bmatrix} \Rightarrow \begin{bmatrix} 0.86 & -3.52 & -1.17 & 0.92 \\ 2.92 & -11.99 & -3.98 & 3.14 \\ -3.60 & 14.77 & 4.90 & -3.87 \\ 1.37 & -5.64 & -1.87 & 1.48 \end{bmatrix}$$

(b) Matrix completion[1]

Applications

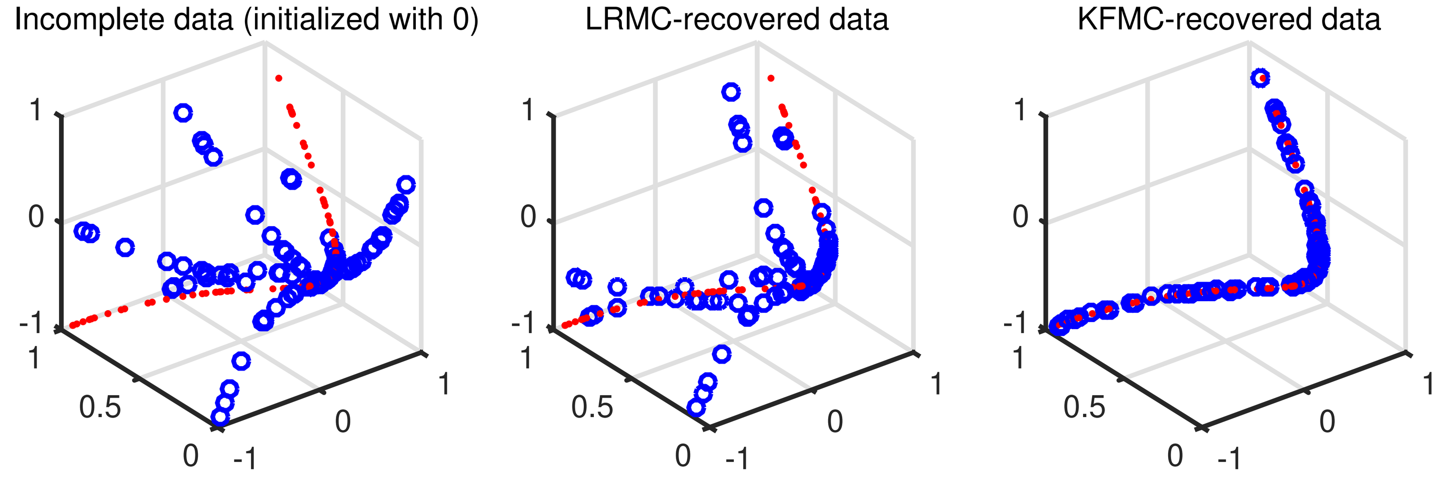
- collaborative filtering (recommendation system)
- classification (especially on incomplete data)
- image inpainting
- other data pre-processing tasks

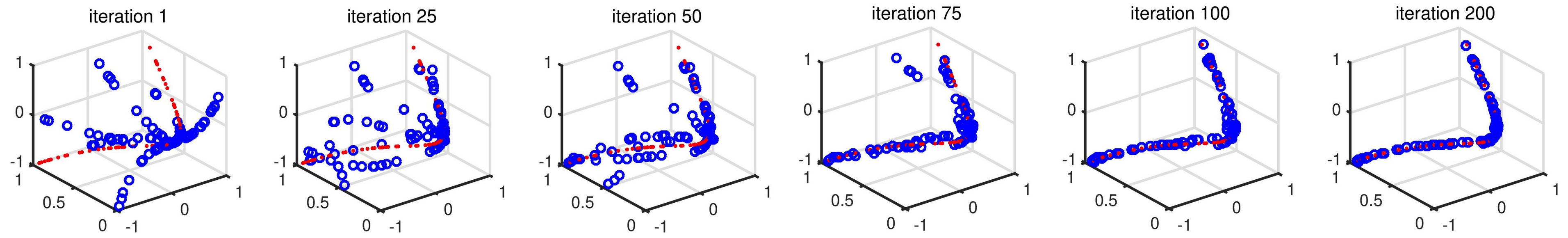
High-rank matrix completion (HRMC)

Example

- the columns of $X \in \mathbb{R}^{3\times 100}$ are generated by $x_1 = s,\ x_2 = s^2,\ x_3 = s^3,\quad s \sim \mathcal{U}(-1,1)$;
- randomly remove one entry of each column of X.

KFMC: kernel factorization based matrix completion





High-rankness of polynomial model in data space

Assumption

The columns of $X \in \mathbb{R}^{m\times n}$ are given by

$$x = \mathbf{f}(s) = [f_1(s), f_2(s), \cdots, f_m(s)]^\top. \quad (1)$$

$s \in \mathbb{R}^d$ ($d \ll m < n$): uncorrelated random variables.

$\{f_i : \mathbb{R}^d \to \mathbb{R}\}_{i=1}^m$: p-order polynomials with random coefficients.

Lemma 1

Suppose the columns of X satisfy (1). Then with probability 1,

$$\mathrm{rank}(X) = \min\{m, n, \binom{d+p}{p}\}.$$

*When p is large, X is of high-rank or even full-rank and hence cannot be recovered by LRMC.

Low-rankness in polynomial feature space

Let $\phi : \mathbb{R}^m \to \mathbb{R}^{\bar{m}}$ be a q-order polynomial feature map: $\phi(x) = \{x_1^{\mu_1} \cdots x_m^{\mu_m}\}_{|\mu|\leq q}$, $\quad \bar{m} = \binom{m+q}{q}$.

Theorem 1

Suppose the columns of X satisfy (1). Then with probability 1,

$$\mathrm{rank}(\phi(X)) = \min\{\bar{m}, n, \binom{d+pq}{pq}\}.$$

*$\phi(X)$ is of low-rank when $\bar{m}$ and n are sufficiently large.

KFMC

Factorization in feature space

$$\underset{X,D,Z}{\text{minimize}}\ \tfrac{1}{2}\|\phi(X) - \phi(D)Z\|_F^2 + \tfrac{\alpha}{2}\|\phi(D)\|_F^2 + \tfrac{\beta}{2}\|Z\|_F^2$$
$$\text{subject to } X_{ij} = M_{ij},\ (i,j) \in \Omega.$$

$D \in \mathbb{R}^{m\times r}$, $Z \in \mathbb{R}^{r\times n}$, and $r = \binom{d+pq}{pq}$.

Kernel representation

$$\underset{X,D,Z}{\text{minimize}}\ \tfrac{1}{2}\mathrm{Tr}\left(K_{XX} - 2K_{XD}Z + Z^\top K_{DD} Z\right) + \tfrac{\alpha}{2}\mathrm{Tr}(K_{DD}) + \tfrac{\beta}{2}\|Z\|_F^2.$$
$$\text{subject to } X_{ij} = M_{ij},\ (i,j) \in \Omega.$$

Online KFMC

- streaming data
- large-scale problem
- data structure changes with time

Out-of-sample extension

- use trained model to recover new data without re-training

Space and time complexity

	Space complexity	Time complexity (per iteration)
VMC [2]	$O(n^2)$	$O(n^3 + mn^2)$
NLMC [3]	$O(n^2)$	$O(n^3 + mn^2)$
KFMC	$O(n^2)$	$O(mn^2 + rmn)$
OL-KFMC	$O(mr + r^2)$	$O(r^3)$
OSE-KFMC	$O(mr + r^2)$	$O(mr)$

* $r, m \ll n$.

Generalization for union of subspaces

Suppose the columns of $X \in \mathbb{R}^{m\times n}$ are given by

$$\{x^{\{k\}} = \mathbf{f}^{\{k\}}(s^{\{k\}})\}_{k=1}^u, \quad (2)$$

where $s^{\{k\}} \in \mathbb{R}^d$ ($d \ll m < n$) are random variables and $\mathbf{f}^{\{k\}} : \mathbb{R}^d \to \mathbb{R}^m$ are p-order polynomial functions for each $k = 1, \ldots, u$. We have

$$\begin{cases} \mathrm{rank}(X) = \min\{m, n, u\binom{d+p}{p}\} \\ \mathrm{rank}(\phi(X)) = \min\{\bar{m}, n, r\} \end{cases}, \quad (3)$$

where $\bar{m} = \binom{m+q}{q}$ and $r = u\binom{d+pq}{pq}$.

Sampling complexity (rule of thumb)

KFMC with q-order polynomial kernel:

$$\rho_{\mathrm{KFMC}} > \left(r/n + r/\bar{m} - r^2/n/\bar{m}\right)^{1/q}, \qquad \bar{m} = \binom{m+q}{q},\ r = u\binom{d+pq}{pq}.$$

LRMC:

$$\rho_{\mathrm{LRMC}} > \left((m+n)r_x - r_x^2\right)/(mn), \qquad r_x = \min\{m, n, u\binom{d+p}{p}\}.$$

Example

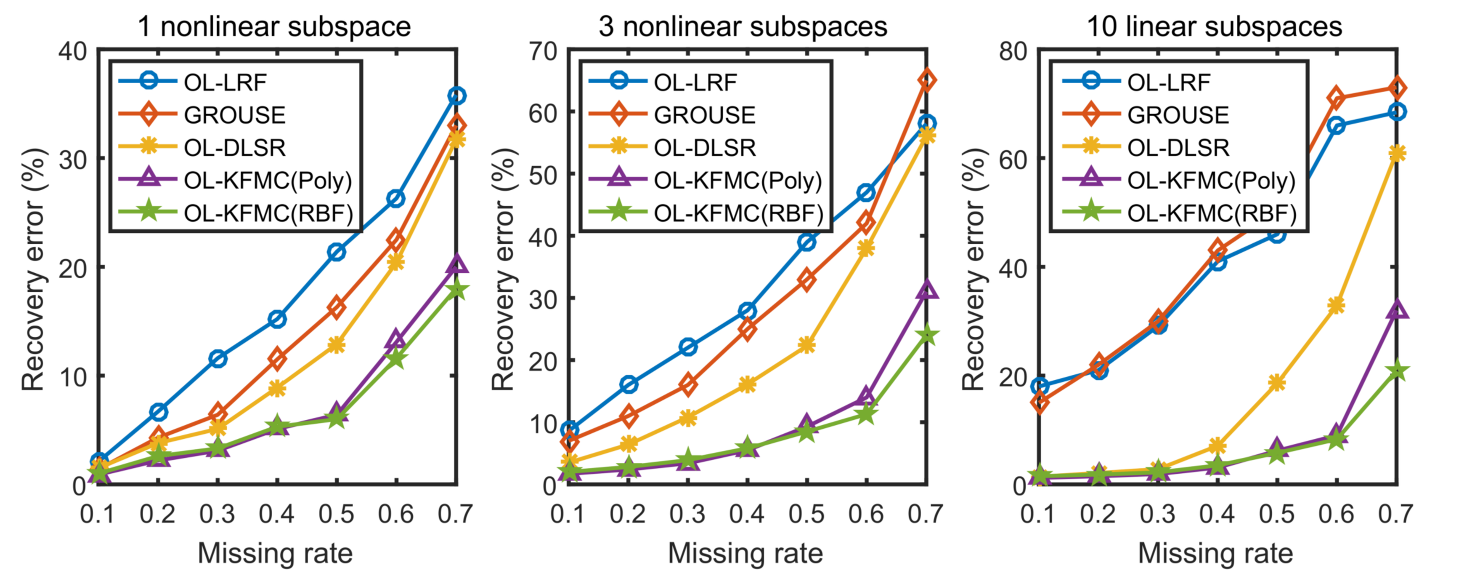
Suppose $m = 20$, $n = 300$, $d = 2$, $p = 2$, $u = 3$, and $q = 2$. Then $\rho_{\mathrm{LRMC}} > 0.91$ and $\rho_{\mathrm{KFMC}} > 0.56$.

Experiments on synthetic data

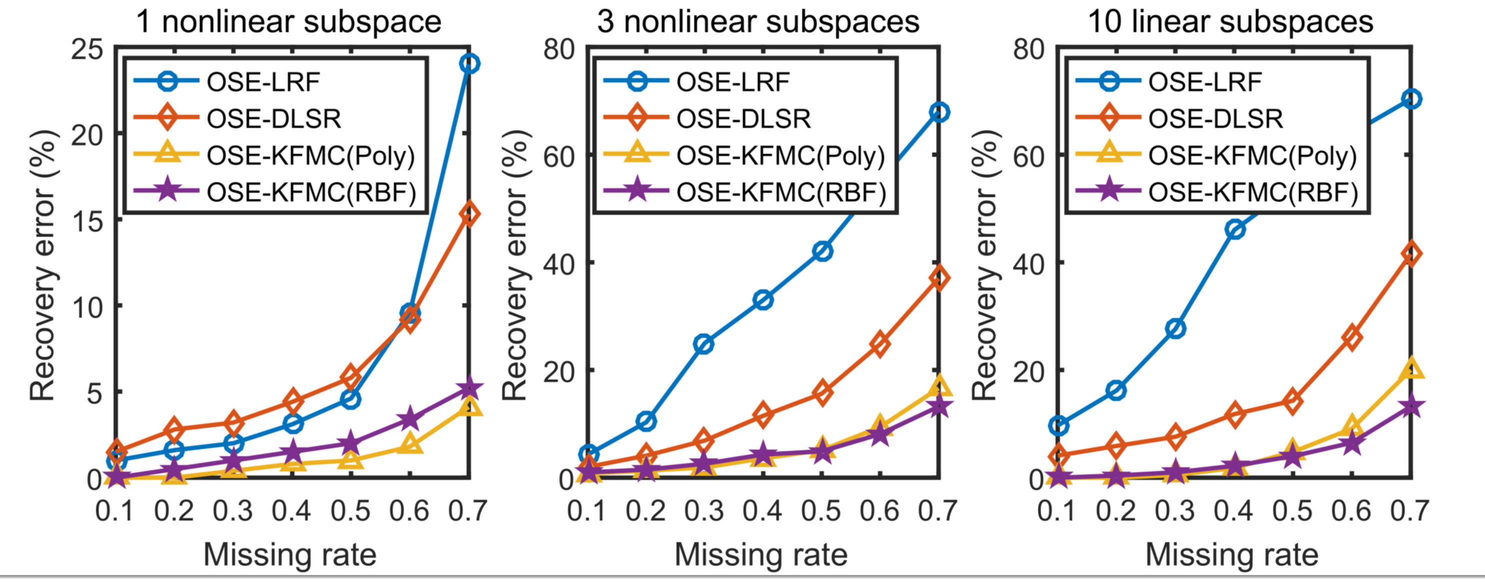
Let $x = \mathbf{f}(s)$, where $s \sim \mathcal{U}(0,1)$ and $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^{30}$ is a p-order polynomial mapping.

- *single nonlinear subspace:* $\quad p = 3,\ u = 1,\ X \in \mathbb{R}^{30\times 100},\ \mathrm{rank}(X) = 19.$
- *union of nonlinear subspaces:* $\quad p = 3,\ u = 3,\ X \in \mathbb{R}^{30\times 300},\ \mathrm{rank}(X) = 30.$
- *union of linear subspaces:* $\quad p = 1,\ u = 10,\ X \in \mathbb{R}^{30\times 1000},\ \mathrm{rank}(X) = 30$

Online matrix completion

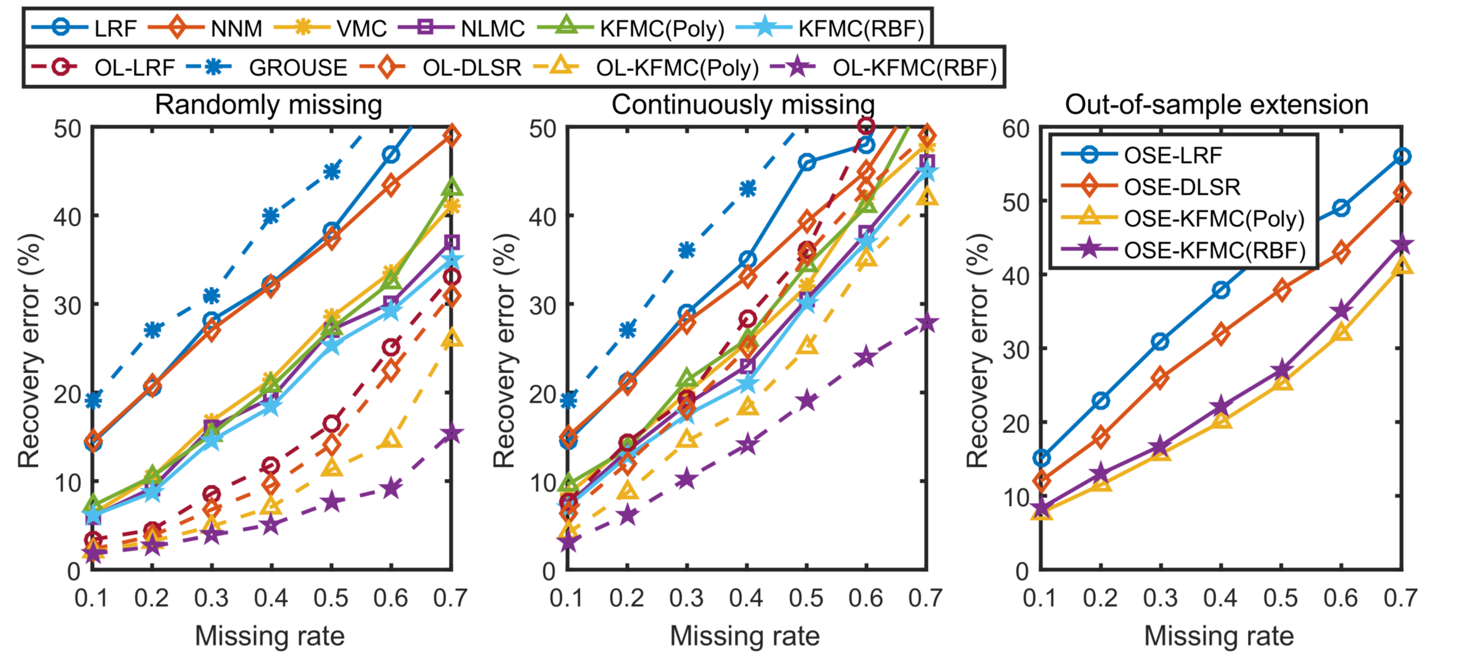


Out-of-sample extension



Experiments on CMU motion capture data

Time-series trajectories of human motions, e.g. walking, running and jumping[4].



Time costs (second) {*matrix size: 62 × 3392; RSVD: randomized singular value decomposition}

VMC[RSVD]	370	NLMC[RSVD]	610
KFMC(Poly)	170	KFMC(RBF)	190
OL-KFMC(Poly)	16	OL-KFMC(RBF)	19

Reference

[1] Emmanuel J. Candès and Benjamin Recht. Foundations of Computational Mathematics, 2009.
[2] Greg Ongie, Rebecca Willett, Robert D. Nowak, and Laura Balzano. ICML, 2017.
[3] Jicong Fan and Tommy W.S. Chow. Pattern Recognition, 2018.
[4] Ehsan Elhamifar. NIPS, 2016.