Matrix Completion on Graphs

Vassilis Kalofolias[1], Xavier Bresson[2], Michael Bronstein[3], and Pierre Vandergheynst[1]
[1]EPFL, Switzerland [2]UNIL, Switzerland [3]USI, Switzerland

Matrix Completion

Given a sparse set Ω of observations, find matrix M.
Assumption: M is low rank:

$$\min_{X \in \mathbb{R}^{m \times n}} \operatorname{rank}(X) \quad \text{s.t.} \quad A_\Omega(X) = A_\Omega(M)$$ ✗ NP-hard!

Relax rank with its tightest convex surrogate [1]:

$$\min_{X \in \mathbb{R}^{m \times n}} ||X||_* \quad \text{s.t.} \quad A_\Omega(X) = A_\Omega(M)$$ ✓ Convex

Observations can be noisy:

$$\min_{X \in \mathbb{R}^{m \times n}} \gamma_n ||X||_* + \frac{1}{2} ||A_\Omega(X - M)||_F^2 \qquad (1)$$

Success guaranteed if [2]: $|\Omega| \geq \mathcal{O}(r(m+n)\log^2(n))$ if $m < n$

Adding structure

- How to go beyond standard sparse recovery problem (1)?
  ⇒ Use "structures" [4, 5]
- Our structure: use similarity information for rows / columns

Assumption: columns / rows of M are points on manifolds [3].

- Discrete representation: **graphs**

$$G_r = (V_r, E_r, W_r), \quad G_c = (V_c, E_c, W_c)$$

Rows graph — nodes = rows of M — edges — Edge weights

- Well connected nodes have similar values:

$$\frac{1}{2}\sum_{j,j'} w^c_{jj'} \|x_j - x_{j'}\|_2^2 = \operatorname{tr}(XL_cX^\top) = \|X\|_{\mathcal{D},c}$$ is small

⇒ Final problem to be solved:

$$\min_X \underbrace{\gamma_n \|X\|_*}_{F(X)} + \underbrace{\frac{1}{2}\|A_\Omega(X-M)\|_F^2 + \frac{\gamma_r}{2}\|X\|_{\mathcal{D},r} + \frac{\gamma_c}{2}\|X\|_{\mathcal{D},c}}_{G(X)} \qquad (2)$$

X is sparse in SV domain and structured according to row/column graphs.

Algorithm

ADMM = (split) + (Augmented Lagrangian method)

- Split problem (2):

$$\min_{X,Y \in \mathbb{R}^{m \times n}} F(X) + G(Y) \quad \text{s.t.} \quad X = Y.$$

non-smooth (F) — smooth (G)

- Augmented Lagrangian:

$$\mathcal{L}(X,Y,Z) = F(X) + G(Y) + \langle Z, X - Y\rangle + \frac{\rho}{2}\|X - Y\|_\mathcal{F}^2$$

- To find a saddle point of $\mathcal{L}$ alternate between:

$$X^{k+1} = \arg\min_X \mathcal{L}(X, Y^k, Z^k) \qquad \text{threshold s-values}$$
$$Y^{k+1} = \arg\min_Y \mathcal{L}(X^{k+1}, Y, Z^k) \qquad \text{graph smoothing}$$
$$Z^{k+1} = Z^k + \rho(X^{k+1} - Y^{k+1})$$

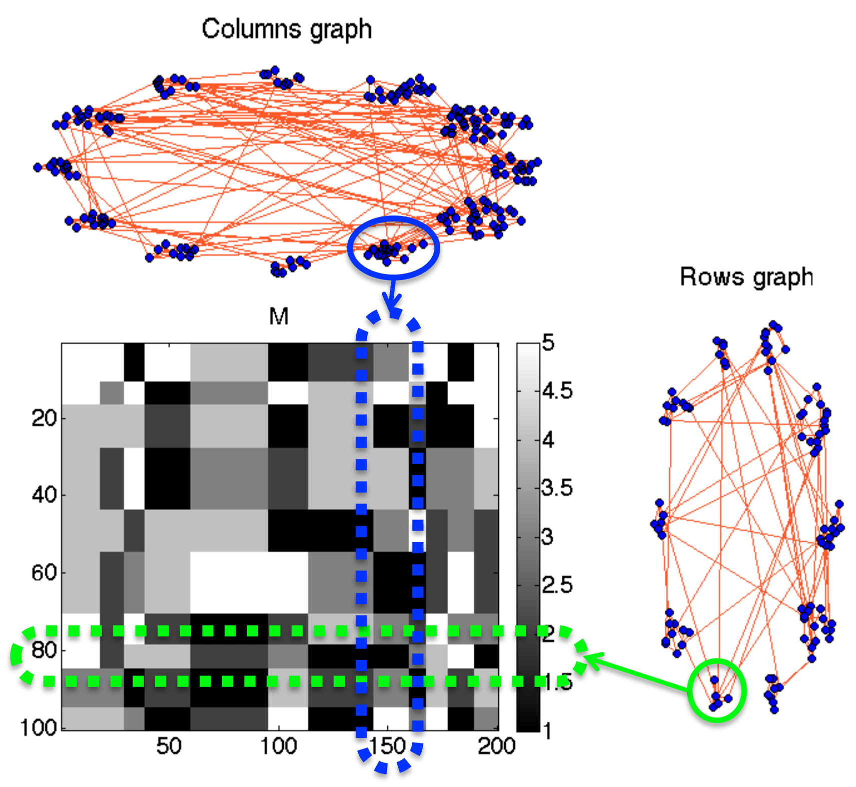
- Approximate solutions of inner steps suffice !

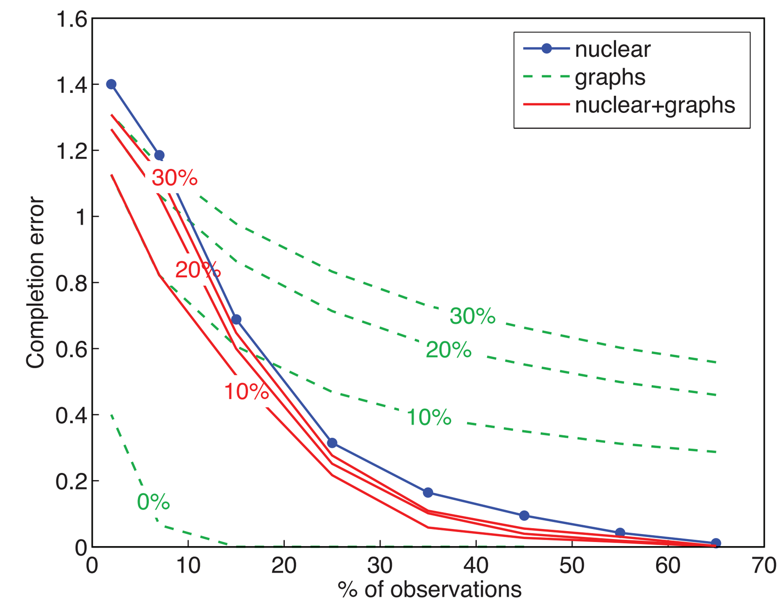
Experiments (artificial data)

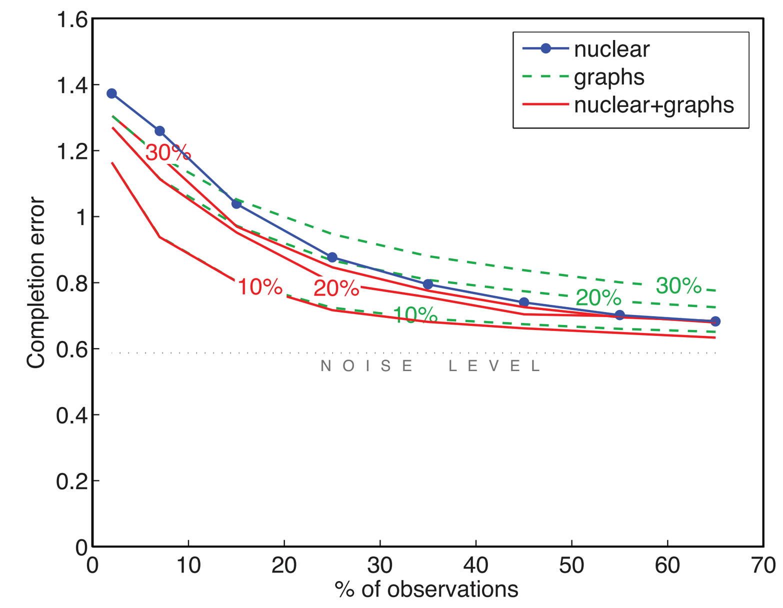
Netflix-like artificial data
- Rank = 10
- Values from {1,…,5}

Structures:
- Communities of users / movies
- Block constant

Add error to:
- Ratings
- Edges (10%, 20%, 30%)







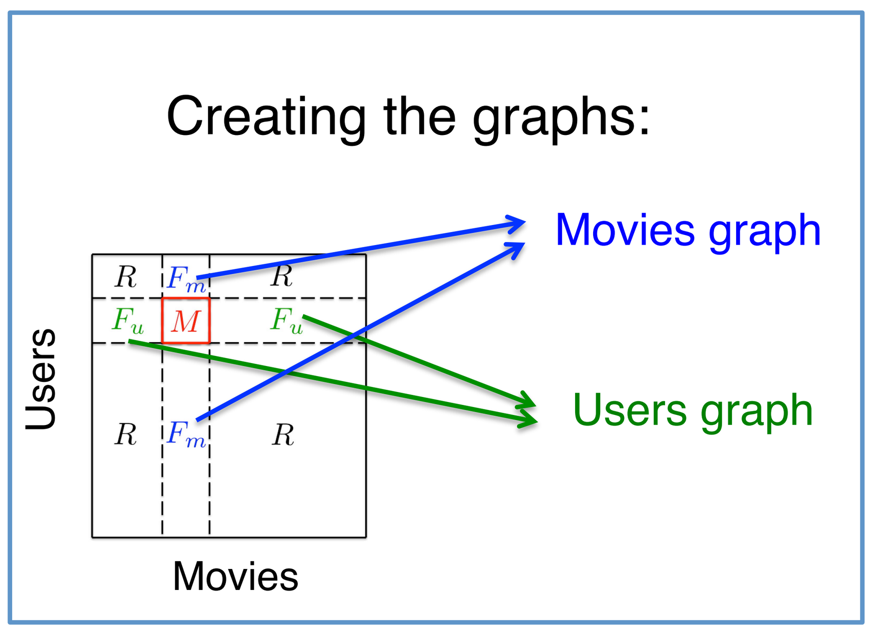
Experiments (Movielens)

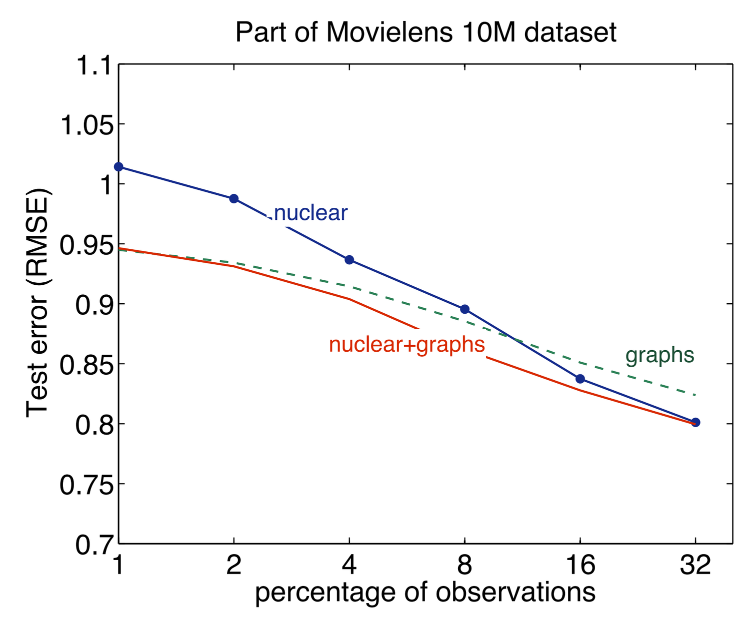
Dataset: 71k users, 10k movies

✗ Graphs are not given: ✓ create them using features!

- Pick 500 users, 500 movies for M.
- Rest of users' ratings: movie features F_m
- Rest of movie ratings: user features F_u

→ distances → edge weights





Conclusions

- Matrix recovery can benefit by row or column similarity information.
- Similarity can be encoded by graphs.
- Convex non-smooth problem solved efficiently by ADMM.
- Method is robust to graph quality.
- Robust to non-uniform sampling (see full paper [6]).

References

[1] E. Candes and B. Recht. Exact matrix completion via convex optimization. FCM, 2009.
[2] B. Recht. A simpler approach to matrix completion. JMLR, 2011.
[3] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Computation, 2003.
[4] R.G. Baraniuk, V. Cevher, M.F. Duarte, C. Hedge. Model-based compressive sensing. IEEE Trans. Information Theory, 2010.
[5] R. Jenatton, J.Y. Audibert, F. Bach. Structured variable selection with sparsity-inducing norms. JMLR, 2011.
[6] V. Kalofolias, X. Bresson, M. Bronstein, P. Vandergheynst. Matrix completion on graphs. arXiv:1408.1717 [cs.LG].

Contact

Vassilis Kalofolias,
PhD student,
LTS2,
vassilis.kalofolias@epfl.ch

Link to full Paper [6]

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

UNIL | Université de Lausanne

Università della Svizzera italiana