Enhanced Matrix Completion with Manifold Learning

Vassilis Kalofolias*, Xavier Bresson*, Michael Bronstein† and Pierre Vandergheynst*
* Signal Processing Laboratory LTS2, Ecole Polytechnique Fédérale de Lausanne (EPFL)
‡ Institute of Computational Science, Faculty of Informatics SI-109, Università della Svizzera Italiana

***Abstract*—We study the problem of matrix completion when information about row or column proximities is available, in the form of weighted graphs. The problem can be formulated as the optimization of a convex function that can be solved efficiently using the alternating direction multipliers method. Experiments show that our model offers better reconstruction than the standard method that only uses a low rank assumption, especially when few observations are available.**

I. Introduction

In matrix completion we have a set of signals in matrix form $M \in \mathbb{R}^{m\times n}$. They are only sparsely observed, i.e. a small set Ω of the elements of M are known, and the goal is to recover the rest. In order for the problem to be solvable, additional assumptions are needed, the most common one being that the matrix is low rank. The problem of finding the matrix of minimum rank that agrees to the given observed values is NP-hard. However, replacing $\operatorname{rank}(X)$ with its complex surrogate nuclear norm $||X||_*$ yields the convex problem

$$\min_{X\in\mathbb{R}^{m\times n}} \|X\|_* \quad \text{s.t.} \quad A_\Omega(X) = A_\Omega(M), \tag{1}$$

where A_Ω denotes the operator that keeps only the observed values of a matrix. Under the assumption that M is sufficiently incoherent, if the indices Ω are uniformly distributed and $|\Omega|$ is sufficiently large, solving the last problem gives an exact solution [1].

II. Matrix completion on graphs

Low rank implies the linear dependence of rows/columns. However, this dependence is unstructured. In many situations, the rows/columns of matrix M possess additional structure that can be incorporated into the completion problem in the form of regularization. In this work, we assume that the row and column signals reside on manifolds approximated by two given graphs. The weight w^c_{ij} of an edge of the columns' graph $G_c = (V_c, E_c, W_c)$ represents how close are the two adjacent columns x_i and x_j. Similarly for the rows, that are nodes of the corresponding graph $G_r = (V_r, E_r, W_r)$. More formally, we want

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

to be small [2], where $L_c = D_c - W_c$ is the Laplacian of the column graph G_c, $D_c = \operatorname{Diag}(\sum_{j'=1}^n w^c_{jj'})$, and $\|\cdot\|_{\mathcal{D},c}$ is the graph Dirichlet semi-norm for columns. Similarly, for the rows we get a corresponding expression $\operatorname{tr}(X^\top L_r X) = \|X\|_{\mathcal{D},r}$ with the Laplacian L_r of the row graph G_r.

These smoothness terms are added to the matrix completion problem as regularization terms, and the final optimization problem we solve is [3]

$$\min_X \gamma_n||X||_* + \frac{1}{2}||A_\Omega(X-M)||_F^2 + \frac{\gamma_r}{2}\|X\|_{\mathcal{D},r} + \frac{\gamma_c}{2}\|X\|_{\mathcal{D},c}, \tag{3}$$

where γ_n, γ_r, γ_c are parameters that can be chosen with model selection methods like cross validation.

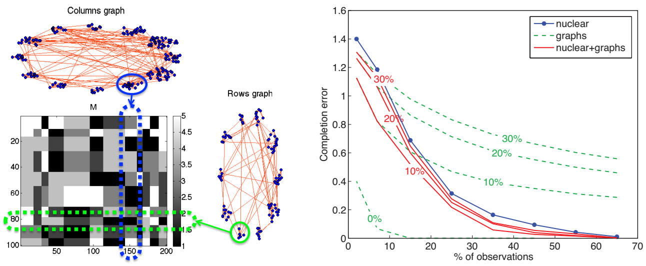

Fig. 1. Artificial low rank dataset (left) and matrix recovery error versus observation level (right). Percentage of erroneous edges in graphs is shown on top of green and red lines.

III. Convex optimization

Problem (3) is a non-smooth convex optimization problem that can not be solved efficiently with a direct approach. Using a splitting approach we can write the objective function of (3) as $F(X)+G(X)$, where $F(X)$ contains the non-smooth part (i.e. the nuclear norm) and $G(X)$ contains the smooth part (all other terms). Then we can equivalently solve the problem

$$\min_{X,Y} \quad F(X) + G(Y) \qquad \text{s.t.} \quad X = Y \tag{4}$$

using an augmented Lagrangian method. The combination of these techniques is known as the Alternating Direction Method of Multipliers (ADMM) [4] that is efficient even when the intermediate steps are only solved approximately.

IV. Experiments

We construct a synthetic block constant rank-10 matrix. Its columns (rows) exhibit a community structure corresponding to an ideal graph connecting pairs of columns (resp. rows) that are identical. We add different levels of erroneous edges (10%, 20%, 30%) sampled from an Erdös Rényi graph. For each of the graphs we perform reconstruction using different levels of observed values of the matrix.

The reconstruction error (RMSE) is reported in fig 1. The blue line corresponds to the standard method making use of only the low rank assumption. The red lines correspond to the results of our method that incorporates structure information of different fidelity levels, corresponding to the percentages of erroneous edges of the graphs. The green lines depict the results for $\gamma_n = 0$, when only graphs and not the low-rank assumption are used. Evidently, reconstruction can benefit when structure information is added to the standard low rank assumption, while the method is robust to the quality of graphs.

References

[1] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *FCM*, vol. 9, no. 6, pp. 717–772, 2009.
[2] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol.6, 2003.
[3] V. Kalofolias, X. Bresson, M. Bronstein, and P. Vandergheynst, "Matrix completion on graphs," *arXiv preprint arXiv:1408.1717*, 2014.
[4] R. Glowinski and A. Marroco, "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires," *ESAIM*, vol. 9, pp. 41–76, 1975.