An IRLS Approach for Low-Rank Matrix Factorization

Paris V. Giampouras*, Athanasios A. Rontogiannis*, Konstantinos D. Koutroumbas*
*IAASARS, National Observatory of Athens, GR-15236, Penteli, Greece
email: {parisg,tronto,koutroum}@noa.gr

***Abstract*—Iterative reweighted methods for sparse recovery and low-rank matrix estimation have been flourished in recent years. As it has been shown, these approaches offer significant merits when it comes both to the recovery performance and the computational efficiency of the derived algorithms. On the other hand, the use of low-rank matrix factorization for encoding low-rankness has been pivotal in numerous machine learning applications that involve big and high dimensional data lately. In this work, a novel iterative reweighted approach for low-rank matrix factorization is presented. The proposed approach gives rise to scalable algorithms for Schatten-p norm minimization with $0 < p \leq 1$. The efficiency of the resulting schemes is empirically verified in a matrix completion problem.**

I. Introduction

Low-rank matrix factorization (LRMF) has been at the heart of several signal processing and machine learning applications over the past few years. LRMF appears in the modeling process of ubiquitous learning tasks such as blind source separation, dictionary learning, subspace learning, etc., [1]. At the same time LRMF offers an efficient way in an effort to find low-rank representations of large-scale and high-dimensional data. The latter has recently given rise to a huge bulk of research works whose primary goal is to provide scalable algorithms for low-rank matrix estimation.

Iterative reweighted methods are old, but they have been recently utilized for seeking parsimonious representations of data. In the seminal work reported in [2], iterative reweighted least squares (IRLS) was put forth for sparse vector recovery providing an efficient alternative to traditional ℓ_1 norm minimization. The IRLS rationale has been recently extended for low-rank matrix estimation in [3], [4]. As it was shown, the resulting algorithms offer competitive runtimes and improved recovery performance as compared to nuclear norm minimization algorithms. In [5], an iterative reweighted type algorithm has been developed to solve a nonconvex and nonsmooth low-rank promoting minimization problem offering improved low-rank matrix recovery results than relevant convex apporaches.

Capitalizing on the advent of iterative reweighted approaches, herein we propose a novel framework for iterative reweighted low-rank matrix factorization. The proposed approach is applied on tight-upper bounds of Schatten-p norms leading to efficient alternative closed-formed updates for the matrix factors at a low computational cost. The resulting algorithms are singular value decomposition (SVD) free and as is empirically shown for the case of matrix completion, they offer promising results in terms of recovery performance and runtime.

II. Problem formulation

LRMF methods are based on the fact that a low-rank matrix can be expressed as $\mathbf{X} = \mathbf{U}\mathbf{V}^T$ ($\mathbf{U} \in \mathcal{R}^{m\times r}, \mathbf{V} \in \mathcal{R}^{n\times r}$) with the inner dimension r of the involved matrices quite smaller than the outer dimensions, i.e., $r \ll \min(m,n)$. An inherent weakness of this approach is that an additional variable is coming up, i.e., the inner dimension r of the factorization that must be learned from the data. In that vein, among other LRMF approaches, the minimization of a tight upper-bound of the nuclear norm defined as

$$\|\mathbf{X}\|_* = \min_{\mathbf{X}=\mathbf{U}\mathbf{V}^T, \boldsymbol{u}_i, \boldsymbol{v}_i} \sum_{i=1}^{d} \|\boldsymbol{u}_i\|_2 \|\boldsymbol{v}_i\|_2 = \min_{\mathbf{X}=\mathbf{U}\mathbf{V}^T, \boldsymbol{u}_i, \boldsymbol{v}_i} \frac{1}{2}\left(\sum_{i=1}^{d} \|\boldsymbol{u}_i\|_2^2 + \|\boldsymbol{v}_i\|_2^2\right) \quad (1)$$

has been widely adopted, [1], [6], where d is an overestimate of r, $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the ith columns of $\mathbf{U}$ and $\mathbf{V}$ and $\|\cdot\|_*, \|\cdot\|_2$ denote the nuclear and the ℓ_2 norms, respectively. We may generalize the above result for the Schatten-p norm (the proof is based on arguments of [7] and is not provided due to space limitations.).

Proposition 1: Schatten-p norms are bounded above as follows,

$$\|\mathbf{X}\|_{\mathcal{S}_p}^p = \min_{\mathbf{X}=\mathbf{U}\mathbf{V}^T, \boldsymbol{u}_i, \boldsymbol{v}_i} \sum_{i=1}^{d} (\|\boldsymbol{u}_i\|_2 \|\boldsymbol{v}_i\|_2)^p = \min_{\mathbf{X}=\mathbf{U}\mathbf{V}^T, \boldsymbol{u}_i, \boldsymbol{v}_i} \frac{1}{2^p} \sum_{i=1}^{d} \left(\|\boldsymbol{u}_i\|_2^2 + \|\boldsymbol{v}_i\|_2^2\right)^p \quad (2)$$

where $\|\cdot\|_{\mathcal{S}_p}$ is the Schatten-p norm and $0 < p \leq 1$.

By using (2) the matrix completion problem is written as follows,

$$\min_{\mathbf{U},\mathbf{V}} \|\mathcal{P}_\Omega(\mathbf{Y}) - \mathcal{P}_\Omega(\mathbf{U}\mathbf{V}^T)\|_F^2 + \lambda \sum_{i=1}^{d} \left(\|\boldsymbol{u}_i\|_2^2 + \|\boldsymbol{v}_i\|_2^2\right)^p, \quad (3)$$

where $\|\cdot\|_F$ is the Frobenius norm, and $\mathcal{P}_\Omega$ denotes the sampling operator on the set Ω of indexes of data matrix $\mathbf{Y}$ where information is available.

III. The Iterative reweighted type algorithm

Next, we apply ideas stemming from *iterative reweighting methods for low-rank matrix recovery* and rewrite the regularization term in (3) as the sum of reweighted Frobenius norms, i.e.,

$$\min_{\mathbf{U},\mathbf{V}} \|\mathcal{P}_\Omega(\mathbf{Y}) - \mathcal{P}_\Omega(\mathbf{U}\mathbf{V}^T)\|_F^2 + \lambda \left(\|\mathbf{U}\mathbf{W}_{(\mathbf{U},\mathbf{V})}^{\frac{1}{2}}\|_F^2 + \|\mathbf{V}\mathbf{W}_{(\mathbf{U},\mathbf{V})}^{\frac{1}{2}}\|_F^2\right) \quad (4)$$

with $\mathbf{W}_{(\mathbf{U},\mathbf{V})}$ defined as

$$\mathbf{W}_{(\mathbf{U},\mathbf{V})} = \mathrm{diag}\Big(\left(\|\boldsymbol{u}_1\|_2^2 + \|\boldsymbol{v}_1\|_2^2\right)^{p-1}, \left(\|\boldsymbol{u}_2\|_2^2 + \|\boldsymbol{v}_2\|_2^2\right)^{p-1}, \ldots, \left(\|\boldsymbol{u}_d\|_2^2 + \|\boldsymbol{v}_d\|_2^2\right)^{p-1} \Big). \quad (5)$$

The resulting Algorithm 1, arises by minimizing (4) w.r.t. matrix factors $\mathbf{U}$ and $\mathbf{V}$ in an alternating fashion, making use of the typical IRLS scheme, [8].

IV. A matrix completion experiment

Next we aim at illustrating the behavior of the proposed Algorithm 1 when it comes to the estimation performance and the speed of convergence for $p = \frac{1}{2}$. In this regard, for the case of the 100K Movielens dataset, the state-of-the-art IRNN [5] and softImpute-ALS [6] algorithms are utilized for comparison purposes (Fig. 1).

Algorithm 1: Alternating iterative reweighted least squares matrix completion (AIRLS-MC) algorithm
Input: $\mathbf{Y}, \lambda$ Initialize: $k = 0, \mathbf{U}_0, \mathbf{V}_0, \mathbf{W}_{(\mathbf{U}_0,\mathbf{V}_0)}$ **repeat** $\mathbf{U}_{k+1} = \mathbf{U}_k - \left(\mathcal{P}_\Omega\left(\mathbf{U}_k\mathbf{V}_k^T - \mathbf{Y}\right)\mathbf{V}_k + \mathbf{U}_k\mathbf{W}_{(\mathbf{U}_k,\mathbf{V}_k)}\right)\left(\mathbf{V}_k^T\mathbf{V}_k + \lambda\mathbf{W}_{(\mathbf{U}_k,\mathbf{V}_k)}\right)^{-1}$ $\mathbf{V}_{k+1} = \mathbf{V}_k - \left(\mathcal{P}_\Omega\left(\mathbf{V}_k\mathbf{U}_{k+1}^T - \mathbf{Y}^T\right)\mathbf{U}_{k+1} + \mathbf{V}_k\mathbf{W}_{(\mathbf{U}_{k+1},\mathbf{V}_k)}\right)\left(\mathbf{U}_{k+1}^T\mathbf{U}_{k+1} + \lambda\mathbf{W}_{(\mathbf{U}_{k+1},\mathbf{V}_k)}\right)^{-1}$ $k = k + 1$ **until** *convergence* Output: $\hat{\mathbf{U}} = \mathbf{U}_{k+1}, \hat{\mathbf{V}} = \mathbf{V}_{k+1}$

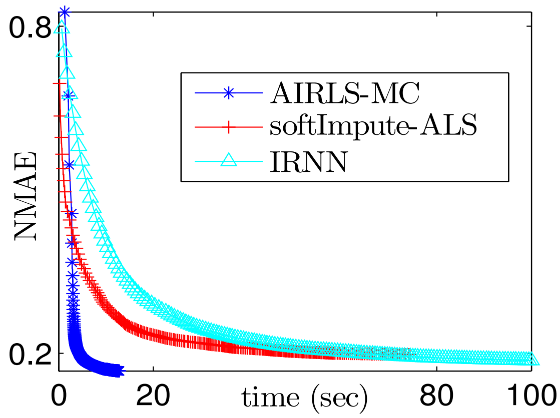

Fig. 1. NMAE vs time evolution (up to 100secs) of the proposed , softImpute-ALS [6] and IRNN [5] on the Movielens 100K validation dataset. AIRLS-MC achieved lower NMAE than its matrix factorization counterpart softImpute-ALS and slightly lower NMAE than IRNN. Moreover, AIRLS-MC required less runtime than both softImpute-ALS and IRNN.

Acknowledgment

We acknowledge support of this work by the project "PROTEAS II - Advanced Space Applications for Exploring the Universe of Space and Earth" (MIS 5002515) which is implemented under the Action Reinforcement of the Research and Innovation InfrastructurE", funded by the Operational Programme Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund).

References

[1] B. Haeffele, E. Young, and R. Vidal, "Structured low-rank matrix factorization: Optimality, algorithm, and applications to image processing," in *International Conference on Machine Learning*, 2014, pp. 2007–2015.

[2] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Gntrk, "Iteratively reweighted least squares minimization for sparse recovery," *Communications on Pure and Applied Mathematics*, vol. 63, no. 1, pp. 1–38. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.20303

[3] K. Mohan and M. Fazel, "Iterative reweighted algorithms for matrix rank minimization," *Journal of Machine Learning Research*, vol. 13, no. Nov, pp. 3441–3473, 2012.

[4] M. Fornasier, H. Rauhut, and R. Ward, "Low-rank matrix recovery via iteratively reweighted least squares minimization," *SIAM Journal on Optimization*, vol. 21, no. 4, pp. 1614–1640, 2011.

[5] C. Lu, J. Tang, S. Yan, and Z. Lin, "Nonconvex nonsmooth low rank minimization via iteratively reweighted nuclear norm," *IEEE Transactions on Image Processing*, vol. 25, no. 2, pp. 829–839, 2016.

[6] T. Hastie, R. Mazumder, J. D. Lee, and R. Zadeh, "Matrix completion and low-rank SVD via fast alternating least squares." *Journal of Machine Learning Research*, vol. 16, pp. 3367–3402, 2015.

[7] M. V. Örnhag, C. Olsson, and A. Heyden, "Bilinear parameterization for differentiable rank-regularization," *arXiv preprint arXiv:1811.11088*, 2018.

[8] P. V. Giampouras, A. A. Rontogiannis, and K. D. Koutroumbas, "Alternating iteratively reweighted least squares minimization for low-rank matrix factorization," *IEEE Transactions on Signal Processing*, vol. 67, no. 2, pp. 490–503, Jan 2019.