Harmonic Mean Iteratively Reweighted Least Squares for Low-Rank Matrix Recovery

Christian Kümmerle
Department of Mathematics
Technische Universität München
Email: christian.kuemmerle@ma.tum.de

Juliane Sigl
Department of Mathematics
Technische Universität München
Email: juliane.sigl@ma.tum.de

A. Introduction

The problem of recovering a low-rank matrix from incomplete linear measurements has gained considerable attention in the last few years due to the omnipresence of low-rank models in different areas of science and applied mathematics [1]. The problem of identifying and reconstructing a low-rank matrix from only few given entries is called *matrix completion* and is a well-known instance of the low-rank matrix recovery problem.

Although the problem is NP-hard in general, several efficient algorithms have been proposed that allow for provable recovery for many classes of matrices, given many measurement operators. The most popular techniques include *nuclear norm minimization* (NNM) [2], [3], which solves a convex relaxation of rank minimization problem, and *alternating minimization* [4]. For NNM, recovery guarantees have been shown for a number of measurements on the order of the information theoretical lower bound $r(d_1 + d_2 - r)$, if r denotes the rank of a $d_1 \times d_2$-matrix [3]. More recently, comparable guarantees have also been derived for several non-convex algorithmic approaches [5]–[7], which are often preferred in practice because of their higher empirical recovery rate and their more efficient implementation.

B. Our Approach

In this spirit, we propose [8] a new Iteratively Reweighted Least Squares (IRLS) algorithm for the low-rank matrix recovery problem striving to minimize the non-convex Schatten-p quasi-norm

$$\min \|X\|_{S_p}^p \text{ subject to } \Phi(X) = Y \tag{1}$$

for $0 < p < 1$, with $\Phi : \mathbb{R}^{d_1 \times d_2} \to \mathbb{R}^m$ being the linear measurement operator. We call it *Harmonic Mean Iteratively Reweighted Least Squares* (HM-IRLS) as its weight matrices can be interpreted as the harmonic mean of left- and right-sided weight matrices that have been discussed in [9], [10] for previous IRLS algorithms aiming to solve the low-rank recovery problem. Unlike [9], [10], our approach uses the information in both the column and the row space of the iterates, which are defined for $n = 1, 2, \dots$ as the minimizers of weighted least squares problems such that

$$X^{(n+1)} = \operatorname{argmin}_{X \in \mathbb{R}^{d_1 \times d_2}, \Phi(X) = Y} \|X_{\text{vec}}\|_{\ell_2(\widetilde{W}^{(n+1)})}^2,$$

given the harmonic mean weight matrix $\widetilde{W}^{(n+1)} = 2\left[U^{(n)}(\bar{\Sigma}^{(n)})^{2-p}U^{(n)*} \oplus V^{(n)}(\bar{\Sigma}^{(n)})^{2-p}V^{(n)*}\right]^{-1}$. In the definition of $\widetilde{W}^{(n+1)}$, the SVD of $X^{(n)} = U^{(n)}\Sigma^{(n)}V^{(n)}$ appears, $\bar{\Sigma}^{(n)}$ is a diagonal matrix containing the smoothed singular values $\bar{\Sigma}_{ii}^{(n)} = (\Sigma_{ii}^{(n)2} + \epsilon^{(n)2})^{\frac{1}{2}}$ with the smoothing parameter $\epsilon^{(n)} = \min(\epsilon^{(n-1)}, \Sigma_{r+1,r+1}^{(n-1)})$; we further use the convention $A \oplus B = \mathbf{I}_{d_2} \otimes A + B \otimes \mathbf{I}_{d_1}$ and the tensor product $\otimes$ in the standard bases. We claim that this choice of the weight matrices $\widetilde{W}^{(n)}$ has several favourable properties for non-convexity parameters $p < 1$. This is the case as HM-IRLS achieves a better alignment of the left-singular and right-singular vectors of the iterates with the ones of the low-rank matrix to be recovered.

C. Theoretical Results

We extend the theoretical guarantees of previous low-rank IRLS algorithms, some of which are based on *null space properties* of the measurement operator, to HM-IRLS. More precisely, we show convergence of the sequence of iterates $(X^{(n)})_{n\in\mathbb{N}}$ to stationary points of the smoothed Schatten-p functional $\sum_{i=1}^{\min d_1,d_2}(\bar{\Sigma}_{ii}^{(\bar{n})})^p$ under the linear constraint.

Furthermore, we show that, unlike the related algorithms [9], [10], HM-IRLS exhibits a *local superlinear convergence rate* (of order $2 - p$) in case of convergence to a low-rank matrix under the mentioned assumptions on the measurement operator. For small Schatten-p parameters $p > 0$, this means that convergence rate is almost quadratic.

D. Numerical Experiments

We conducted extensive numerical experiments comparing the efficiency of HM-IRLS with related algorithms as well as with algorithms based on different concepts (cf. Fig. 1) in terms of needed number of measurements for reconstructing matrices of a given rank. In the experiments, we focussed on the matrix completion setting due to the popularity of this model, even though our theoretical guarantees do not apply directly to this model.

We observe that surprisingly, HM-IRLS needs fewer given entries to complete low rank matrices with high empirical probability than all the state-of-the-art algorithms we included in our experiments. Computationally, the most expensive operation of our algorithm is the solution of a $m \times m$ sparse linear system, if m is the number of given entries. This allows us to recover low-rank matrices up to e.g. $d_1 = d_2 = 1000$ on a single machine given very few entries.

Also, we verify the theoretically predicted superlinear convergence rate in our experiments accurately (cf. Fig. 2), noting that the proposed algorithm converges often to Frobenius errors as small as 10^{-10} in 10 or 20 iterations, while other iterative algorithms need hundreds or thousands for comparable errors.

E. Conclusion

We consider HM-IRLS as the presently best extension of [16], as we achieve their local convergence rate, which is unprecedented for both IRLS-type and other types of algorithms solving the Schatten-p minimization problem. Unlike for the algorithm of [16], we observe that global convergence of HM-IRLS is not lost for small $p < 0.5$, even not for p as small as $p = 0.01$. The numerical experiments indicate that this behaviour of HM-IRLS is complemented by a superior performance in terms of the number of measurements needed for successful recovery compared to state-of-the-art methods.

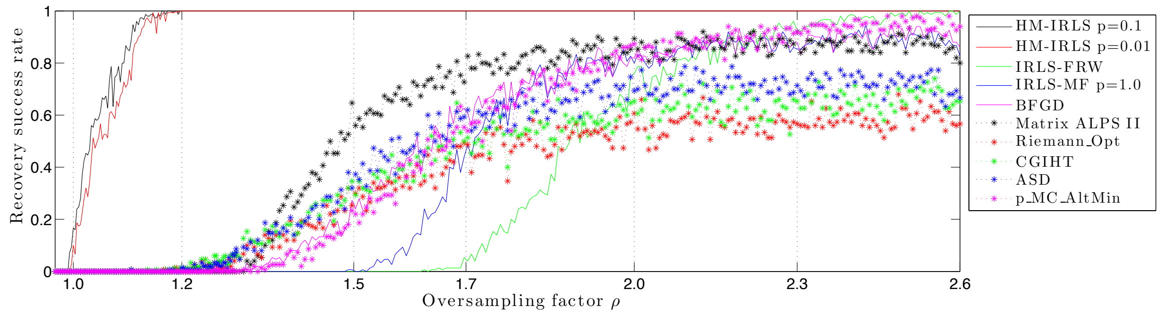
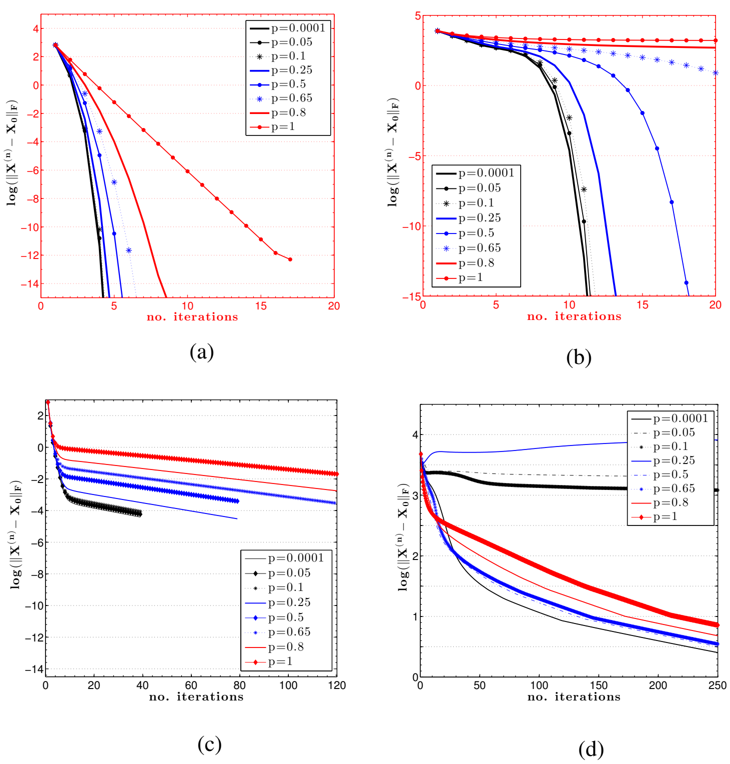


Fig. 1: Recovery success rate with varying oversampling factor ρ for state-of-the-art algorithms
We consider low-rank matrices $X_0 \in \mathbb{R}^{d_1 \times d_2}$, $\operatorname{rank}(X_0) = r$ with $d_1 = d_2 = 100$, $r = 8$, and take matrix completion measurements, sampling $m = \rho \cdot r(d_1 + d_2 - r)$ entries of X_0. We examine the recovery performance for various types of algorithms such as IRLS algorithms `IRLS-FRW`, `IRLS-MF` [9], [10], Riemannian optimization (`Riemann_Opt` [11]), alternating minimization (`p_MC_AltMin`, `ASD` and `BFGD` as in [12]–[14]), and iterative hard thresholding (`MatrixALPSII`, `CGIHT_Matrix` [6], [15]). We draw 150 instances of $X_0 = UV^*$, where $U \in \mathbb{R}^{d_1 \times r}$ and $V \in \mathbb{R}^{r \times d_2}$ are random matrices with i.i.d. standard Gaussian entries, and matrix completion sampling operators Φ with oversampling factor ρ from 0.975 to 2.60, and define successful recovery as a relative Frobenius error of smaller than 10^{-3}.



(a) (b) (c) (d)

Fig. 2: Logarithmic error plot vs. number of iterations for `HM-IRLS` [(a) and (b)] and `IRLS-FRW` [9] [(c) and (d)] and oversampling factors $\rho = 2.0$ [(a) and (c)] and $\rho = 1.2$ [(b) and (d)]
We consider low-rank matrices $X_0 \in \mathbb{R}^{d_1 \times d_2}$, $\operatorname{rank}(X_0) = r$ with $d_1 = d_2 = 40$, $r = 10$, and take matrix completion measurements, sampling $m = \rho \cdot r(d_1 + d_2 - r)$ entries of X_0. We run different variants of IRLS, namely `HM-IRLS` and `IRLS-FRW` [9], to recover X_0 for different the choices for the paramter $p = \{0.0001, 0.05, 0.1, 0.25, 0.5, 0.65, 0.8, 1.0\}$, steering the non-convexity of the minimization problem (1).

References

[1] M. Davenport and J. Romberg, "An overview of low-rank matrix recovery from incomplete observations," *IEEE J. Sel. Topics Signal Process.*, vol. 10, pp. 608–622, 2016.
[2] M. Fazel, "Matrix rank minimization with applications," Ph.D. dissertation, Stanford University, 2002.
[3] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, 2010.
[4] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proc. ACM Symp. Theory Comput. (STOC)*, 2013, pp. 665–674.
[5] Q. Zheng and J. Lafferty, "A convergent gradient descent algorithm for rank minimization and semidefinite programming from random linear measurements," in *Advances in Neural Information Processing Systems (NIPS)*, 2015, pp. 109–117.
[6] J. D. Blanchard, J. Tanner, and K. Wei, "CGIHT: Conjugate gradient iterative hard thresholding for compressed sensing and matrix completion," *Inf. Inference*, vol. 4, no. 4, pp. 289–327, 2015.
[7] S. Tu, R. Boczar, M. Simchowitz, M. Soltanolkotabi, and B. Recht, "Low-rank Solutions of Linear Matrix Equations via Procrustes Flow," 2015, preprint, arXiv:1507.03566 [math.OC].
[8] C. Kümmerle and J. Sigl, "Harmonic Mean Iteratively Reweighted Least Squares for Low-Rank Matrix Recovery," *ArXiv e-prints*, Mar. 2017, preprint, arXiv:1703.05038 [math.NA]. [Online]. Available: https://arxiv.org/abs/1703.05038
[9] M. Fornasier, H. Rauhut, and R. Ward, "Low-rank matrix recovery via iteratively reweighted least squares minimization," *SIAM J. Optim.*, vol. 21, no. 4, pp. 1614–1640, 2011.
[10] K. Mohan and M. Fazel, "Iterative reweighted algorithms for matrix rank minimization," *J. Mach. Learn. Res.*, vol. 13, pp. 3441–3473, 2012.
[11] B. Vandereycken, "Low-rank matrix completion by Riemannian optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1214–1236, 2013.
[12] J. P. Haldar and D. Hernando, "Rank-constrained solutions to linear matrix equations using powerfactorization," *IEEE Signal Process. Lett.*, vol. 16, no. 7, pp. 584–587, July 2009.
[13] J. Tanner and K. Wei, "Low rank matrix completion by alternating steepest descent methods," *Appl. Comput. Harmon. Anal.*, vol. 40, no. 2, pp. 417 – 429, 2016.
[14] D. Park, A. Kyrillidis, C. Caramanis, and S. Sanghavi, "Finding Low-rank Solutions to Matrix Problems, Efficiently and Provably," 2016, preprint, arXiv:1606.03168 [math.OC].
[15] A. Kyrillidis and V. Cevher, "Matrix recipes for hard thresholding methods," *J. Math. Imaging Vision*, vol. 48, no. 2, pp. 235–265, 2014.
[16] I. Daubechies, R. DeVore, M. Fornasier, and C. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Commun. Pure Appl. Math.*, vol. 63, pp. 1–38, 2010.