RANK-ONE AND SPARSE MATRIX DECOMPOSITION FOR DYNAMIC MRI

XIANCHAO XIU AND LINGCHEN KONG

Department of Applied Mathematics
Beijing Jiaotong University, Beijing 100044, China

Abstract. We introduce a rank-one and sparse matrix decomposition model for dynamic magnetic resonance imaging (MRI). Since $\ell_p$-norm ($0 < p < 1$) is generally nonconvex, nonsmooth, non-Lipschitz, we propose reweighted $\ell_1$-norm to surrogate $\ell_p$-norm. Based on this, we put forward a modified alternative direction method. Numerical experiments are also given to illustrate the efficiency of our algorithm.

1. Introduction. Dynamic magnetic resonance imaging (MRI) reconstructs a temporal series of images to resolve the motion or the variation of the imaged object. It is often used in cardiac, perfusion, functional, and gastro-intestinal imaging. However, achieving high spatio-temporal resolutions is challenging and inspiring in dynamic MRI due to the hardware limitations and the risk of peripheral nerve stimulation. This is because the data in MRI, samples in $k$-space of the spatial Fourier transform of the object, are acquired sequentially in time. A variety of MR techniques therefore aim to reduce the number of data required for accurate reconstruction.

Recently, R. Otazo et al. [12] applied low-rank and sparse model into MRI, and considered the following optimization problem

$$
\min \quad \text{rank}(L) + \beta ||TS||_0
$$

s.t. \quad \text{E}(L + S) = d,

where $|| \cdot ||_0$ is $\ell_0$-norm, which counts the number of nonzero entries. $\beta \in R$ is a real positive weighting parameter to balance the weights of rank and sparsity, and $d$ is the undersampled $k - t$ data. $L \in \mathbb{C}^{m \times n}$ is a low-rank matrix, which is better at capturing the global structure. $S \in \mathbb{C}^{m \times n}$ is a sparse matrix, which can capture the local structure. Generally, we assume that the dynamic component $S$ has a sparse representation in some known sparsifying transformation $T$, hence the idea of minimizing $TS$ and not $S$ itself. For acquisition with multiple receiver coils, $E$ is given by the frame-by-frame multicoil encoding operator, which performs a multiplication by coil sensitivities followed by a Fourier transform according to the sampling pattern.

2010 Mathematics Subject Classification. Primary: 90C90, 90C26; Secondary: 65K10.

Key words and phrases. Dynamic MRI, rank-one and sparse matrix decomposition, modified alternative direction method.

This research was partially supported by the National Natural Science Foundation of China (11431002, 11171018).
It will be nice if one can solve (1), however, it is NP-hard due to the combinatorial nature of the \( \ell_0 \)-norm. Instead one typically resorts to its relaxations that will be solvable in polynomial time. Fortunately, Chandrasekaran et al. [5] suggested to replace rank with nuclear norm and \( \ell_0 \) norm with \( \ell_1 \) norm, hence, (1) can be reformulated as follows

\[
\min \quad \|L\|_* + \beta \|TS\|_1 \\
\text{s.t.} \quad E(L + S) = d,
\]

where \( \| \cdot \|_* \) is the nuclear norm (defined as the sum of all singular values) of a matrix, and \( \| \cdot \|_1 \) is the \( \ell_1 \) norm in componentwise (defined as the sum of absolute values of all entries).

In fact, \( \ell_1 \) norm is a loose approximation of \( \ell_0 \) norm and often leads to an overpenalized problem, consequently, some further improvements are required. Among such efforts, a very natural improvement is the suggestion of the use of \( \ell_p \) (\( 0 < p < 1 \)) norm, which is defined as \( \|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p} \) for any \( x \in \mathbb{C}^n \). Even though the \( \ell_p \) norm is generally nonconvex, nonsmooth, non-Lipschitz, and NP-hard (see, e.g. [6, 9]), it is shown in [8, 7] that under some assumptions, the sequence generated by iterative reweighted \( \ell_1 \) or \( \ell_2 \) minimization algorithms converges to the sparest solution, which is also the global minimizer. Compared with the convex \( \ell_1 \) norm regularization, it has significant advantages in many fields such as finance, econometrics and signal processing. Sequently,

\[
\min \quad \|L\|_* + \beta \|TS\|_p \\
\text{s.t.} \quad E(L + S) = d.
\]

X. Li et al. [10] considered the following classical low-rank and sparse matrix decomposition model for surveillance video

\[
\min \quad \text{rank}(L) + \|S\|_0 \\
\text{s.t.} \quad L + S = M,
\]

and suggested to use rank-one matrix to surrogate low-rank matrix because the background component has relative small changes over a short period, that is,

\[
\min \quad \|S\|_0 \\
\text{s.t.} \quad u_1^T + S = M,
\]

where \( u \in \mathbb{R}^m, \ 1 \in \mathbb{R}^n \), and \( L = u_1^T \).

Inspired by their work, in this paper, we focus our attention on the rank-one and sparse matrix decomposition model

\[
\min \quad \|TS\|_p \\
\text{s.t.} \quad E(u_1^T + S) = d,
\]

where \( L = u_1^T, \ u \in \mathbb{C}^m, \ 1 \in \mathbb{R}^n \).

For classical low rank and sparse matrix decomposition model (2), Wright et al. [13] adopted the iterative thresholding technique to solve. However, the iterative thresholding scheme converges extremely slowly. Lin et al. [11] presented the novel augmented Lagrange multipliers, and established its convergence analysis. Building on the above work, Yuan and Yang [16] showed alternating direction method of multipliers (ADMM) to handle it. After that, ADMM began to be revisited frequently due to its success in the emerging application of structured convex optimization problems, like statistical learning [3], graphical model selection [2], matrix
completion [11]. Xiu et al. [15] proposed modified iterative reweighted algorithm (MIRL1), which worked well in surveillance video. An interesting question occurs: if we can combine ADMM with MIRL1 to deal with dynamic MRI?

We will give an affirmative answer in this paper, and the rest of this paper is organized as follows. We will analyze (3) and derive modified alternative direction method in Section 2. Section 3 presents experimental results, including comparisons with other algorithm and computational performance. Finally, Section 4 concludes our paper.

2. Model and Algorithm. We begin this section with model analysis. It is easy to write the Lagrangian function associated with (3),

$$L_\lambda(u,S,M) = \frac{1}{2} \|E(u^T + S) - d\|_2^2 + \lambda \|TS\|_p^p$$

with $\lambda \in \mathbb{R}$ being a tuning parameter.

Now, we show how to apply modified alternating direction method (MADM) to solve (4). We first divide this optimization problem into three subproblems:

$$\begin{cases}
    u^k = \text{argmin}_u \frac{1}{2} \|u^T + S^{k-1} - M^{k-1}\|_2^2, \\
    S^k = \text{argmin}_S \frac{1}{2} \|u^k^T + S - M^{k-1} + \lambda \|TS\|_p^p, \\
    M^k = u^k^T + S^k - E^*(E(u^k^T + S^k) - d),
\end{cases}$$

where $M^0 = E^*d$ and $M^k$ can be derived from the third subproblem. Here, we call these subproblems in (5) as u-subproblem, S-subproblem, M-subproblem, respectively.

- u-subproblem has solution

$$u^k = \text{mean}(M^{k-1} - S^{k-1}),$$

where for any $P \in \mathbb{C}^{m \times n}$, $\text{mean}(\cdot) \in \mathbb{C}^m$ is defined as

$$(\text{mean}(P))_i = \frac{1}{n} \sum_{j=1}^n P_{ij}, \quad i = 1, 2, \cdots, m.$$

- S-subproblem seems much more troublesome. As we know, $\|TS\|_p^p$ ($0 < p < 1$) is not Lipschitz continuous at some points, and this poses the difficulty to design algorithms for solving (4). Much of the recent research has concentrated on the reweighted $\ell_1$ minimization problems [1, 4, 17], and shown that it performs well in recovering sparse signal. Thus, in this paper, for a fixed $u^k$, the reweighted $\ell_1$ version for S-subproblem is revised to

$$\min \frac{1}{2} \|E(u^k^T + S) - d\|_2^2 + \lambda \|W \circ (TS)\|_1,$$

where $\circ$ denotes the Hadamard product, $W$ is the reweighted matrix and $0 < W_{ij} \leq 1$ for entries in $W$.

Following the similar idea of [15], we can derive the following modified iterative reweighted $\ell_1$ algorithm (MIRL1) for the S-subproblem (6).

By exploiting the method from Literature [18] to update the weight, which is different from classic iterative reweighted $\ell_1$ algorithm [4], we take $W^\tau$ as

$$W^\tau_{ij} = \begin{cases}
    \left[ \frac{\max_{(i,j) \in (\Omega^\tau)^C} \|H_{ij}\|_2^2 + \varepsilon}{\max_{(i,j) \in (\Omega^\tau)^C} \|H_{ij}\|_2^2} \right]^{p-1}, & (i,j) \in \Omega^\tau; \\
    1, & (i,j) \in (\Omega^\tau)^C,
\end{cases}$$

where $\Omega^\tau$ is the set of indexes with weights in the previous iteration.
Algorithm 1: Modified Iterative Reweighted $\ell_1$ Algorithm for (6)

**Initialize:** $\lambda, m, \varepsilon_0 \geq 0, S^0 = M^{k-1} - u^k 1T$, and $W^0$;

For $\tau = 1 : m$

Initialize $S^{\tau,1} = S^{\tau-1}$;

While $\|S^{\tau,k+1} - S^{\tau,k}\|_2 > \varepsilon_0$

Compute $S^{\tau,k+1} = T^{-1}(\max \{ S^{\tau,k} - \lambda W^{\tau}, 0 \} \frac{S^{\tau,k}}{|S^{\tau,k}|})$;

End

Update $S^{\tau} = S^{\tau,k+1}$;

Update $W^{\tau+1}$ from $S^{\tau-1}, S^{\tau}$;

End

where

$$H^{\tau} = (S^T)^{\tau} - (S^T)^{\tau-1}, \quad k^{\tau} = |\text{supp}(S^{\tau})|,$$

$\Omega^{\tau}$ denotes the indices of the $k^{\tau}$ largest entries (in absolute) of $H^{\tau}$ and $(\Omega^{\tau})^C$ notates its complementary set, and $0 < p < 1, \varepsilon > 0$ is sufficiently small.

- M-subproblem ensures data consistency, which is

$$M^k = u^k 1T + S^k - E^*(E(u^k 1T + S^k) - d).$$

Based on the above arguments, the framework of modified alternative direction method (MADM) for (4) is obtained as below.

Algorithm 2: Modified Alternative Direction Method for (3)

**Initialize:** $\lambda, m, \varepsilon_0 \geq 0, M^0 = E^*d, S^0 = 0$, and $W^0$;

**While** not converged **do**

Compute $u$-subproblem by

$u^k = \text{mean}(M^{k-1} - S^{k-1})$;

Compute $S$-subproblem by Algorithm 1;

Compute $M$-subproblem by

$$M^k = u^k 1T + S^k - E^*(E(u^k 1T + S^k) - d);$$

End while

**Output:** $S$

Unfortunately, the convergence of Algorithm 2 in unknown, because it is different from the classical ADMM. Here, we only give the convergence of Algorithm 1.

**Theorem 2.1.** For given $\lambda > 0$, and $W^\tau$, the inter loops of Algorithm 1 have the global convergence, namely, any accumulation point of $\{S^{\tau,k}\}$ is a global minimizer of $\min \frac{1}{2}\|E(u^k 1T + S) - d\|_2^2 + \lambda\|W^{\tau} \circ (TS)\|_1$.

**Proof.** See [15] for more details.

3. Numerical Experiments. In this section, we conduct numerical experiments to compare the performance of the MADM and IST in [12]. All experiments are performed in MATLAB version 8.2 with Intel Core I5 2.60GHz CPU and 8GB of RAM. We test cardiac perfusion data, which are available at the web site cair2r.net/resources/software/ls-reconstruction-matlab-code. Some example frames of the test data can be found in Figure 1.
In this paper, we choose parameter $\lambda = 0.01$ as suggested in [16], stopping criterion $\varepsilon_0 = 5\varepsilon - 4$. The following Figures 1–4 show the computational results. Figure 1 shows the computational results of cardiac perfusion by IST, Figures 2–4 show the computational results by MADM with $p = 0.1, 0.5, 0.9$, respectively. In these figures, the first row is original frames, the second row is reconstructed background parts, and the third is reconstructed moving parts.

![Some frames of cardiac perfusion by IST.](image)

From the numerical experiments, we see that: 1) MADM is easier to find the pathological parts than IST. See the third row of Figures 1–4; 2) The smaller $p$ we choose, the more obvious interested parts we get. See the third row of Figures 2–4.

4. Conclusions. In this paper, we introduced a new rank-one and sparse matrix decomposition for dynamic MRI. Rather than handling the nonconvex $\ell_p$-norm directly, we applied the convex reweighted $\ell_1$-norm. Combining MIRLI and ADM, we developed MADM for (4), and illustrated the effectiveness of the proposed algorithm.

Acknowledgement. We would like to thank the referees very much for their valuable comments and suggestions.
Figure 2. Some frames of cardiac perfusion by MADM with $p=0.1$.

Figure 3. Some frames of cardiac perfusion by MADM with $p=0.5$. 
REFERENCES

[1] K. Amin, W. Xu, A. Avestimehr and B. Hassibi, Weighted \( \ell_1 \) minimization for sparse recovery with prior information, *IEEE International Symposium on Information Theory*, 2 (2009), 483–487.

[2] O. Banerjee, L. Ghaoui and A. D’Aspremont, Model selection through sparse maximum likelihood estimation for multivariate Gaussian or Binary data, *The Journal of Machine Learning Research*, 9 (2008), 485–516.

[3] S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein, Distributed optimization and statistical learning via alternating direction method of multipliers, *Foundations and Trends in Machine Learning*, 3 (2011), 1–122.

[4] E. Candès, M. Wakin and S. Boyd, Enhancing sparsity by reweighted \( \ell_1 \) minimization, *Journal of Fourier Analysis and Applications*, 14 (2008), 877–905.

[5] V. Chandrasekaran, S. Sanghavi, P. Parrilo and A. Willsky, Rank-sparsity incoherence for matrix decomposition, *SIAM Journal on Optimization*, 21 (2011), 572–596.

[6] X. Chen, D. Ge, Z. Wang and Y. Ye, Complexity of unconstrained \( L_2-L_p \) minimization, *Mathematical Programming*, 143 (2014), 371–383.

[7] I. Daubechies, R. DeVore, M. Fornasier and C. Giuntírk, Iteratively reweighted least squares minimization for sparse recovery, *Communications on Pure and Applied Mathematics*, 63 (2010), 1–38.

[8] S. Foucart and M. Lai, Sparest solutions of underdetermined linear systems via \( \ell_p \) minimization for \( 0 < q < 1 \), *Applied and Computational Harmonic Analysis*, 26 (2009), 395–407.

[9] D. Ge, X. Jiang and Y. Ye, A note on the complexity of \( \ell_p \) minimization, *Mathematical Programming*, 129 (2011), 285–299.

[10] X. Li, M. Ng and X. Yuan, Nuclear-norm-free variational models for background extraction from surveillance video, submitted to IEEE Transactions on Image Processing, 2013.

[11] Z. Lin, M. Chen and Y. Ma, The augmented Lagrange multiplier method for exact recovery of a corrupted low-rank matrices, Preprint, 2010.
[12] R. Otazo, E. Candès and D. Sodickson, Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components, *Magnetic Resonance in Medicine*, **73** (2015), 1125–1136.

[13] J. Wright, A. Ganesh, S. Rao, Y. Peng and Y. Ma, Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization, *Advances in Neural Information Processing Systems*, (2009), 2080–2088.

[14] S. Wright, R. Nowak and M. Figueiredo, Sparse reconstruction by separable approximation, *IEEE Transactions on Signal Processing*, **57** (2009), 2479–2493.

[15] X. Xiu, L. Kong and S. Zhou, Modified iterative reweighted $\ell_1$ algorithm for surveillance video, Preprint, 2014.

[16] X. Yuan and J. Yang, Sparse and low-rank matrix decomposition via alternating direction methods, *Pacific Journal of Optimization*, **9** (2013), 167–180.

[17] Y. Zhao and D. Li, Reweighted $\ell_1$-minimization for sparse solutions to underdetermined linear systems, *SIAM Journal on Optimization*, **22** (2012), 1065–1088.

[18] S. Zhou, N. Xiu, Y. Wang and L. Kong, Exact recovery for sparse signal via weighted $\ell_1$ minimization, Preprint, 2014.

Received November 2014; 1st revision January 2015; final revision April 2015.

E-mail address: xianchaoxiu@163.com
E-mail address: konglchen@126.com