Iteratively reweighted penalty alternating minimization methods with continuation for image deblurring

Tao Sun, Dongsheng Li, Hao Jiang, Zhe Quan

NUDT& HNU

We consider a class of nonconvex problems

$$\min_{x,y}\{\Psi(x,y) := f(x) + \sum_{i=1}^{N} h(g(y_i)),\ \text{s.t. } Ax + By = c\}. \qquad (1)$$

where $x \in \mathbb{R}^M, y \in \mathbb{R}^N$, and functions f, g and h satisfy the following assumptions:

- **A.1** $f : \mathbb{R}^M \to \mathbb{R}$ is a closed proper convex function and $\inf_{x\in\mathbb{R}^M} f(x) > -\infty$.
- **A.2** $g : \mathbb{R} \to \mathbb{R}$ is a convex function, and the proximal map of g is easy to calculated.
- **A.3** $h : \mathrm{Im}(g) \to \mathbb{R}$ is a concave function and $\inf_{t\in\mathrm{Im}(g)} h(t) > -\infty$.

This problem is frequently used in deblurring task.

Although ADMM can be applied to this nonconvex problem, several drawbacks still exist.

1. The convergence guarantees of nonconvex ADMMs require a very large Lagrange dual multiplier. Worse still, the large multiplier makes the nonconvex ADMM run slowly.

2. When applying nonconvex ADMMs to the nonconvex TV deblurring model, by direct checks, the convergence requires TV operator to be full row-rank; however, the TV operator cannot promise such an assumption.

3. The previous analyses show that the sequence converges to a critical point of an auxiliary function under several assumptions. But the relationship between the auxiliary function and the original one is unclear in the nonconvex settings.

We consider the penalty function as

$$\min_{x,y}\{\Phi_\gamma(x,y) := f(x) + \sum_{i=1}^{N} h(g(y_i)) + \frac{\gamma}{2}\|Ax + By - c\|_2^2\}. \quad (2)$$

The difference between problem (1) and (2) is determined by the parameter γ. They are identical if $\gamma = +\infty$.

The classical algorithm solving this problem is the Alternating Minimization (AM) method, i.e., minimizing one variable while fixing the other one.

However, if directly applying AM to model (2), the subproblem may still be nonconvex; the minimizer is hard to obtain in most cases.

Considering the structure of the problem, we use a linearized technique for the nonsmooth part $\sum_{i=1}^{N} h(g(y_i))$. This method was inspired by the reweighted algorithms. To derive the sufficient descent, we also add a proximal term. And we apply the continuation technique to the penalty parameter.

Scheme: parameters $\bar{\gamma} > 0, a > 1, \delta > 0$

Initialization: $z^0 = (x^0, y^0)$, $\gamma_0 > 0$

for $k = 0, 1, 2, \ldots$

$x^{k+1} \in \arg\min_x \{f(x) + \frac{\gamma_k}{2}\|Ax + By^k - c\|_2^2\}$

$w_i^k \in -\partial(-h(g(y_i^k)))$, $i \in [1, 2, \ldots, N]$

$y^{k+1} \in$
$\arg\min_y \{\sum_i^N w_i^k g(y_i) + \frac{\gamma_k}{2}\|Ax^{k+1} + By - c\|_2^2 + \frac{\delta\gamma_k\|y-y^k\|_2^2}{2}\}$

$\gamma_{k+1} = \min\{\bar{\gamma}, (a\gamma_k)\}$

end for

Output x^k

- **A.4** $f(x) + \frac{1}{2}\|Ax\|_2^2$ is strongly convex with ν.

Convergence: Assume that $(z^k)_{k\geq 0}$ is generated by IRPAMC and Assumptions **A.1**, **A.2**, **A.3** and **A.4** hold, and $\delta > 0$. Then we have the following results.

(1) It holds that

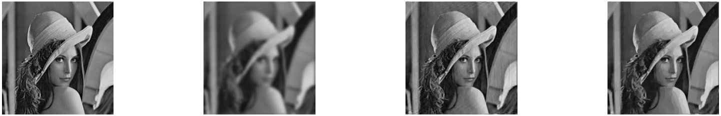
$$\begin{aligned}&\Phi_{\bar{\gamma}}(x^k, y^k) - \Phi_{\bar{\gamma}}(x^{k+1}, y^{k+1}) \\ &\quad \geq \min\{\bar{\gamma}, \nu\bar{\gamma}\} \cdot \|x^{k+1} - x^k\|_2^2 + \frac{\delta\bar{\gamma}\|y^{k+1} - y^k\|_2^2}{2}.\end{aligned}$$

for $k > K$ with $K = \ulcorner \log_a(\frac{\bar{\gamma}}{\gamma_0}) \urcorner$.

(2) $\sum_k (\|x^{k+1} - x^k\|_2^2 + \|y^{k+1} - y^k\|_2^2) < +\infty$, which implies that

$$\lim_k \|x^{k+1} - x^k\|_2 = 0, \ \ \lim_k \|y^{k+1} - y^k\|_2 = 0.$$

we apply the proposed algorithm to image deblurring and compare the performance with the nonconvex ADMM. The Lena image is used in the numerical experiments.



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

Figure: Deblurring results for Lena under Gaussian operator by using the two algorithms. (a) Original image; (b) Blurred image; (c) IRPAMC 16.0dB; (d) nonconvex ADMM 14.4dB.