Power-Iterative Strategy for $\ell_p$-$\ell_2$ Optimization for Compressive Sensing: Towards Global Solution

Jie Yan and Wu-Sheng Lu

Department of Electrical and Computer Engineering
University of Victoria, Victoria, BC, Canada

November 7, 2011
INTRODUCTION

OUTLINE

1. INTRODUCTION
2. POWER-Iterative Algorithms
3. SIMULATIONS
4. CONCLUSIONS
5. KEY REFERENCES
Background

Compressive sensing: seeking an accurate or approximate solution to an underdetermined linear system, while requiring the solution to have fewest nonzero components.

The recovery problem can be solved in an $\ell_1$-$\ell_2$ formulation as

$$\min \lambda \| s \|_1 + \| \Theta s - y \|_2^2$$
Fast iterative-shrinkage-thresholding algorithm (FISTA)

FISTA provides a convergence rate of $O(1/k^2)$, compared to $O(1/k)$ by ISTA

MFISTA: an enhanced version of FISTA with monotone convergence
Background Cont’d

- \( \ell_p \) minimization with \( 0 < p < 1 \) outperforms \( \ell_1 \) minimization

- We propose algorithms for

\[
\min F(s) = \lambda \|s\|_p^p + \|\Theta s - y\|_2^2
\]

with \( 0 < p < 1 \).
New ingredients of our algorithms

1. We associate (1) with an $\ell_p - \ell_2$ P-P objective function, and devise a fast formulation to secure the its global minimizer

2. A parallel implementation with significantly accelerated computational speed

3. An algorithm developed in the framework of MFISTA (denoted as M-MFISTA)

4. A power-iterative strategy designed towards a global solution of (1)
Within the framework MFISTA, we associate each iteration to a P-P function

\[ Q_p(s, b_k) = f(b_k) + \langle s - b_k, \nabla f(b_k) \rangle + \frac{L}{2} \| s - b_k \|^2 + \lambda \| s \|^p \]

where \( f(s) = \| \Theta s - y \|_2^2 \).

Up to a constant, minimizing \( Q_p(s, b_k) \) can be cast as

\[ \min \hat{Q}_p(s, b_k) = \frac{L}{2} \| s - c_k \|^2 + \lambda \| s \|^p \quad (2) \]

with \( c_k = b_k - \frac{1}{L} \nabla f(b_k) \).

As \( p < 1 \), conventional soft shrinkage fails to work in general.
Global solver for $\hat{Q}_p(s, b_k) = \frac{L}{2}||s - c_k||^2_2 + \lambda||s||^p_p$

- The problem is reduced to a series of $N$ 1-D problems of the form

  $$\min u(s) = \frac{L}{2}(s - c)^2 + \lambda|s|^p.$$  \hspace{1cm} (3)

- By our proposed method in [1], global solution of (3) is obtained as

  $$z = \text{gsol}(c, L, \lambda, p) \quad \text{when } c \geq 0$$

  $$z = -\text{gsol}(-c, L, \lambda, p) \quad \text{when } c < 0$$

- Drawback: Low efficiency, as one needs to solve $N$ 1-D problems
Parallel implementation for \( \min \hat{Q}_p(s, b_k) = \frac{L}{2} \| s - c_k \|_2^2 + \lambda \| s \|_p^p \)

- Idea: process the data in a vector-wise rather than component-wise manner
- The proposed \( \ell_p-\ell_2 \) solver is highly parallel with only \(|\Omega|\) calls for 1-D solver \( g_{sol} \).
- \(|\Omega|\) is typically much smaller than \( N \)
- Complexity is considerably reduced compared with that required by applying 1-D solver \( g_{sol} \) \( N \) times.
### Power-Iterative Strategy for $\ell_p$-$\ell_2$ Optimization for Compressive Sensing: Towards Global Solution

**Power-Iterative Algorithms**

<table>
<thead>
<tr>
<th>Input Data</th>
<th>$c_k, L, \lambda$ and $p$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output Data</td>
<td>$z_k = \arg\min \hat{Q}_p(s, b_k)$.</td>
</tr>
<tr>
<td><strong>Step 1</strong></td>
<td>Set $\theta = \text{sign}(c_k)$ and $c = \theta \cdot c_k$.</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td>If $p = 0$, compute $\vartheta = \left[ \frac{L}{2} c^2 &gt; (\lambda \cdot 1) \right]$, set $z = c \cdot \vartheta$ and do Step 4; otherwise do Step 3.</td>
</tr>
</tbody>
</table>
| **Step 3** | 1. Compute $s_c = \left[ \lambda p (1 - p)/L \right]^{1/(2-p)}$, set $\vartheta = \left[ (s_c \cdot 1) < c \right]$ and $z = \vartheta$.  
2. Define $\Lambda = \{i : \vartheta(i) = 1\}$ and update $c \leftarrow c(\Lambda)$.  
3. Compute $v = L(s_c \cdot 1 - c) + \lambda p s_c^{p-1} \cdot 1$, update $\vartheta = [v \geq 0]$ and set $\tilde{s} = s_c \cdot \vartheta$.  
4. Define $\Omega = \{i : \vartheta(i) = 0\}$. For each $i \in \Omega$, replace the $i$th component of $\tilde{s}$ by the global solution of (3) over $[s_c, c]$ with $c = c(i)$.  
5. Set $\vartheta = \left[ \frac{L}{2} c^2 > \frac{L}{2} (\tilde{s} - c)^2 + \lambda \tilde{s}^p \right]$ and $\tilde{z} = \tilde{s} \cdot \vartheta$.  
6. update $z(\Lambda) = \tilde{z}$. |
| **Step 4** | $z_k = \theta \cdot z$. |
Modified MFISTA (M-MFISTA)

- By replacing the conventional soft shrinkage with the global-parallel $\ell_p-\ell_2$ solver in MFISTA, we construct M-MFISTA.

- In each iteration the M-MFISTA minimizes the P-P function globally.

- Yet, a solution of problem (1) is not guaranteed globally optimal.

- We propose a power-iterative strategy towards global solution.
### POWER-ITERATIVE ALGORITHMS

**Input Data**
\(\lambda, p, \Theta\) and \(y\).

<table>
<thead>
<tr>
<th>Output Data</th>
<th>Local solution of problem (1).</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Take (L = 2\lambda_{\text{max}}(\Theta\Theta^T)) as the Lipschitz constant of (\nabla f). Set initial iterate (s_0) and the number of iterations (K_m). Set (b_1 = s_0), (k = 1) and (t_1 = 1).</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Step 2</th>
<th>Compute minimizer (z_k) of (2) using the parallel global solver. Then update</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(t_{k+1} = (1 + \sqrt{1 + 4t_k^2})/2), (s_k = \text{argmin} {F(s) : s = z_k, s_{k-1}}), (b_{k+1} = s_k + (t_k/t_{k+1})(z_k - s_k) + [(t_k - 1)/t_{k+1}](s_k - s_{k-1})).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 3</th>
<th>If (k = K_m), stop and output (s_k) as the solution; otherwise set (k = k + 1) and repeat from Step 2.</th>
</tr>
</thead>
</table>
The Power-Iterative Strategy

1. Obtain a global solution $s^{(0)}$ by solving the convex $\ell_1$-$\ell_2$ problem.
2. $s^{(0)}$ is then used as the initial point to start the next $\ell_p$-$\ell_2$ problem with a $p$ close to but slightly less than one. Use M-MFISTA to obtain a solution $s^{(1)}$.
3. $s^{(1)}$ is served as an initial point for the next $\ell_p$-$\ell_2$ problem with $p$ further reduced slightly.
4. This process continues until the target power value $p_t$ is reached.
The global solutions of (1) associated with powers $p$ and $p + \Delta p$ should be close to each other with $\Delta p$ sufficiently small.

Based on this intuitive observation, the above power iterative technique is developed to produce solutions likely globally optimal.
OUTLINE

1 INTRODUCTION

2 POWER-ITERATIVE ALGORITHMS

3 SIMULATIONS

4 CONCLUSIONS

5 KEY REFERENCES
Parameter Settings

- Generate $K$-sparse test signal with $N = 32$.
- A total of 20 values of $K$ from 1 to 20 were used.
- Set the number of measurements to $M = 20$. Construct a Gaussian measurement matrix $\Phi$ of size $M \times N$.
- The power-iterative strategy in conjunction with M-MFISTA was applied to reconstruct $s$. 
A sequence of power $p$ was set from 1 to 0 with a decrement of $d = 0.1$.

A recovered signal $\hat{s}$ was deemed perfect if the relative solution error $\|\hat{s} - s\|_2 / \|s\|_2$ was less than $1e-5$.

For each value of $K$, the number of perfect reconstructions were counted over 100 runs.
Fig 1: Rate of perfect reconstruction for $\ell_p$-$\ell_2$ problems with $p = 1, 0.9, 0.8, 0.7, 0.4$ and $0$ over 100 runs for signals of length $N = 32$ and number of random measurements $M = 20$. 
Fig 2: Average relative reconstruction errors for $\ell_p$-$\ell_2$ problems with $p = 1, 0.9, 0.8, 0.7, 0.4$ and $0$ over 100 runs for signals of length $N = 32$ and number of random measurements $M = 20$. 
Observations

1. Perfect reconstruction rate increases and the average relative reconstruction error reduces as a smaller power $p$ was used.

2. Considerable improvement as $p$ reduces from 1 to 0.9.

3. The best reconstruction performance was achieved at $p = 0$. 
SIMULATIONS

Power-Iterative Strategy for $\ell_p-\ell_2$ Optimization for Compressive Sensing: Towards Global Solution

---

The graphs illustrate the perfect reconstruction rate as a function of sparsity, $K$. The plots compare different optimization strategies for $\ell_p-\ell_2$ minimization, with $p=1$, $p=0$, and $p=0.9$. The results are shown for two methods: (LS) and (ZERO). The graphs demonstrate how the reconstruction rate varies with increasing sparsity for each method, indicating the effectiveness of these strategies in recovering sparse signals from compressively sensed data.
Importance of Initial Point

- Compare the $\ell_0$ (and $\ell_{0.9}$) solution obtained by the power-iterative strategy with an $\ell_0$ (and $\ell_{0.9}$) solution obtained by M-MFISTA with the least-squares solution or the zero vector as the initial point.
- Considerable performance gain achieved by the proposed method with an adequate initial point.
- Choosing an initial point greatly affects reconstruction performance.
CONCLUSIONS

OUTLINE

1 INTRODUCTION
2 POWER-ITERATIVE ALGORITHMS
3 SIMULATIONS
4 CONCLUSIONS
5 KEY REFERENCES
A parallel global solver devised for the $\ell_p-\ell_2$ P-P function

2. A modified MFISTA (M-MFISTA) developed for local solution of the $\ell_p-\ell_2$ problem

3. A power-iterative strategy towards global solution of the $\ell_p-\ell_2$ problem

4. Experimental results for CS signal recovery are presented to show the superiority of the proposed algorithms compared with the conventional BP benchmark, and to demonstrate that the solutions obtained are highly likely to be globally optimal.
### KEY REFERENCES


