Accelerated Iterative Thresholding Algorithm for Regularization of Linear Inverse Problems with Sparsity Constraints

Presenter: Pham, Quy Muoi

Seminar, 10.04.2010
Outline

1. Motivation
2. Notations
3. Accelerated iterative thresholding algorithm
4. The convergence of the accelerated algorithm
5. Numerical Experiments
The aim is to compute approximate solution of an operator equation

\[ T(u) = y, \quad (1) \]

where \( T : X \rightarrow Y \) is a linear and ill-posed operator between Hilbert spaces \( X \) and \( Y \). Moreover, only noisy data \( y^\delta \) with

\[ \| y^\delta - y \| \leq \delta \]

are available.

We assume that the solution of the problem has a \emph{sparse expansion} in some orthonormal basis \( \{ \varphi_k \}_{k \in \mathbb{N}} \).
Regularization with sparsity constraints

Regularization for the problem with sparsity constraints leads to minimize the functional

$$\min_{u \in X} F(u) := \| T(u) - y^\delta \|^2 + \alpha \sum_{k \in \mathbb{N}} \omega_k | \langle u, \varphi_k \rangle |^p. \quad (2)$$

where $\omega_k > \omega_0 > 0$, $\forall k \in \mathbb{N}$, and $1 \leq p \leq 2$. 
To solve problem (2) the iterative thresholding algorithm proposed in [2] as follows

\[ u^{n+1} = S_{\beta \alpha w, p}(u^n + \beta T^*(y^\delta - Tu^n)), \]

starting from an arbitrary \( u^0 \), where \( S \) is the soft thresholding operation, see (5).

We now modify this method to obtain an accelerated algorithm as follows

\[ u^{n+1} = S_{\beta \alpha w, p}(u^n + \beta^n T^*(y^\delta - Tu^n)), \quad (3) \]

starting from an arbitrary \( u^0 \).
Notations

- shrinkage functions: \( S_{\tau,p} : \mathbb{R} \to \mathbb{R} : \)

\[
S_{\tau,p}(x) = \begin{cases} 
\text{sgn}(x) \max(|x| - \frac{\tau}{2}, 0) & : p = 1 \\
G_{\tau,p}^{-1}(x) & : p \in (1, 2]
\end{cases}
\]

where

\[
G_{\tau,p}(x) = x + \frac{\tau p}{2} \text{sgn}(x)|x|^{p-1} \text{ for } 1 < p \leq 2,
\]

- Denote \( \omega = \{\omega_k\}_{k \in \mathbb{N}} \), the soft thresholding operation is defined by

\[
S_{\omega,p}(u) = \sum_{k \in \mathbb{N}} S_{\omega_k,p}(\langle u, \varphi_k \rangle) \varphi_k,
\]

\[
(4) \\
(5)
\]
Accelerated iterative thresholding algorithm

- As we introduce above, we aim at solving approximately the problem

\[
\min_{u \in X} F(u) := \| T(u) - y^\delta \|^2 + \alpha \sum_{k \in \mathbb{N}} \omega_k \langle u, \varphi_k \rangle^p.
\]

where \( \omega_k > \omega_0 > 0, \forall k \in \mathbb{N}, \alpha > 0 \) and \( 1 \leq p \leq 2 \).

- The main idea of accelerated method: we replace \( F(u) \) by a sequence functions \( F_{\beta_n}(u, u_n) \) and consider the sequence problems as follows

\[
u^0 : \text{arbitrary} ; u^{n+1} = \min_{u \in X} F_{\beta_n}(u, u_n).
\]  \hspace{1cm} (6)

- What is \( F_{\beta_n} \)? how to choose \( \beta^n \) and solve the minimizer of \( F_{\beta_n}(., u^n) \), which the sequence \( u^n \) converges fastly toward the minimizer of \( F \).
How to choose $F_{\beta_n}$ and $\beta^n$

- $F_{\beta_n}$ are defined by

$$F_{\beta}(f, a) = F(f) - \| T(f - a) \|^2 + \frac{1}{\beta} \| f - a \|^2. \quad (7)$$

- The parameters $\beta^n$ are chosen as follows

**Definition**

We say that the sequence $\{\beta^n\}_{n \in \mathbb{N}}$ satisfy Condition (B) with respect to the sequence $\{u^n\}_{n \in \mathbb{N}}$ if there exists $n_0$ so that:

- $(B_1)$ $\bar{\beta} := \sup\{\beta^n; n \in \mathbb{N}\} < \infty$ and $\inf\{\beta^n; n \in \mathbb{N}\} \geq 1$

- $(B_2)$ $\beta^n \| T(u^{n+1} - u^n) \|^2 \leq r \| u^{n+1} - u^n \|^2$

- $(B_3)$ $\beta_n \cdot r \leq 1.$

where $r = \| T \|^2 < 1.$
Some remarks

- Note that
  \[
  \| T(u^{n+1} - u^n) \|^2 \leq r \| u^{n+1} - u^n \|^2,
  \]
  and by condition \((B_1)\) we ensure
  \[
  \| T(u^{n+1} - u^n) \|^2 \leq \beta^n \| T(u^{n+1} - u^n) \|^2.
  \]

- The idea of adding condition \((B_2)\) is to find the largest number \(\beta^n \geq 1\) such that
  \[
  0 \leq -\| T(u^{n+1} - u^n) \|^2 + \frac{r}{\beta^n} \| u^{n+1} - u^n \|^2
  \]
  is as small as possible. The reason can be verified below in the definition of the surrogate functional \(\Phi_\beta\) in Lemma 2. The goal is to ensure that \(\Phi^n_\beta\) is not too far off \(F(u^n)\).

- Condition \((B_3)\) use to prove strong convergence of \(\{u^n\}_{n \in \mathbb{N}}\).
The minimizer of $F_\beta$

**Lemma**

Assume $\|T\| < 1$ and $\beta \geq 1$. For arbitrary fixed $a \in X$, define the functional $F_\beta(., a)$ is defined by (7). Then $F_\beta(f, a)$ has a unique minimizer in $X$. This minimizer is given by

$$f = \mathcal{S}_{\beta_\omega,p}(a + \beta(T^*(y - Ta))).$$

- By Lemma 2, the sequence $\{u^n\}$ defined by (6) is given by

$$u^{n+1} = \mathcal{S}_{\beta_\omega,p}(u^n + \beta^n T^*(y^\delta - Tu^n)),$$

starting from an arbitrary $u^0$. 
**Optimality condition**

**Lemma**

The functional $F$ defined by (2) has at least one minimizer $\bar{u} \in X$. Its minimizers are characterized by

$$\bar{u} = S_{\beta \alpha \omega, p}(\bar{u} + \beta T^*(y - T\bar{u}))$$

for any $\beta > 0$. (9)

**Lemma**

Assume that $\|T\| < 1$ and $u^{n+1}$ is given by

$$u^{n+1} = S_{\beta \alpha \omega, p}(u^n + \beta^n T^*(y - Tu^n)),$$

where the $\beta^n$ satisfy Condition (B) with respect to $\{u^n\}_{n \in \mathbb{N}}$; then the sequence $F(u^n)$ is monotonically decreasing, $\lim_{n \to \infty} \|u^{n+1} - u^n\| = 0$ and the $\|u^n\|$ are bounded uniformly in $n$. 
Weak and strong convergence

Because the sequence $\{u^n\}_{n \in \mathbb{N}}$ are bounded, it must have weak accumulation points.

**Lemma**

If $u^\star$ is a weak accumulation point of $\{u^n\}_{n \in \mathbb{N}}$, then $u^\star$ is a minimizer of $F$.

**Lemma**

With assumption as in Lemma 4, the sequence $\{u^n\}_{n \in \mathbb{N}}$ has a convergence subsequence in norm and the limit of the subsequence is a minimizer of $F$. 
Some algorithmic aspects

- Conditions \((B_1)\) and \((B_2)\) are inspired by a classical length step in the steepest descent algorithm for unconstrained functional \(\| Tu - y \|^2 \) leading to an accelerated Landweber iteration \( u^{n+1} = u^n + \gamma^n T^* (y - Tu^n) \), for which \( \gamma^n \) is picked

\[
\gamma^n = \| T^* (y - Tu^n) \|^2 / \| TT^* (y - Tu^n) \|^2 =: \| r^n \|^2 / \| Tr^n \|^2 \tag{10}
\]

with \( r^n = T^* (y - Tu^n) \).

- On the other hand, by Condition \((B_3)\) we have \( \beta^n \leq \frac{1}{r} \). Therefore, an explicit (but somewhat “greedy”) guess for \( \beta^n \) is given by

\[
\beta^n = \min \{ \gamma^n; \frac{1}{r} \} \tag{11}
\]

- If this choice fulfills \((B_2)\) and \((B_3)\) as well, it is retained; if it does not, it can be gradually decreased (by multiplying it with a factor slightly smaller than 1 until \((B_2)\) and \((B_3)\) are satisfied).
### Accelerated iterative thresholding algorithm
#### Algorithm 1.

<table>
<thead>
<tr>
<th>Given</th>
<th>operator $T$, data $y$, some initial guess $u^0$, $\alpha$ and $\omega$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization</td>
<td>$r = | T |^2$, set $q = 0.9$ (as an example)</td>
</tr>
<tr>
<td>Iteration</td>
<td>for $n = 0, 1, 2, \ldots$ until a preassigned precision/maximum number of the iteration</td>
</tr>
<tr>
<td></td>
<td>1. $\beta^n = \min \left{ \frac{| r^n |_2^2}{| T r^n |_2^2}, \frac{1}{r} \right}$ with $r^n = T^*(y - Tu^n)$</td>
</tr>
<tr>
<td></td>
<td>2. $u^{n+1} = S_{\beta^n \alpha \omega, \rho}(u^n + \beta^n T^*(y - Tu^n))$</td>
</tr>
<tr>
<td></td>
<td>3. verify $(B_2)$: $\beta^n | T(u^{n+1} - u^n) |_2^2 \leq r | u^{n+1} - u^n |_2^2$</td>
</tr>
<tr>
<td></td>
<td>if $(B_2)$ is satisfied increase $n$ and go to 1.</td>
</tr>
<tr>
<td></td>
<td>otherwise set $\beta^n = q.\beta^n$ and go to 2.</td>
</tr>
</tbody>
</table>
We aim at the reconstruction of a function \( x \in L^2([0, 1]) \) from the integral equation of the first kind

\[
Tu(t) := \int_0^t u(s) ds = y(t), \quad t \in [0, 1]
\]

The data \( f \) is given as \( (f(t_k))_{k=1,\ldots,N} \) with \( t_k = \frac{1}{N} k \). We discretized the operator \( K \) by the matrix

\[
T = \frac{1}{N} \begin{pmatrix}
1 & 0 & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
1 & \ldots & \ldots & 1
\end{pmatrix}, \quad K : \mathbb{R}^N \to \mathbb{R}^N.
\]
Setting

- we set $\alpha = 5 \times 10^{-4}$, $w_i = 1$, $\forall i$, $p = 1$ and $N = 500$.
- The true solution $\bar{u}$ is given by

$$
\bar{u}(s) = \begin{cases} 
30, & s \in [0.2, 0.25] \\
-60, & s \in [0.5, 0.55] \\
80, & s \in [0.65, 0.7] \\
-10, & s \in [0.85, 0.9] \\
0, & \text{the otherwise}
\end{cases}
$$

- the data $y^\delta = T\bar{u} + \delta$ is a noisy function. Figure 1 shows the true solution and our sample data $y^\delta$ with $\delta = 0.05 \cdot \text{randn}(N, 1)$. 
The true solution and noisy data
The values of $\beta^n$
Decreasing speed of $F$
We use the relative error $\|u^n - \bar{u}\|/\|\bar{u}\|$ for stopping criterion.

<table>
<thead>
<tr>
<th>Relative error</th>
<th>Algorithm 1</th>
<th>Iterative thresholding al.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>time</td>
</tr>
<tr>
<td>0.9</td>
<td>4</td>
<td>0.02</td>
</tr>
<tr>
<td>0.8</td>
<td>10</td>
<td>0.04</td>
</tr>
<tr>
<td>0.7</td>
<td>18</td>
<td>0.07</td>
</tr>
<tr>
<td>0.5</td>
<td>47</td>
<td>0.16</td>
</tr>
<tr>
<td>0.2</td>
<td>797</td>
<td>3</td>
</tr>
</tbody>
</table>

**Table:** Table illustrating the relative performance of two algorithms.
Some recovered solutions

Left to right: recovered solution by Acc.ITA and ITA with relative error 0.5
Some recovered solutions

Left to right: recovered solution by Acc.ITA and ITA with relative error 0.2
References


