Dynamics-driven error reduction for extremely large problems in geophysics

Emmanouil Daskalakis
Felix J. Herrmann
Rachel Kuske
Mengmeng Yang
Motivation: least-squares migration

Consider

$$\min_x \|x\|_1$$

s.t.

$$\sum_{i=1}^{n_s} \|J_i[m_0, q_i]C^*x - b_i\|_2 \leq \sigma$$

- $x$ is the vector of Curvelet coefficients,
- $J_i$ is the Born modelling operator,
- $m_0$ is the background model for the velocity,
- $b_i$ is the vectorized reflection of the $i$-th shot,
- $C^*$ is the transpose of the curvelet transform,
- $\sigma$ is the tolerance for noise.
Motivation: least-squares migration

Consider

\[
\min_x \|x\|_1 \\
\text{s.t.} \\
\sum_{i=1}^{n_s} \|J_i[m_0, q_i]C^*x - b_i\|_2 \leq \sigma
\]

- \( x \) is the vector of Curvelet coefficients,
- \( J_i \) is the Born modelling operator,
- \( m_0 \) is the background model for the velocity,
- \( b_i \) is the vectorized reflection of the \( i \)-th shot,
- \( C^* \) is the transpose of the curvelet transform,
- \( \sigma \) is the tolerance for noise.

Solution w/ Linearized Bregman (LB)

\[
\begin{align*}
    z_{k+1} &= z_k - t_k A_k^T(A_kx_k - b_k) \\
    x_{k+1} &= S_\lambda(z_{k+1}).
\end{align*}
\]

where

\[
S_\lambda(z_k) = \max(|z_k| - \lambda, 0) \text{sign}(z_k)
\]

and

\[
t_k = \frac{\|A_kx_k - b_k\|_2^2}{\|A_k^T A_kx_k - b_k\|_2^2}
\]
Motivation: least-squares migration

Consider

$$\min_{\mathbf{x}} \| \mathbf{x} \|_1$$

s.t.

$$\sum_{i=1}^{n_s} \| \mathbf{J}_i [\mathbf{m}_0, \mathbf{q}_i][\mathbf{C}^* \mathbf{x} - \mathbf{b}_i] \|_2 \leq \sigma$$

- $\mathbf{x}$ is the vector of Curvelet coefficients,
- $\mathbf{J}_i$ is the Born modelling operator,
- $\mathbf{m}_0$ is the background model for the velocity,
- $\mathbf{b}_i$ is the vectorized reflection of the i-th shot,
- $\mathbf{C}^*$ is the transpose of the curvelet transform,
- $\sigma$ is the tolerance for noise.

Solution w/ Linearized Bregman (LB)

$$z_{k+1} = z_k - t_k A_k^T (A_k x_k - b_k)$$

$$x_{k+1} = S_\lambda(z_{k+1})$$

where

$$S_\lambda(z_k) = \max(|z_k| - \lambda, 0) \text{sign}(z_k)$$

and

$$t_k = \frac{\|A_k x_k - b_k\|_2^2}{\|A_k^T (A_k x_k - b_k)\|_2^2}$$

NOTE: Subsampling is necessary for large data sets
Motivation: least-squares migration

Consider

\[
\min_x \|x\|_1 \\
\text{s.t.} \sum_{i=1}^{n_s} \|J_i[m_0, q_i]C^*x - b_i\|_2 \leq \sigma
\]

- \(x\) is the vector of Curvelet coefficients,
- \(J_i\) is the Born modelling operator,
- \(m_0\) is the background model for the velocity,
- \(b_i\) is the vectorized reflection of the i-th shot,
- \(C^*\) is the transpose of the curvelet transform,
- \(\sigma\) is the tolerance for noise.

Solution w/ Linearized Bregman (LB)

\[
\begin{align*}
z_{k+1} &= z_k - t_k A_k^T(A_k x_k - b_k) \\
x_{k+1} &= S_\lambda(z_{k+1}).
\end{align*}
\]

\[
x = b \\
A \rightarrow \text{A}_r(k) \rightarrow x = b_r(k)
\]
Motivation: least-squares migration

Consider

\[
\min_x \|x\|_1 \\
\text{s.t.} \sum_{i=1}^{n_s} \|J_i[m_0, q_i]C^*x - b_i\|_2 \leq \sigma
\]

- $x$ is the vector of Curvelet coefficients,
- $J_i$ is the Born modelling operator,
- $m_0$ is the background model for the velocity,
- $b_i$ is the vectorized reflection of the $i$-th shot,
- $C^*$ is the transpose of the curvelet transform,
- $\sigma$ is the tolerance for noise.

Solution w/ Linearized Bregman (LB)
Motivation: least-squares migration

Consider

$$\min_x \|x\|_1$$

s.t.

$$\sum_{i=1}^{n_s} \|J_i[m_0, q_i]C^*x - b_i\|_2 \leq \sigma$$

- $x$ is the vector of Curvelet coefficients,
- $J_i$ is the Born modelling operator,
- $m_0$ is the background model for the velocity,
- $b_i$ is the vectorized reflection of the $i$-th shot,
- $C^*$ is the transpose of the curvelet transform,
- $\sigma$ is the tolerance for noise.
**LB Method:** w/ weighted increment

**LB method**

\[
\begin{align*}
  z_{k+1} &= z_k - t_k A_k^T (A_k x_k - b_k) \\
  x_{k+1} &= S_\lambda(z_{k+1}),
\end{align*}
\]

where

\[
S_\lambda(z_k) = \max(|z_k| - \lambda, 0) \text{sign}(z_k)
\]

and

\[
t_k = \frac{||A_k x_k - b_k||_2^2}{||A_k^T (A_k x_k - b_k)||_2^2}
\]
**LB Method:** w/ weighted increment

\[
\begin{align*}
z_{k+1} &= z_k - t_k A_k^T (A_k x_k - b_k) \\
x_{k+1} &= S_\lambda(z_{k+1}),
\end{align*}
\]

where

\[
S_\lambda(z_k) = \max(|z_k| - \lambda, 0) \text{sign}(z_k)
\]

and

\[
t_k = \frac{||A_k x_k - b_k||_2^2}{||A_k^T (A_k x_k - b_k)||_2^2}
\]

**LB method w/ weighted increment**

\[
\begin{align*}
z_{k+1} &= z_k - \tau_k \odot A_k^T (A_k x_k - b_k) \\
x_{k+1} &= S_\lambda(z_{k+1}),
\end{align*}
\]

where

\[
S_\lambda(z_k) = \max(|z_k| - \lambda, 0) \text{sign}(z_k)
\]

and

\[
\tau_k^i = t_k \sum_{j=1}^k \text{sign}(A_j^T (A_j x_j - b_j))
\]

with

\[
t_k = \frac{||A_k x_k - b_k||_2^2}{||A_k^T (A_k x_k - b_k)||_2^2}
\]
LB method

LB method w/ weighted increment
LB method

LB method w/ weighted increment
Toy problem

$A \in \mathbb{R}^{20000 \times 1000}$

Gaussian matrix

$x \in \mathbb{R}^{1000}$

Sparse vector

$\ast$

$b \in \mathbb{R}^{20000}$

Noisy data vector

$x \ast A = b$
Optimization problems

1. $l_1$-minimization problem (consistent)
   \[ \min_x \|x\|_1 \quad \text{s.t.} \quad Ax = b \]

2. BPDN problem (inconsistent)
   \[ \min_x \|x\|_1 \quad \text{s.t.} \quad \|Ax - b\|_2 \leq \sigma \]
Intuition: gradient entry for weighted increments

(for consistent problems)

\[
[A_k^T (A_k x_k - b_k)]_{136}
\]

\[
[A_k^T (A_k x_k - b_k)]_{147}
\]

Largest entry of the exact solution \( x^* \)

One of the small entries of the exact solution \( x^* \)
**Intuition:** gradient entry for weighted increments  
(for consistent problems)  

\[
\begin{align*}
[A_k^T(A_kx_k - b_k)]_{136} & \quad [A_k^T(A_kx_k - b_k)]_{147}
\end{align*}
\]

Largest entry of the exact solution \(x^*\)  

One of the small entries of the exact solution \(x^*\)  

Dashed line represents 1 pass from the data
Intuition: gradient entry for weighted increments

(for inconsistent problem)

\[ [A^T_k (A_k x_k - b_k)]_{136} \]

\[ [A^T_k (A_k x_k - b_k)]_{147} \]

Largest entry of the exact solution \( x^* \)

One of the small entries of the exact solution \( x^* \)
Intuition: Behaviour of the new weighted increment

\[ [A^T_k (A_k x_k - b_k)]_{136} \]

Weighted increment

\[
\tau_k^i = t_k \frac{\left| \sum_{j=1}^{k} \text{sign}([A_j^T (A_j x_j - b_j)]_i) \right|}{k}
\]
Intuition: Behaviour of the new weighted increment

\[ \left[ A_k^T(A_k x_k - b_k) \right]_{136} \]
**Intuition:** Behaviour of the new weighted increment

\[ [A_k^T (A_k x_k - b_k)]_{147} \]
Intuition: Behaviour of the entries of the solution

Consistent problem

Inconsistent problem
**Intuition:** Behaviour of the entries of the solution

- **Consistent problem**
- **Inconsistent problem** (w/ weighted increment)

\[ z_k(136) \]
\[ z_k(147) \]
Effect on the LSRTM problem

LB

LB w/ weighted increment
Effect on the LSRTM problem

LB

LB w/ weighted increment
Effect on the LSRTM problem

LB

LB w/ weighted increment
Effect on the LSRTM problem

LB

LB w/ weighted increment
Effect on the LSRTM problem

LB

LB w/ weighted increment
Effect on the LSRTM problem

LB

LB w/ weighted increment
Effect on large problems

Problem a

- The vector $x$ corresponds to a known vector of curvelet coefficients
- $A_k \in \mathbb{R}^{500000 \times 2040739}$
- The signal to noise ratio for the data in both problems is the same.

Problem a: $A \in \mathbb{R}^{1000000 \times 2040739}$

Problem b: $A \in \mathbb{R}^{7000000 \times 2040739}$
Effect on large problems

Problem b

Problem a: \( A \in \mathbb{R}^{1000000 \times 2040739} \)

Problem b: \( A \in \mathbb{R}^{7000000 \times 2040739} \)

- The vector \( x \) corresponds to a known vector of curvelet coefficients
- \( A_k \in \mathbb{R}^{500000 \times 2040739} \)
- The signal to noise ratio for the data in both problems is the same.
Future goals

- Using the Weighted Increment in other iterative methods (SGD, Kaczmarz etc) is possible.
- Improve the convergence speed of Weighted Increment.
Acknowledgement

This research was carried out as part of the SINBAD project with the support of the member organizations of the SINBAD Consortium.