An Iterative, Projection-Based Algorithm for General Form Tikhonov Regularization

Misha Kilmer, Tufts University
Per Christian Hansen, Technical University of Denmark
Malena Español, Tufts University
Outline

- Problem
- Background
- Algorithm
- Numerical Examples
- Conclusion and future work
Discrete Ill-Posed problem

Find $x_{true}$, given $A$, $b$ and the model

$$Ax_{true} = b = b_{true} + e,$$

where $A \in R^{m \times n}$ is a large, ill-conditioned matrix

Properties:
• Decaying singular values without gap
• $e$ is unknown (white) noise
• Discrete Picard condition holds
Need for regularization

Let \( A = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T \) be the SVD.

The exact solution is given by

\[
\chi_{\text{true}} = \sum_{i=1}^{n} \frac{u_i^T b_{\text{true}}}{\sigma_i} v_i + \frac{u_i^T e}{\sigma_i} v_i
\]
Tikhonov Regularization

The Tikhonov Regularized Problem:

$$\min_x \left\{ \|Ax - b\|_2^2 + \lambda^2 \|Lx\|_2^2 \right\}$$

or

$$\min_x \left\| \begin{bmatrix} A \\ \lambda L \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2^2.$$

Conditioning depends on regularization parameter \( \lambda > 0 \).
Tikhonov Method: choosing $\lambda$

L-curve (Lawson-Hansen)

\[
\log \| Lx \|_2
\]

\[
\log \| Ax - b \|_2
\]
Tikhonov Regularization

- In standard form Tikhonov, \( L = I \).
- If \( L \neq I \):

\[
\min_y \left\{ \| AL_A^+ y - b \|_2^2 + \lambda^2 \| y \|_2^2 \right\}, \quad x^{(\lambda)} = L_A^+ y^{(\lambda)}
\]

where \( L_A^+ \) is the A-weighted pseudo-inverse.
- \( L \) is often a (scaled) discrete derivative operator.
- Iterative solvers (CGLS, LSQR) employed.
Bidiagonalization

Consider the QR factorization

\[
\begin{bmatrix}
A \\
L
\end{bmatrix} = \begin{bmatrix}
Q_A \\
Q_L
\end{bmatrix} R, \quad Q_A^T Q_A + Q_L^T Q_L = I.
\]

- Bidiagonalization of \( Q_A \) and \( Q_L \) can be computed simultaneously.
  Obtaining a lower bidiagonal form of \( Q_A \) and an upper bidiagonal form of \( Q_L \).
- No need of forming QR factorization explicitly
  Golub - Kahan '65, Zha '96, Paige and Saunders '82
Relating $A, L$

Theorem: Let $A \in R^{m \times n}$ and $L \in R^{p \times n}$. There exist unitary $m \times m$ matrix $U$, unitary $p \times p$ matrix $\hat{U}$, a lower bidiagonal $B \in R^{m \times n}$, an upper bidiagonal $\overline{B} \in R^{p \times n}$, and invertible $n \times n$ matrix $Z$ such that

$$A = U B Z^{-1}, \quad L = \hat{U} \overline{B} Z^{-1}.$$
Projected Problem

Replace the Tikhonov problem with

$$\min_{x \in Z_k} \left\| \begin{bmatrix} A \\ \lambda \mathbf{L} \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2,$$

with $Z_k = \text{span} \{z_1, \ldots, z_k\}$.

But $U_{k+1}(\beta_1 e_1) = b$ so with its solution $x_k = Z_k y_k$,

$y_k$ is the solution of the projected problem

$$\min_{y} \left\| \begin{bmatrix} B_k \\ \lambda \mathbf{B}_k \end{bmatrix} y \right\|_2,$$

Only $x_k$ and $y_k$ depend on $\lambda$. 

Choosing $\lambda$

L - curve

Theorem: for $x_k, y_k$ as defined for each $\lambda$.

$$\|Ax_k - b\|_2^2 = \|B_k y_k - \beta_1 e_1\|_2^2 \quad \text{and} \quad \|Lx_k\|_2^2 = \|\overline{B}_k y_k\|_2^2$$
Iterative Method

For $k = 1, 2, \ldots$

1. compute the $k$th step of joint bidiagonalization,
2. compute $y_k$, $(x_k$, if desired) for each $\lambda$ from quantities computed at step $k - 1$,
3. update residual and constraint norms,
4. check for convergence.
Regularizing algorithm

• We can show that the algorithm works similar to LSQR applied to the problem
  \[
  \min_{y} \| Q_A y - b \|_2^2, \quad \text{with } x = R^{-1} y.
  \]

• \(Q_A\) inherits conditioning properties from \(A\): the singular values have similar decaying behavior.

• with the CS decomposition of \(\{B_k, \overline{B_k}\}\) we get an approximation of the GSVD of \(\{A, L\}\)
TGSVD of the Projected Problem

Recall the Projected Problem:

\[
\min_y \left\| \begin{bmatrix} B_k \\ \lambda \overline{B}_k \end{bmatrix} y - \begin{bmatrix} \beta_1 e_1 \\ 0 \end{bmatrix} \right\|_2, \text{ with } x^k = Z_k y_k.
\]

Let

\[
\begin{bmatrix} B_k \\ \overline{B}_k \end{bmatrix} = \begin{bmatrix} U_{B_k} & 0 \\ 0 & U_{\overline{B}_k} \end{bmatrix} \begin{bmatrix} N \\ M \end{bmatrix} H^{-1}
\]

be the GSVD of \( \{B_k, \overline{B}_k\} \).

The TGSVD solution to the projected problem is

\[
y^k_l = \sum_{i=1}^l \frac{(u_{B_k})^T_i (\beta_1 e_1)}{\eta_i} h_i \text{ and then } x^k_{\text{reg}} = Z_k y^k_l
\]
Numerical Examples

2D Image deblurring, Matlab, Regularization Tools by Hansen.

\[ A = \text{blur()}, \quad k(A) = O(10^0) \]

\( x_{true} = \text{"rice"} \) and \( b = Ax_{true} + e \) with \( e \approx N(0, 0.01 \|b\|) \).

\[ L = \begin{bmatrix} I \otimes L_1 \\ L_1 \otimes I \end{bmatrix}, \quad \text{where } L_1 = \text{derivativeop. of dim 1 or } L = \Delta. \]

\( A, L \) scaled so that \( \|A\|_2 = \|L\|_2 = 1 \)
Original and blurred images

Original

Blurred + Error
Restoration $L = \text{derivative operator}$
Restoration $L = \text{Laplacian}$
Comparing with \( L=I \) and \( L=L\text{Laplacian} \)
Conclusion and Future work

- Iterative algorithm that avoids standard - form transformation useful when:
  - $L^+_A$ is difficult to compute/store
  - $\lambda$ not known a priori

- Future work:
  - Inner iteration accuracy
  - Preconditioning