Reconstruction methods for sparse-data tomography

Part D: Regularization

Samuli Siltanen

Department of Mathematics and Statistics
University of Helsinki, Finland
samuli.siltanen@helsinki.fi
www.siltanen-research.net

INDAM intensive period
Computational Methods for Inverse Problems in Imaging
Como, Italy, June 26–28, 2018
Inverse problem of X-ray tomography: given noisy sinogram, find a stable approximation to $f$

Model space $X = \mathbb{R}^{32 \times 32}$

Data space $Y = \mathbb{R}^{32 \times 49}$

$D(A)$

$A$ (operator)

$Af = m$

$A(D(A))$
Robust solution of ill-posed inverse problems requires regularization.

Model space $X = \mathbb{R}^{32 \times 32}$

Data space $Y = \mathbb{R}^{39 \times 49}$

We need to define a family of continuous functions $\Gamma_\alpha : Y \to X$ so that the reconstruction error $\|\Gamma_{\alpha(\delta)}(m) - x\|_X$ vanishes asymptotically at the zero-noise level $\delta \to 0$. 

\[ D(A) \quad \Gamma_\alpha(m) \quad \Gamma_\alpha \quad \Gamma_{\alpha(\delta)}(m) \quad A \quad Af \quad m \]
Outline

Tikhonov regularization

Frame-sparsity methods

Hospital case study: diagnosing osteoarthritis
Tikhonov regularization is the classical method for noise-robust tomographic reconstruction.

Write a penalty functional

\[ \Phi(f) = \|Af - m\|_2^2 + \alpha\|f\|_2^2, \]

where \(0 < \alpha < \infty\) is a regularization parameter. Define \(\Gamma_\alpha(m)\) by

\[ \Phi(\Gamma_\alpha(m)) = \min_{f \in X} \{ \Phi(f) \}. \]

We denote

\[ \Gamma_\alpha(m) = \arg\min_{f \in X} \{ \|Af - m\|_2^2 + \alpha\|f\|_2^2 \}. \]
Tikhonov regularization can be expressed as filtering the singular values of the matrix $A$

$\Gamma_\alpha(m) = V \begin{bmatrix}
\frac{d_1}{d_1^2 + \alpha} & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \frac{d_{\min\{k,n\}}}{d_{\min\{k,n\}}^2 + \alpha}
\end{bmatrix} U^T m$

In large-scale computations it is better to use the formula

$\Gamma_\alpha(m) = (A^T A + \alpha I)^{-1} A^T m$. 
Standard Tikhonov regularization

\[
\arg\min_{f \in \mathbb{R}^n} \left\{ \|Af - m\|_2^2 + \alpha \|f\|_2^2 \right\}
\]

Original phantom sampled at 32×32 resolution

Reconstruction

Relative square norm error 12%
Non-negative Tikhonov regularization

$$\text{arg min}_{f \in \mathbb{R}^n_+} \left\{ \| Af - m \|_2^2 + \alpha \| f \|_2^2 \right\}$$

Original phantom sampled at 32×32 resolution

Reconstruction
Relative square norm error 10%
Recall the \( L^p \) norms for \( \mathbb{R}^n \)

Let \( f \in \mathbb{R}^n \). The \( L^p \) norms for \( 1 \leq p < \infty \) are defined by

\[
\| f \|_p = \left( \sum_{j=1}^{n} |f_j|^p \right)^{1/p}.
\]

In particular we use the following two cases:

\[
\| f \|_2^2 = \sum_{j=1}^{n} |f_j|^2, \quad \| f \|_1 = \sum_{j=1}^{n} |f_j|.
\]
Total variation (TV) regularization is a technique for preserving edges in the reconstruction

We consider calculating the minimizer of the TV functional

\[ \| Af - m \|_2^2 + \alpha \{ \| L_H f \|_1 + \| L_V f \|_1 \} \]

\[ = \| Af - m \|_2^2 + \alpha \left\{ \sum_j \sum_i (| f_{i(j+1)} - f_{ij} | + | f_{(i+1)j} - f_{ij} |) \right\} \]

where \( L_H \) and \( L_V \) are horizontal and vertical first-order difference matrices. [Rudin, Osher and Fatemi 1992]
TV tomography: \( \text{arg min}_{f \in \mathbb{R}^n} \{ \| Af - m \|_2^2 + \alpha \| \nabla f \|_1 \} \)

1992 Rudin, Osher & Fatemi: denoise images by taking \( A = I \)
1998 Delaney & Bresler
2001 Persson, Bone & Elmqvist
2003 Kolehmainen, S, Järvenpää, Kaipio, Koistinen, Lassas, Pirttilä & Somersalo (first TV work with measured X-ray data)
2006 Kolehmainen, Vanne, S, Järvenpää, Kaipio, Lassas & Kalke
2006 Sidky, Kao & Pan
2008 Liao & Sapiro
2008 Sidky & Pan
2008 Herman & Davidi
2009 Tang, Nett & Chen
2009 Duan, Zhang, Xing, Chen & Cheng
2010 Bian, Han, Sidky, Cao, Lu, Zhou & Pan
2011 Jensen, Jørgensen, Hansen & Jensen
2011 Tian, Jia, Yuan, Pan & Jiang
2012–present: hundreds of articles indicated by Google Scholar
There are many computational approaches for computing the minimum

**Primal-dual algorithms** Chambolle, Chan, Chen, Esser, Golub, Mulet, Nesterov, Zhang

**Thresholding** Candès, Chambolle, Chaux, Combettes, Daubechies, Defrise, DeMol, Donoho, Pesquet, Starck, Teschke, Vese, Wajs

**Bregman iteration** Cai, Burger, Darbon, Dong, Goldfarb, Mao, Osher, Shen, Xu, Yin, Zhang

**Splitting approaches** Chan, Esser, Fornasier, Goldstein, Langer, Osher, Schönlieb, Setzer, Wajs

**Nonlocal TV** Bertozzi, Bresson, Burger, Chan, Lou, Osher, Zhang

We found that *quadratic programming* works well for us.
The minimizer of the functional

$$\arg\min_{f \in \mathbb{R}^n_+} \left\{ \|Af - m\|_2^2 + \alpha \|L_Hf\|_1 + \alpha \|L_Vf\|_1 \right\}$$

can be transformed into the standard form

$$\arg\min_{z \in \mathbb{R}^{5n}} \left\{ \frac{1}{2} z^T Q z + c^T z \right\}, \quad z \geq 0, \quad Ez = b,$$

where $Q$ is symmetric and $E$ implements equality constraints.

Large-scale primal-dual interior point QP method was developed in Kolehmainen, Lassas, Niinimäki & S (2012) and Hämäläinen, Kallonen, Kolehmainen, Lassas, Niinimäki & S (2013).
Reduction to \( \arg \min_{z \in \mathbb{R}^{5n}} \left\{ \frac{1}{2} z^T Q z + c^T z \right\} \)

Denote horizontal and vertical differences by

\[ L_H f = u_H^+ - u_H^- \quad \text{and} \quad L_V f = u_V^+ - u_V^- , \]

where \( u_H^\pm, u_V^\pm \geq 0 \). TV minimization is now

\[ \arg \min_{f \in \mathbb{R}_+^n} \left\{ f^T A^T A f - 2 f^T A^T m + \alpha 1^T (u_H^+ + u_H^- + u_V^+ + u_V^-) \right\} , \]

where \( 1 \in \mathbb{R}^n \) is vector of all ones. Further, we denote

\[ z = \begin{bmatrix} f \\ u_H^+ \\ u_H^- \\ u_V^+ \\ u_V^- \end{bmatrix} , \quad Q = \begin{bmatrix} \frac{1}{\sigma^2} A^T A & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} , \quad c = \begin{bmatrix} -2 A^T m \\ \alpha 1 \\ \alpha 1 \\ \alpha 1 \\ \alpha 1 \end{bmatrix} . \]
Non-negative TV regularization

\[
\arg \min_{f \in \mathbb{R}^n_+} \left\{ \| Af - m \|_2^2 + \alpha \| \nabla f \|_1 \right\}
\]

Original phantom sampled at 32\times32 resolution

TV regularized reconstruction
Relative square norm error 7%
Let’s consider a square phantom

\[ f \in \mathbb{R}^{32 \times 32} \]

\[ Af \in \mathbb{R}^{49 \times 39} \]
Naive reconstruction using the Moore-Penrose pseudoinverse; data has 0.1% relative noise.

Original phantom, values between zero (black) and one (white).

Naive reconstruction with minimum $-14.9$ and maximum $18.5$. 
Standard Tikhonov regularization

$$\arg \min_{f \in \mathbb{R}^n} \left\{ \|Af - m\|_2^2 + \alpha \|f\|_2^2 \right\}$$

Original phantom

Reconstruction

Relative square norm error 35%
Constrained Tikhonov regularization

\[
\arg \min_{f \in \mathbb{R}_+^n} \left\{ \| Af - m \|^2_2 + \alpha \| f \|^2_2 \right\}
\]

Original phantom

Reconstruction
Relative square norm error 13%
Constrained total variation (TV) regularization

\[
\text{arg min}_{f \in \mathbb{R}^n_+} \left\{ \| Af - m \|^2_2 + \alpha \left( \| L_h f \|_1 + \| L_v f \|_1 \right) \right\}
\]
In variational regularization, the penalty term expresses *a priori* knowledge about the unknown

**Standard Tikhonov regularization:**

\[
\underset{f \in \mathbb{R}^n}{\arg\min} \left\{ \| Af - m \|^2_2 + \alpha \| f \|^2_2 \right\}
\]

**Non-negativity constrained Tikhonov regularization:**

\[
\underset{f \in \mathbb{R}_+^n}{\arg\min} \left\{ \| Af - m \|^2_2 + \alpha \| f \|^2_2 \right\}
\]

**Non-negativity constrained Total Variation (TV) regularization:**

\[
\underset{f \in \mathbb{R}_+^n}{\arg\min} \left\{ \| Af - m \|^2_2 + \alpha \| \nabla f \|_1 \right\}
\]
Outline

Tikhonov regularization

Frame-sparsity methods

Hospital case study: diagnosing osteoarthritis
Daubechies, Defrise and de Mol introduced a revolutionary inversion method in 2004

Consider the sparsity-promoting variational regularization

\[
\arg \min_{f \in \mathbb{R}^n} \left\{ \|Af - m\|_2^2 + \mu \|Wf\|_1 \right\},
\]

where \(W\) is an orthonormal wavelet transform. The minimizer can be computed using the iteration

\[
f_{j+1} = W^{-1} S_\mu W \left( f_j + A^T (m - Af_j) \right),
\]

where the soft-thresholding operation

\[
S_\mu(x) = \begin{cases} 
  x + \frac{\mu}{2} & \text{if } x \leq -\frac{\mu}{2}, \\
  0 & \text{if } |x| < \frac{\mu}{2}, \\
  x - \frac{\mu}{2} & \text{if } x \geq \frac{\mu}{2},
\end{cases}
\]

is applied to each wavelet coefficient separately.
We modify the method so that non-negativity constraint has rigorous mathematical foundation

The minimizer

$$\arg\min_{f \in \mathbb{R}_+^n} \left\{ \frac{1}{2} \| A f - m \|_2^2 + \mu \| W f \|_1 \right\}$$

can be computed using this iteration:

$$y^{(i+1)} = \mathbb{P}_C \left( f^{(i)} - \tau \nabla g(f^{(i)}) - \lambda W^T v^{(i)} \right)$$

$$v^{(i+1)} = \left( I - S_{\mu} \right) \left( W y^{(i+1)} + v^{(i)} \right)$$

$$f^{(i+1)} = \mathbb{P}_C \left( f^{(i)} - \tau \nabla g(f^{(i)}) - \lambda W^T v^{(i+1)} \right)$$

where $\tau > 0$, $\lambda > 0$ and $g(f) = \frac{1}{2} \| A f - m \|_2^2$. Here $\mathbb{P}_C$ denotes projection to the non-negative “quadrant.”

[Loris & Verhoeven 2011], [Chen, Huang & Zhang 2016]
Illustration of the Haar wavelet transform
Sparse-data reconstruction of the walnut using Haar wavelet sparsity

Filtered back-projection

Constrained Besov regularization

\[
\arg\min_{f \in \mathbb{R}^n_+} \left\{ \|Af - m\|_2^2 + \alpha \|f\|_{B_{11}} \right\}
\]
How to choose the thresholding parameter $\mu$? Here it is too small.
How to choose the thresholding parameter $\mu$? Here it is too large.
Automatic parameter choice using controlled wavelet-domain sparsity (CWDS)

Assume given the *a priori* sparsity level $0 \leq C_{pr} \leq 1$. Denote by $C_j$ the sparsity of the $j$th iterate $f_j \in \mathbb{R}^n$:

$$C_j = \frac{\text{(number of nonzero elements in } Wf_j \in \mathbb{R}^n)}{n}.$$  

The CWDS iteration is based on proportional-integral-derivative (PID) controllers:

$$\mu^{(i+1)} = \mu^{(i)} + \beta(C^{(i)} - C_{pr}).$$

[Purisha, Rimpeläinen, Bubba & S 2018]
CWDS choice of the thresholding parameter $\mu$
CWDS choice of the thresholding parameter $\mu$
Shearlet coefficients at coarse scale 1/8

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at coarse scale 3/8

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at coarse scale $5/8$

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at coarse scale 7/8

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at coarse scale 8/8

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at medium scale 4/8

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at medium scale 6/8

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale $1/16$

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale 2/16

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale $3/16$

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale 5/16

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale 6/16

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale 8/16

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale 9/16

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale 12/16

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale $13/16$

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
Shearlet coefficients at fine scale 14/16

We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
We use Shearlab [Kutyniok, Shahram & Zhuang 2012].
The shearlet transform gives multi-resolution and orientation-aware building blocks for image data.

Schematic diagram of the frequency plane tiling of several elements of a 2D shearlet system, for different values of dilation and shearing parameters.
Outline

Tikhonov regularization

Frame-sparsity methods

Hospital case study: diagnosing osteoarthritis
This is a joint work with

Tatiana Bubba, University of Helsinki, Finland

Sakari Karhula, Oulu University Hospital, Finland

Juuso Ketola, Oulu University Hospital, Finland

Maximilian März, TU Berlin

Miika T. Nieminen, University of Oulu, Finland

Zenith Purisha, University of Helsinki, Finland

Juho Rimpeläinen, University of Helsinki, Finland

Simo Saarakkala, Oulu University Hospital, Finland
We consider small specimens of human bone imaged using microtomography.

Slice of 3D reconstruction by FDK based on **596 angles**

Three-dimensional structure
We pick out a smaller region of interest for osteoarthritis analysis.

Slice of 3D reconstruction by FDK based on **596 angles**
Slice of 3D region of interest after binary thresholding
We use two numerical quality measures applied to segmented three-dimensional bone structure.

- Trabecular thickness
- Trabecular separation

[Bouxsein, Boyd, Christiansen, Guldberg, Jepsen, & Müller 2010]
The goal is to reduce measurement time by recording fewer radiographs

3D FDK reconstruction based on **40 angles**

3D shearlet-sparsity reconstruction based on **40 angles**
Bone quality parameters from ground truth

![Image with bone projections]

<table>
<thead>
<tr>
<th>Thickness</th>
<th>Separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.34</td>
<td>0.71</td>
</tr>
<tr>
<td>0.37</td>
<td>0.35</td>
</tr>
</tbody>
</table>

[Purisha, Karhula, Rimpeläinen, Nieminen, Saarakkala & S, submitted]
Results from FDK reconstructions

<table>
<thead>
<tr>
<th>Projections</th>
<th>Thickness</th>
<th>Separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>0.34</td>
<td>0.71</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[Projections: 300]

<table>
<thead>
<tr>
<th>Projections</th>
<th>Thickness</th>
<th>Separation</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>0.37</td>
<td>0.35</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[Prousha, Karhula, Rimpeläinen, Nieminen, Saarakkala & S, submitted]
Results from 3D shearlet-sparsity reconstructions

- Projections: 300
  - Thickness: 0.34
  - Separation: 0.71
- Projections: 50
  - Thickness: 0.37
  - Separation: 0.35
- Projections: 30
  - Thickness: 0.37
  - Separation: 0.35

[Purisha, Karhula, Rimpeläinen, Nieminen, Saarakkala & S, submitted]
Thank you for your attention!