Variational Methods in CT Reconstruction

Chapter 11.4 and 11.5

Yiqiu Dong

DTU Compute
Technical University of Denmark

20 Jan. 2021
MAP estimate

The **MAP estimation** problem is

\[
\min_x \frac{1}{2} \|A x - b\|_2^2 + \alpha J(x).
\]

- The term \( \frac{1}{2} \|A x - b\|_2^2 \) is called the *data-fidelity* term.
- The term \( J(x) \) is called the *regularization* term.
- \( \alpha > 0 \) is the regularization parameter.
- *Tikhonov regularization*: \( J(x) = \frac{1}{2} \|x\|_2^2 \).
- *Tikhonov regularization in general form*: \( J(x) = \frac{1}{2} \|Dx\|_2^2 \).
An example using a continuous function

Consider the piecewise linear function

\[ f(t) = \begin{cases} 
0, & 0 \leq t < \frac{1}{2}(1 - h) \\
\frac{t}{h} - \frac{1-h}{2h}, & \frac{1}{2}(1 - h) \leq t \leq \frac{1}{2}(1 + h) \\
1, & \frac{1}{2}(1 + h) < t \leq 1 
\end{cases} \]

which increases linearly from 0 to 1 in \([\frac{1}{2}(1 - h), \frac{1}{2}(1 + h)]\).
Norms of the First Derivative

It is easy to show that the 1- and 2-norms of $f'(t)$ satisfy

$$
\|f'\|_1 = \int_0^1 |f'(t)| \, dt = \int_0^h \frac{1}{h} \, dt = 1,
$$

$$
\|f'\|_2^2 = \int_0^1 f'(t)^2 \, dt = \int_0^h \frac{1}{h^2} \, dt = \frac{1}{h}.
$$

Note that $\|f'\|_1$ is independent of the slope of the middle part of $f(t)$, while $\|f'\|_2$ penalizes steep gradients (when $h$ is small).

- The 2-norm of $f'(t)$ will not allow any steep gradients and therefore it produces a smooth solution.
- The 1-norm, on the other hand, allows some steep gradients – but not too many – and it is therefore able to produce a less smooth solution, and even a discontinuous solution.
Total Variation (TV) Regularization

- **In 1D:**

\[
TV(x) = \| D_n x \|_1 ,
\]

where \( \| x \|_1 \) denotes 1-norm of vector \( x \) defined as \( \| x \|_1 = |x_1| + \cdots + |x_n| \).

- **In 2D:**

\[
TV_a(x) = \| (I_N \otimes D_M)x \|_1 + \| (D_N \otimes I_M)x \|_1 \\
= \sum_{i=1}^{n} \left( |i_i^T (I_N \otimes D_M)x| + |i_i^T (D_N \otimes I_M)x| \right),
\]

\[
TV_i(x) = \sum_{i=1}^{n} \sqrt{ (i_i^T (I_N \otimes D_M)x)^2 + (i_i^T (D_N \otimes I_M)x)^2 },
\]

where \( i_i \) denotes the \( i \)th column of the identity matrix \( I_n \) with \( n = M \times N \).
TV regularized CT reconstruction problem

$$\min_{x} \frac{1}{2} \| A x - b \|_2^2 + \alpha \text{TV} □ (x) ,$$

where □ stands for either “a” or “i”.

- By allowing occasional larger jumps in the reconstruction, it leads to piecewise constant results with sharp edges.
- The price is that the problem is non-differentiable.
Total Variation in Continuous Setting (1D)

The total variation of \( f : [a, b] \to \mathbb{R} \) is defined as

\[
TV(f) = \int_a^b |Df| = \sup_{p \in \mathcal{P}} \sum_{i=1}^{n_p} |f(t_i) - f(t_{i-1})|,
\]

where \( \mathcal{P} = \{ p = \{t_0, \cdots, t_{n_p}\} | p \) is a partition of \([a, b] \} \). The supremum is taken over all partitions of the interval \([a, b] \).

- If \( f \) is differentiable, then \( TV(f) = \int_a^b |f'(t)|\ dt \)
- Example: If \( f(t) = \begin{cases} -1, & \text{if } -1 \leq t \leq 0, \\ 1, & \text{if } 0 < t \leq 1 \end{cases} \), then \( TV(f) = 2 \).
- Bounded variation (BV) space is defined as

\[
f \in BV([a, b]) \iff TV(f) < +\infty.
\]

- \( W^{1,1}([a, b]) \subset BV([a, b]) \subset L^1([a, b]) \).
- In discrete case, \( TV(f) = \sum_{i=1}^{n} |f'(t_i)| \).
Sparse regularization

**Idea:** Look for a “sparse” solution with many zero elements.

\[
\min_x \frac{1}{2} \|Ax - b\|_2^2 + \alpha \|x\|_0
\]

- \(\|x\|_0 = \#\{i = 1, \ldots, n : x_i \neq 0\}\).
- It is computational very challenging.
- The 0-norm is often approximated by
  - the \(q\)-norm with \(0 < q < 1\) (from bridge regression);
  - the 1-norm (most commonly used).
Sparse regularization

**Idea:** Look for a “sparse” solution with many zero elements.

\[
\min_x \frac{1}{2} \| A x - b \|_2^2 + \alpha \| x \|_1
\]

(Lasso)

- The minimization problem is non-differentiable.
Sparse Regularization

In practice, sparsity regularization is often applied with respect to a certain orthonormal basis.

We consider an orthogonal basis \( \{ \psi_s \} \), e.g. wavelet basis, then formally we have

\[
\langle x, \psi_s \rangle = \langle A^\dagger b, \psi_s \rangle = \langle \bar{x}, \psi_s \rangle + \langle A^\dagger e, \psi_s \rangle.
\]

If the basis \( \{ \psi_s \} \) is efficient to represent \( \bar{x} \), i.e., the most of the coefficient \( \langle \bar{x}, \psi_s \rangle \) are close to zero. But noise leads to a large set of non-zero coefficients, so we can use the sparse regularization in order to obtain sparse coefficients.

\[
\min_x \frac{1}{2} \| A x - b \|_2^2 + \alpha \sum_s \| \langle x, \psi_s \rangle \|_1.
\]
Stair-casing artifacts from TV

\[
\min_x \frac{1}{2} \| b - A x \|_2^2 + \alpha \text{TV}_\square(x)
\]

Due to the relation between the sparsity and the 1-norm, TV regularization can be interpreted as sparsity regularization on the gradient, which leads to a piecewise constant solution.

- **Pro:** It can reconstruct sharp edges.
- **Con:** Smoothly varying parts in \( \bar{x} \) are often approximated by piecewise constant structures. This phenomena is called **stair-casing artifacts**.