

Inverse Problems 1: Convolution and Deconvolution

Lesson 9: Total Variation regularization

Rashmi Murthy

University of Helsinki

October 1, 2019

Outline

- Introduction
 - Motivation
 - Problem Statement
 - Scope and Objectives
- Literature Review
 - Existing Research
 - Gaps in Knowledge
- Methodology
 - Research Design
 - Data Collection
 - Analysis Techniques
- Results and Discussion
 - Key Findings
 - Interpretation of Results
- Conclusion
 - Summary of Findings
 - Implications and Future Work
- References
- Appendix
 - Additional Data
 - Supplementary Figures

Regularization

- Singular Value decomposition

Singular Value Decomposition

$A = UDV^T$ U, V orthogonal matrices

$D = \text{diag}\{d_1, d_2, \dots, d_n\}, \quad d_1 \geq d_2 \geq d_3 \dots \geq d_n$

- If $d_n > 0$, the the problem is well posed.
- If on the other hand $d_r > 0$, but $d_{r+1} = d_{r+2} = \dots = d_n = 0$, then existence, uniqueness and stability fails.
- We need regularization techniques.

Two Techniques so far

- Truncated Singular Value Decomposition (TSVD)

$$T_r(m) = \sum_{i=0}^r \frac{m^T u_i}{d_i} v_i$$

where d_i are the singular values and u_i and v_i are singular vectors.

- Tikhonov regularization

$$\begin{aligned} T_\alpha(m) &= \underset{f}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Af - m\|_2^2 + \alpha \|f\|_2^2 \right\} \\ &= \sum_{i=1}^k \frac{d_i(m^T u_i)}{d_i^2 + \alpha} v_i \end{aligned}$$

where α is the regularization parameter.

Generalized Tikhnov Regularization

- Generalized Tikhnov regularization

$$T_{\alpha}(m) = \underset{f}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Af - m\|_2^2 + \alpha \|Lf\|_2^2 \right\}$$

where L for instance can be the discrete differentiation matrix:

$$L = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & -1 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & & \vdots & \\ \vdots & & & \ddots & & \vdots & \\ 0 & & \dots & 0 & -1 & 1 & 0 \\ 0 & & \dots & 0 & 0 & -1 & 1 \\ 1 & & \dots & 0 & 0 & 0 & -1 \end{pmatrix}$$

Choice of Regularization parameter

- Morozov's Discrepancy Principle: Consider the residual

$$r(\alpha) = \|A f_{\alpha, \delta} - m_\delta\|, \quad f_{\alpha, \delta} = T_\alpha m_\delta$$

Morozov's Principle

Select α such that $r(\alpha) = \delta$

- Alternative formulation of Morozov:

$$r(\alpha) = \tau \delta$$

where $\tau = 1.2$

- L-curve method

Outline

- 1. Introduction
 - 1.1. Background and Motivation
 - 1.2. Scope and Objectives
- 2. Literature Review
 - 2.1. Existing Research
 - 2.2. Gaps and Contributions
- 3. Methodology
 - 3.1. Research Design
 - 3.2. Data Collection
 - 3.3. Analysis Techniques
- 4. Results
 - 4.1. Descriptive Statistics
 - 4.2. Inferential Analysis
 - 4.3. Key Findings
- 5. Discussion
 - 5.1. Interpretation of Results
 - 5.2. Implications and Applications
- 6. Conclusion
 - 6.1. Summary of Findings
 - 6.2. Future Research Directions
- 7. References
- 8. Appendix
 - 8.1. Additional Data
 - 8.2. Supplementary Figures

Total Variation

Definition

Let f be a real valued function defined on the interval $[a, b]$. The total variation of f , denoted by $TV(f)$, is defined to be

$$TV(f) = \sup \sum_{i=1}^k |f(x_i) - f(x_{i-1})|$$

where the supremum is over all partitions $a = x_0 < x_1 < \dots < x_k = b$ of $[a, b]$

- If f is piecewise continuous function with a finite number of jump discontinuities, then the total variation is the sum of the magnitude of the jumps.

Total Variation

Definition

Let f be a real valued function defined on the interval $[a, b]$. The total variation of f , denoted by $TV(f)$, is defined to be

$$TV(f) = \sup \sum_{i=1}^k |f(x_i) - f(x_{i-1})|$$

where the supremum is over all partitions $a = x_0 < x_1 < \dots < x_k = b$ of $[a, b]$

- If f is piecewise continuous function with a finite number of jump discontinuities, then the total variation is the sum of the magnitude of the jumps.

Total Variation

- If f is differentiable function, then let $\Delta x_i = x_i - x_{i-1}$ Consider

$$TV(f) = \sup \sum_{i=1}^k \frac{|f(x_i) - f(x_{i-1})|}{|x_i - x_{i-1}|} |x_i - x_{i-1}|$$

Taking the limit $\Delta x_i \rightarrow 0$, we get in one dimension,

$$TV(f) = \int_a^b |f'(x)| dx$$

- In higher dimensions, that is $\Omega = [a, b]^n$ The formula generalizes to

$$TV(f) = \int_{\Omega} |\nabla f(x)| dx$$

Total Variation

- If f is differentiable function, then let $\Delta x_i = x_i - x_{i-1}$ Consider

$$TV(f) = \sup \sum_{i=1}^k \frac{|f(x_i) - f(x_{i-1})|}{|x_i - x_{i-1}|} |x_i - x_{i-1}|$$

Taking the limit $\Delta x_i \rightarrow 0$, we get in one dimension,

$$TV(f) = \int_a^b |f'(x)| dx$$

- In higher dimensions, that is $\Omega = [a, b]^n$ The formula generalizes to

$$TV(f) = \int_{\Omega} |\nabla f(x)| dx$$

Total Variation

- If f is differentiable function, then let $\Delta x_i = x_i - x_{i-1}$ Consider

$$TV(f) = \sup \sum_{i=1}^k \frac{|f(x_i) - f(x_{i-1})|}{|\Delta x_i|} |\Delta x_i|$$

Taking the limit $\Delta x_i \rightarrow 0$, we get in one dimension,

$$TV(f) = \int_a^b |f'(x)| dx$$

- In higher dimensions, that is $\Omega = [a, b]^n$ The formula generalizes to

$$TV(f) = \int_{\Omega} |\nabla f(x)| dx$$

Variational Forms

Variational problems are of the form:

$$T_{\alpha}(m) = \operatorname{argmin}_f \left\{ \frac{1}{2} \|Af - m\|_2^2 + \alpha \mathcal{R}(f) \right\}$$

where $\mathcal{R}(f)$ includes *a priori* information on the unknown f .

- Tikhnov regularization: $\mathcal{R}(f) = \|f\|_2^2$
- Generalized Tikhnov regularization: $\mathcal{R}(f) = \|Lf\|_2^2$
- Total variation regularization: $\mathcal{R}(f) = \|Lf\|_1$

Variational Forms

Variational problems are of the form:

$$T_\alpha(m) = \operatorname{argmin}_f \left\{ \frac{1}{2} \|Af - m\|_2^2 + \alpha \mathcal{R}(f) \right\}$$

where $\mathcal{R}(f)$ includes *a priori* information on the unknown f .

- Tikhnov regularization: $\mathcal{R}(f) = \|f\|_2^2$
- Generalized Tikhnov regularization: $\mathcal{R}(f) = \|Lf\|_2^2$
- Total variation regularization: $\mathcal{R}(f) = \|Lf\|_1$

Variational Forms

Variational problems are of the form:

$$T_\alpha(m) = \operatorname{argmin}_f \left\{ \frac{1}{2} \|Af - m\|_2^2 + \alpha \mathcal{R}(f) \right\}$$

where $\mathcal{R}(f)$ includes *a priori* information on the unknown f .

- Tikhnov regularization: $\mathcal{R}(f) = \|f\|_2^2$
- Generalized Tikhnov regularization: $\mathcal{R}(f) = \|Lf\|_2^2$
- Total variation regularization: $\mathcal{R}(f) = \|Lf\|_1$

Variational Forms

Variational problems are of the form:

$$T_\alpha(m) = \operatorname{argmin}_f \left\{ \frac{1}{2} \|Af - m\|_2^2 + \alpha \mathcal{R}(f) \right\}$$

where $\mathcal{R}(f)$ includes *a priori* information on the unknown f .

- Tikhnov regularization: $\mathcal{R}(f) = \|f\|_2^2$
- Generalized Tikhnov regularization: $\mathcal{R}(f) = \|Lf\|_2^2$
- Total variation regularization: $\mathcal{R}(f) = \|Lf\|_1$

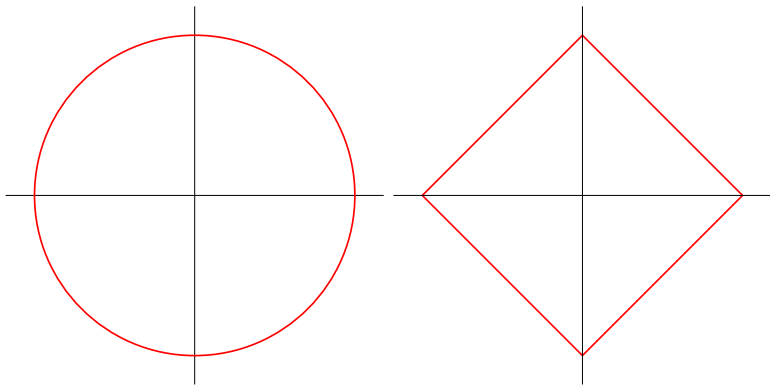
A brief history of l_p norms

Let $f \in \mathbb{R}^n$. The l_p norms for $1 \leq p < \infty$ is defined by

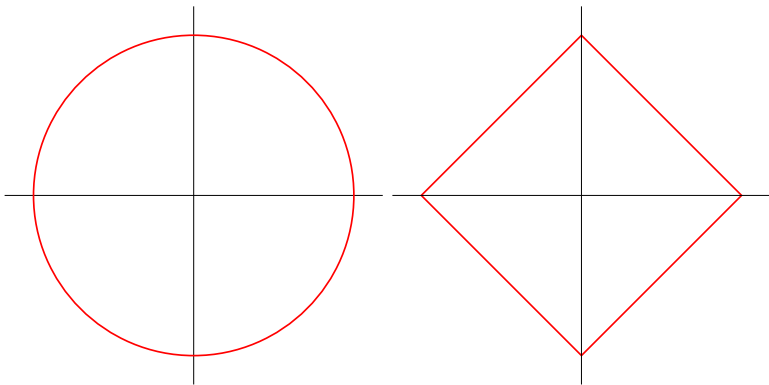
$$\|f\|_p = \left(\sum_{j=1}^n |f_j|^p \right)^{1/p}$$

In particular,

$$\|f\|_2^2 = \sum_{j=1}^n |f_j|^2, \quad \|f\|_1 = \sum_{j=1}^n |f_j|$$



- By constraining our least squares minimization problem to minimize the total variation of the solution, we expect the solution to have a fewer oscillations.
- We expect the solution to be more of a block as there is no minimizing the function amplitude or derivative directly.



- By constraining our least squares minimization problem to minimize the total variation of the solution, we expect the solution to have a fewer oscillations.
- We expect the solution to be more of a block as there is no minimizing the function amplitude or derivative directly.

Outline

- Introduction
 - Motivation
 - Problem Statement
 - Scope and Objectives
- Literature Review
 - Existing Research
 - Gaps in Knowledge
- Methodology
 - Research Design
 - Data Collection
 - Analysis Techniques
- Results and Discussion
 - Key Findings
 - Interpretation of Results
- Conclusion
 - Summary of Findings
 - Implications and Future Work
- References
- Appendix
 - Additional Data
 - Supplementary Figures

What does Total Variation do?

Total Variation regularization can be understood as a balance between two requirements

- $T_\alpha(m)$ should give a small residual $AT_\alpha(m) - m$.
- $LT_\alpha(m)$ should be small in l_1 norm.

TV regularization

Total Variation (TV) regularization promotes sparsity in the derivative favouring piecewise constant functions.

What does Total Variation do?

Total Variation regularization can be understood as a balance between two requirements

- $T_\alpha(m)$ should give a small residual $AT_\alpha(m) - m$.
- $LT_\alpha(m)$ should be small in l_1 norm.

TV regularization

Total Variation (TV) regularization promotes sparsity in the derivative favouring piecewise constant functions.

What does Total Variation do?

Total Variation regularization can be understood as a balance between two requirements

- $T_\alpha(m)$ should give a small residual $AT_\alpha(m) - m$.
- $LT_\alpha(m)$ should be small in l_1 norm.

TV regularization

Total Variation (TV) regularization promotes sparsity in the derivative favouring piecewise constant functions.

What does Total Variation do?

Total Variation regularization can be understood as a balance between two requirements

- $T_\alpha(m)$ should give a small residual $AT_\alpha(m) - m$.
- $LT_\alpha(m)$ should be small in l_1 norm.

TV regularization

Total Variation (TV) regularization promotes sparsity in the derivative favouring piecewise constant functions.

Outline

- 1. Introduction
 - 1.1. Background and Motivation
 - 1.2. Scope and Objectives
- 2. Literature Review
 - 2.1. Existing Research
 - 2.2. Gaps and Contributions
- 3. Methodology
 - 3.1. Research Design
 - 3.2. Data Collection
 - 3.3. Analysis Techniques
- 4. Results and Discussion
 - 4.1. Findings
 - 4.2. Interpretation
- 5. Conclusion
 - 5.1. Summary
 - 5.2. Implications
 - 5.3. Future Research

Example1

- Consider a piecewise continuous function

$$h(x) = \begin{cases} 0 & 0 \leq x \leq 1/2 \\ 2 & 1/2 < x \leq 1 \end{cases}$$

- Choose $n = 10$ and discretize the function at the points $x_1 = 0, x_2 = \frac{1}{10}, \dots, x_{11} = \frac{9}{10}$.
- Define the derivative matrix L
- Compute the l_1 and l_2 norm of $\|Lh\|$

Example 2

- Consider a piecewise continuous function

$$f(x) = 2x$$

- Choose $n = 10$ and discretize the function at the points $x_1 = 0, x_2 = \frac{1}{10}, \dots, x_{11} = \frac{9}{10}$.
- Define the derivative matrix L
- Compute the l_1 and l_2 norm of $\|Lf\|$

Example 3

- Consider a piecewise continuous function

$$f(x) = \sin(2\pi x)$$

- Choose $n = 10$ and discretize the function at the points $x_1 = 0, x_2 = \frac{1}{10}, \dots, x_{11} = \frac{9}{10}$.
- Define the derivative matrix L
- Compute the l_1 and l_2 norm of $\|Lf\|$
- Repeat the experiment by choosing different n values such as $n = 21, 51, 101$

Example 4

- Consider a piecewise continuous function

$$h(x) = \begin{cases} 0 & x = 0 \\ 1 & 0 < x < 0.5 \\ 0 & x = 0.5 \\ -1 & 0.5 < x < 1 \\ 0 & x = 1 \end{cases}$$

- Choose $n = 10$ and discretize the function at the points $x_1 = 0, x_2 = \frac{1}{10}, \dots, x_{11} = \frac{9}{10}$.
- Define the derivative matrix L
- Compute the l_1 and l_2 norm of $\|Lh\|$
- Repeat the experiment by choosing different n values such as $n = 21, 51, 101$