Math 4/896: Seminar in Mathematics
Topic: Inverse Theory

Instructor: Thomas Shores
Department of Mathematics

Lecture 24, April 11, 2006
AvH 10

Outline

Regularization...Sort Of

Basic Idea:

Use prior knowledge about the nature of the solution to restrict it:

- Most common restrictions: on the magnitude of the parameter values. Which leads to the problem:
- Minimize $f(\mathbf{m})$
  subject to $l \leq m \leq u$.
- One could choose $f(\mathbf{m}) = \|G\mathbf{m} - \mathbf{d}\|_2$(BVLS)
- One could choose $f(\mathbf{m}) = \mathbf{c}^T \cdot \mathbf{m}$with additional constraint $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$.

Regularization...Sort Of

Basic Idea:

Use prior knowledge about the nature of the solution to restrict it:

- Most common restrictions: on the magnitude of the parameter values. Which leads to the problem:
- Minimize $f(\mathbf{m})$
  subject to $l \leq m \leq u$.
- One could choose $f(\mathbf{m}) = \|G\mathbf{m} - \mathbf{d}\|_2$(BVLS)
- One could choose $f(\mathbf{m}) = \mathbf{c}^T \cdot \mathbf{m}$with additional constraint $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$.

Regularization...Sort Of

Basic Idea:

Use prior knowledge about the nature of the solution to restrict it:

- Most common restrictions: on the magnitude of the parameter values. Which leads to the problem:
- Minimize $f(\mathbf{m})$
  subject to $l \leq m \leq u$.
- One could choose $f(\mathbf{m}) = \|G\mathbf{m} - \mathbf{d}\|_2$(BVLS)
- One could choose $f(\mathbf{m}) = \mathbf{c}^T \cdot \mathbf{m}$with additional constraint $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$.

Regularization...Sort Of

Basic Idea:

Use prior knowledge about the nature of the solution to restrict it:

- Most common restrictions: on the magnitude of the parameter values. Which leads to the problem:
- Minimize $f(\mathbf{m})$
  subject to $l \leq m \leq u$.
- One could choose $f(\mathbf{m}) = \|G\mathbf{m} - \mathbf{d}\|_2$(BVLS)
- One could choose $f(\mathbf{m}) = \mathbf{c}^T \cdot \mathbf{m}$with additional constraint $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$.

Regularization...Sort Of

Basic Idea:

Use prior knowledge about the nature of the solution to restrict it:

- Most common restrictions: on the magnitude of the parameter values. Which leads to the problem:
- Minimize $f(\mathbf{m})$
  subject to $l \leq m \leq u$.
- One could choose $f(\mathbf{m}) = \|G\mathbf{m} - \mathbf{d}\|_2$(BVLS)
- One could choose $f(\mathbf{m}) = \mathbf{c}^T \cdot \mathbf{m}$with additional constraint $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$.

Example 3.3

Contaminant Transport

Let $C(x,t)$ be the concentration of a pollutant at point x in a linear stream, time t, where $0 \leq x < \infty$ and $0 \leq t \leq T$. The defining model

$$
\begin{aligned}
\frac{\partial C}{\partial t} &= D\frac{\partial^2 C}{\partial x^2} - v\frac{\partial C}{\partial x} \\
C(0,t) &= C_{in}(t) \\
C(x,t) &\to 0, \quad x \to \infty \\
C(x,0) &= C_0(x)
\end{aligned}
$$

Solution

Solution:

In the case that $C_0(x) \equiv 0$, the explicit solution is

$$C(x, T) = \int_0^T C_{in}(t)\, f(x, T - t)\, dt,$$

where

$$f(x, \tau) = \frac{x}{2\sqrt{\pi D \tau^3}} e^{-(x - v\tau)^2/(4D\tau)}$$

Inverse Problem

Problem:

Given simultaneous measurements at time T, to estimate the contaminant inflow history. That is, given data

$$d_i = C\left(x_i, T\right), \; i = 1, 2, \ldots, m,$$

to estimate

$$C_{in}\left(t\right), \; 0 \leq t \leq T.$$

Change the startupfile path to Examples/chap7/examp1 execute it and examp.

Outline

A Better Idea (?)

Entropy:

$$E(\mathbf{m}) = -\sum_{j=1}^{n} m_j \ln(w_j m_j),\ \mathbf{w} \text{ a vector of positive weights.}$$

- Motivated by Shannon's information theory and Bolzmann's theory of entropy in statistical mechanics. A measure of uncertainty about which message or physical state will occur.
- Shannon's entropy function for a probability distribution $\{p_i\}_{i=1}^n$ is $H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \ln(p_i)$.
- Bayesian Maximum Entropy Principle: least biased model is one that maximizes entropy subject to constraints of testable information like bounds or average values of parameters.

A Better Idea (?)

Entropy:

$E(\mathbf{m}) = -\sum_{j=1}^{n} m_j \ln(w_j m_j)$, $\mathbf{w}$ a vector of positive weights.

- Motivated by Shannon's information theory and Bolzmann's theory of entropy in statistical mechanics. A measure of uncertainty about which message or physical state will occur.
- Shannon's entropy function for a probability distribution $\{p_i\}_{i=1}^{n}$ is $H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \ln(p_i)$.
- Bayesian Maximimum Entropy Principle: least biased model is one that maximizes entropy subject to constraints of testable information like bounds or average values of parameters.

A Better Idea (?)

Entropy:

$E(\mathbf{m}) = -\sum_{j=1}^{n} m_j \ln(w_j m_j)$, $\mathbf{w}$ a vector of positive weights.

- Motivated by Shannon's information theory and Bolzmann's theory of entropy in statistical mechanics. A measure of uncertainty about which message or physical state will occur.
- Shannon's entropy function for a probability distribution $\{p_i\}_{i=1}^{n}$ is $H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \ln(p_i)$.
- Bayesian Maximimum Entropy Principle: least biased model is one that maximizes entropy subject to constraints of testable information like bounds or average values of parameters.

A Better Idea (?)

Entropy:

$E(\mathbf{m}) = -\sum_{j=1}^{n} m_j \ln(w_j m_j)$, $\mathbf{w}$ a vector of positive weights.

- Motivated by Shannon's information theory and Bolzmann's theory of entropy in statistical mechanics. A measure of uncertainty about which message or physical state will occur.
- Shannon's entropy function for a probability distribution $\{p_i\}_{i=1}^{n}$ is $H(\mathbf{p}) = -\sum_{i=1}^{n} p_i \ln(p_i)$.
- Bayesian Maximimum Entropy Principle: least biased model is one that maximizes entropy subject to constraints of testable information like bounds or average values of parameters.

Maximum Entropy Regularization

Maximize Entropy:

That is, our version. So problem looks like:

- Maximize $-\sum_{j=1}^{n} m_j \ln(w_j m_j)$
- Subject to $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$ and $\mathbf{m} \geq \mathbf{0}$.
- In absence of extra information, take $w_i = 1$. Lagrange multipliers give:
- Minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \sum_{j=1}^{n} m_j \ln(w_j m_j)$,
- subject to $\mathbf{m} \geq \mathbf{0}$.

Change the startupfile path to Examples/chap7/examp2 execute it and examp.

Maximum Entropy Regularization

Maximize Entropy:

That is, our version. So problem looks like:

- Maximize $-\sum_{j=1}^{n} m_j \ln (w_j m_j)$
- Subject to $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$ and $\mathbf{m} \geq \mathbf{0}$.
- In absence of extra information, take $w_i = 1$. Lagrange multipliers give:
- Minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \sum_{j=1}^{n} m_j \ln (w_j m_j)$,
- subject to $\mathbf{m} \geq \mathbf{0}$.

Change the startupfile path to Examples/chap7/examp2 execute it and examp.

Maximize Entropy:

That is, our version. So problem looks like:

- Maximize $-\sum_{j=1}^{n} m_j \ln(w_j m_j)$
- Subject to $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$ and $\mathbf{m} \geq \mathbf{0}$.
- In absence of extra information, take $w_i = 1$. Lagrange multipliers give:
- Minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \sum_{j=1}^{n} m_j \ln(w_j m_j)$,
- subject to $\mathbf{m} \geq \mathbf{0}$.

Change the startupfile path to Examples/chap7/examp2 execute it and examp.

Maximum Entropy Regularization

Maximize Entropy:

That is, our version. So problem looks like:

- Maximize $-\sum_{j=1}^{n} m_j \ln\left(w_j m_j\right)$
- Subject to $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$ and $\mathbf{m} \geq \mathbf{0}$.
- In absence of extra information, take $w_i = 1$. Lagrange multipliers give:
- Minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \sum_{j=1}^{n} m_j \ln\left(w_j m_j\right)$,
- subject to $\mathbf{m} \geq \mathbf{0}$.

Change the startupfile path to Examples/chap7/examp2 execute it and examp.

Maximum Entropy Regularization

Maximize Entropy:

That is, our version. So problem looks like:

- Maximize $-\sum_{j=1}^{n} m_j \ln(w_j m_j)$
- Subject to $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$ and $\mathbf{m} \geq \mathbf{0}$.
- In absence of extra information, take $w_i = 1$. Lagrange multipliers give:
- Minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \sum_{j=1}^{n} m_j \ln(w_j m_j)$,
- subject to $\mathbf{m} \geq \mathbf{0}$.

Change the startupfile path to Examples/chap7/examp2 execute it and examp.

Maximum Entropy Regularization

Maximize Entropy:

That is, our version. So problem looks like:

- Maximize $-\sum_{j=1}^{n} m_j \ln(w_j m_j)$
- Subject to $\|G\mathbf{m} - \mathbf{d}\|_2 \leq \delta$ and $\mathbf{m} \geq \mathbf{0}$.
- In absence of extra information, take $w_i = 1$. Lagrange multipliers give:
- Minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \sum_{j=1}^{n} m_j \ln(w_j m_j)$,
- subject to $\mathbf{m} \geq \mathbf{0}$.

Change the startupfile path to Examples/chap7/examp2 execute it and examp.

Outline

TV Regularization

We only consider total variation regularization from this section.

Regularization term:

$\mathrm{DV}(\mathbf{m}) = \sum_{j=1}^{n-1} |m_{j+1} - m_j| = \|L\mathbf{m}\|_1$, where L is the matrix used in first order Tikhonov regularization.

- Problem becomes: minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha \|\mathbf{m}\|_1$
- Better yet: minimize $\|G\mathbf{m} - \mathbf{d}\|_1 + \alpha \|\mathbf{m}\|_1$.
- Equivalently: minimize $\left\| \begin{bmatrix} G \\ \alpha L \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_1$.
- Now just use IRLS (iteratively reweighted least squares) to solve it and an L-curve of sorts to find optimal α.
-

TV Regularization

We only consider total variation regularization from this section.

Regularization term:

$$\mathrm{DV}(\mathbf{m}) = \sum_{j=1}^{n-1} |m_{j+1} - m_j| = \|L\mathbf{m}\|_1,$$ where L is the matrix used in first order Tikhonov regularization.

- Problem becomes: minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha \|\mathbf{m}\|_1$
- Better yet: minimize $\|G\mathbf{m} - \mathbf{d}\|_1 + \alpha \|\mathbf{m}\|_1$.
- Equivalently: minimize $\left\| \begin{bmatrix} G \\ \alpha L \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_1$.
- Now just use IRLS (iteratively reweighted least squares) to solve it and an L-curve of sorts to find optimal α.
-

TV Regularization

We only consider total variation regularization from this section.

Regularization term:

$\mathrm{DV}(\mathbf{m}) = \sum_{j=1}^{n-1} |m_{j+1} - m_j| = \|L\mathbf{m}\|_1$, where L is the matrix used in first order Tikhonov regularization.

- Problem becomes: minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha \|\mathbf{m}\|_1$
- Better yet: minimize $\|G\mathbf{m} - \mathbf{d}\|_1 + \alpha \|\mathbf{m}\|_1$.
- Equivalently: minimize $\left\| \begin{bmatrix} G \\ \alpha L \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_1$.
- Now just use IRLS (iteratively reweighted least squares) to solve it and an L-curve of sorts to find optimal α.
-

TV Regularization

We only consider total variation regularization from this section.

Regularization term:

$\mathrm{DV}(\mathbf{m}) = \sum_{j=1}^{n-1} |m_{j+1} - m_j| = \|L\mathbf{m}\|_1$, where L is the matrix used in first order Tikhonov regularization.

- Problem becomes: minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha \|\mathbf{m}\|_1$
- Better yet: minimize $\|G\mathbf{m} - \mathbf{d}\|_1 + \alpha \|\mathbf{m}\|_1$.
- Equivalently: minimize $\left\| \begin{bmatrix} G \\ \alpha L \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_1$.
- Now just use IRLS (iteratively reweighted least squares) to solve it and an L-curve of sorts to find optimal α.

TV Regularization

We only consider total variation regularization from this section.

Regularization term:

$$\mathrm{DV}(\mathbf{m}) = \sum_{j=1}^{n-1} |m_{j+1} - m_j| = \|L\mathbf{m}\|_1,$$ where L is the matrix used in first order Tikhonov regularization.

- Problem becomes: minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha \|\mathbf{m}\|_1$
- Better yet: minimize $\|G\mathbf{m} - \mathbf{d}\|_1 + \alpha \|\mathbf{m}\|_1$.
- Equivalently: minimize $\left\| \begin{bmatrix} G \\ \alpha L \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_1$.
- Now just use IRLS (iteratively reweighted least squares) to solve it and an L-curve of sorts to find optimal α.

TV Regularization

We only consider total variation regularization from this section.

Regularization term:

$\mathrm{DV}(\mathbf{m}) = \sum_{j=1}^{n-1} |m_{j+1} - m_j| = \|L\mathbf{m}\|_1$, where L is the matrix used in first order Tikhonov regularization.

- Problem becomes: minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha \|\mathbf{m}\|_1$
- Better yet: minimize $\|G\mathbf{m} - \mathbf{d}\|_1 + \alpha \|\mathbf{m}\|_1$.
- Equivalently: minimize $\left\| \begin{bmatrix} G \\ \alpha L \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_1$.
- Now just use IRLS (iteratively reweighted least squares) to solve it and an L-curve of sorts to find optimal α.

Total Variation

Key Property:

- TV doesn't smooth discontinuities as much as Tikhonov regularization.

Change startupfile path to Examples/chap7/examp3 execute it and examp.

Total Variation

Key Property:

- TV doesn't smooth discontinuities as much as Tikhonov regularization.

Change startupfile path to Examples/chap7/examp3 execute it and examp.

Outline

Basic Problem:

Solve the system of equations represented in vector form as

$$\mathbf{F}(\mathbf{x}) = \mathbf{0}.$$

- Here $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), \ldots, F_m(\mathbf{x}))$ and $x = (x_1, \ldots, x_m)$.

Basic Problem:

Solve the system of equations represented in vector form as

$$\mathbf{F}(\mathbf{x}) = \mathbf{0}.$$

- Here $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), \ldots, F_m(\mathbf{x}))$ and $x = (x_1, \ldots, x_m)$.