Inverse Problem Study in One Order ODEs

Estimation Parameter Functions in Linear Problem

Wenchao. Zhang Qikun. Wu Jiale. Wang Tao. Luo

South University of Science and Technology of China

Second Project's Report Presentations, Nov.30th 2012

Outline

Outline

Work List

1. To know about inverse problem
2. Make basic concepts in IPs clear
3. Study regularization method for ill-posed problems
4. Discretizing continuous inverse problems for computing
5. Applying Regularization in our problems
6. Introduce some iterative methods
7. Introduce interpolation methods for approximation
8. Interpolation approach to specific parameter estimation problems by MATLAB
9. $P(x)$ estimation by using MATLAB interpolation method

Outline

Questions in Inverse Problem

- According to the output can we determine at least a model (existence)?
- How can we obtain parameters from such model (approximation)?
- Is the model existing the only one model (uniqueness)?
- Can the model endure little error of the output (stability)?

Model Space and Data Space

A problem system can be written as following equation

$$d = G(m) \tag{1}$$

$d = f(m, t)$, usually is a collection of discrete observations. G is operator.
Problem system with errors:

$$d = G(m_{true}) + \eta = d_{true} + \eta$$

Linear Systems

Definition

(Linear System)Let d and m be connected the equation (1),a linear system $G(m)$ satisfies following two conditions:

1. $G(m_1 + m_2) = G(m_1) + G(m_2)$ for all m_1 and $m_2 \in M$
2. $G(\alpha m) = \alpha G(m)$

Fredholm Integral Equations of the First Kind

In a continuous linear inverse problem, G can be expressed as a linear integral operator

$$\int_a^b g(s,x)m(x)dx = d(s) \tag{2}$$

$g(s,x)$ is called the kernel.

Well-posed Problems

Hadamard first raised this idea:

Definition

Let M and D be normed spaces, $G : M \to D$ a mapping. The equation $Gm = d$ is called properly posed or well-posed if:

1. Existence. For every $d \in D$ there exist $m \in M$ such that $Gm = d$.
2. Uniqueness. For every $d \in D$ there is only one $m \in M$ such that $Gm = d$.
3. Stability. The solution depends continuously on d, i.e., for every sequence $(m_n) \subset M$ with $Gm_n \to Gm$, it follows that $x_n \to x (n \to \infty)$

Equations which cannot satisfy all three properties are called *ill-posed*.

Idea of Regularization

- To solve ill-posed problems
- Imposing additional constraints that bias the solution
- Regularization is frequently essential to producing a usable solution.

Outline

Quadrature Methods

We first assume that $d(s)$ is known at a finite number of points $s_1, s_2, ..., s_m$.We can write the inverse problem as

$$d_i = d(s_i) = \int_a^b g(s_i, x)m(x)dx = \int_a^b g_i(x)m(x)dx \qquad (3)$$

The simplest quadrature rule is the midpoint rule,where we divide the interval $[a, b]$ into n subintervals and pick points $x_1, x_2, ..., x_n$ in the middle of each subinterval. The points are given by

$$x_j = a + \frac{\Delta x}{2} + (j-1)\Delta x, \; where \, \Delta x = \frac{b-a}{n}$$

Thus,we have an approximate formula:

$$d_i \approx \sum_{j=1}^{n} g_i(x_j)m(x_j)\Delta x, i = 1, 2, 3, ..., m \qquad (4)$$

Expansion in Terms of Representers

In the Gram matrix technique for discretizing a linear inverse problem, a continuous model $m(x)$ is written as:

$$m(x) = \sum_{j=1}^{m} \alpha_j g_j(x) \tag{5}$$

Then substitute this equation into equation (3), we will get

$$d(s_i) = \int_a^b g_i(x) \sum_{j=1}^{m} \alpha_j g_j(x)\, dx = \sum_{j=1}^{m} \alpha_j \int_a^b g_i(x)\, g_j(x)\, dx \tag{6}$$

$for\ i = 1, 2, \ldots, m$

Expansion in Terms of Representers

Gram matrix with elements are defined as following:

$$\Gamma_{i,j} = \int_a^b g_i(x) g_j(x) dx, \text{ for } i,j = 1,2,...,m$$

So we can discretize IFKs:

$$\Gamma \alpha = d$$

Actually,we usually use the expansion in terms of orthonormal bases function to discretize $m(x)$.In that case $||m(x)||_2 = ||\alpha||_2$.

Outline

Singular Value Decomposition

In the SVD an m by n matrix G is factored into

$$G = USV^T \tag{7}$$

where

1. U is an m by m orthogonal matrix with columns that are unit basis vectors spanning the data space, R^m.
2. V is an n by n orthogonal matrix with columns that are basis vectors spanning the model space, R^n.
3. S is an m by n diagonal matrix with nonnegative diagonal elements called singular values.(sqrt. of eigenvalues of $G^\star G$).

The SVD can be used to compute a generalized inverse of G, called the *Moore–Penrose pseudoinverse*($G^\dagger = V_p S_p^{-1} U_p^T$).

Outline

Origins of Tikhonov Regularization

In Tikhonov regularization,we consider all solutions with $||Gm-d||_2 \leq \delta$,and select the one that minimizes the norm of m:

$$||Gm-d||_2 \leqslant \delta$$

$$min\|m\|_2$$

It can be shown[1] that for appropriate choices of δ and α,above method can be equal to the following one:

$$\min\left(J_\alpha\left(m\right)\right)=\min(||Gm-d||_2^2+\alpha^2\left\|m\right\|_2^2) \tag{8}$$

[1] Karush-Kuhn-Tucker Conditions

Tikhonov Regularization

We have know that G^*G has nonnegative eigenvalues(by the def. of singular vaule), $G^*G + \alpha^2 I (\alpha \neq 0)$ have all positive eigenvalues. The we transform the ill-posed problem into a well-posed problem.

$$(G^*G + \alpha^2 I) m_\alpha = G^* d \tag{9}$$

The following Thm. states that this solution fixes in condition (8):

Theorem

Let $G : D \to M$ be a linear and bounded operator between Hilbert spaces and $\alpha \neq 0$. Then the Tikhonov functional J_α has a unique minimum $m_\alpha \in M$. This minimum m_α is the unique solution of the normal equation (9).

Determination of Regularization Parameter alpha

- Discrepancy Principle of Morozov: $||Gm_\alpha^\delta - d^\delta||_2 = \delta$;
- Engl's Criterien:Find an α which make the following formula minimum:

$$\varphi(\alpha) = \frac{||Gm_\alpha^\delta - d^\delta||_2}{\alpha}$$

- Arcangeli Criterien: $||Gm_\alpha^\delta - d^\delta||_2 = \frac{\delta}{\alpha}$;
- Hanke's L-Curves

Outline

Iterative Methods

SVD and Tikhonov regularization solutions become impractical when we consider larger problems in which G has thousands of rows and columns.For these problems Iterative Methods become efficient.

- Kaczmarz's Algorithm
- Algebraic Reconstruction Technique
- Selective Internal Radiation Therapy
- The Conjugate Gradient Method
- Conjugate Gradient Least Squares Method

Outline

Linear Interpolation

we can use some functions called interpolant to estimate the y value, and this method is called interpolation.
this method creates linear function connecting two points $(x_{j-1}, y_{j-1});(x_j, y_j)$

$$y = y_{j-1} + \frac{(x - x_{j-1})(y_j - y_{j-1})}{(x_j - x_{j-1})}$$

for any x inbetween, we can get corresponding y through this linear function.

Polynomial interpolation

Assume we have n points, there exists unique nth order polynomial that can connect these n points, this nth order polynomial is called Polynomial interpolant .
Lagrange polynomial (one method to get this nth order polynomial):

$$P(x) = \sum_{j=1}^{n} \prod_{i=1, i \neq j}^{n} \frac{x - x_j}{x_i - x_j} y_i$$

Cubic Spline interpolation

Assume we have n+1 points,(x_0, y_0),...(x_n, y_n).
let $S_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i$ be the CS function between $[x_i, x_{i+1})$. Since for

- $\forall S_i(x)\ S(x_i) = y_i, S(x_{i+1}) = y_{i+1}$ (2n linear equations),
- $\forall x_i\ S'_{i-1}(x_i) = S'_i(x_i) i = 1, 2, ..., n-1$ (n-1 linear equations),
- $\forall x_i\ S''_{i-1}(x_i) = S''_i(x_i), i = 1, 2, ..., n-1$(n-1 linear equations).

two dimesion of freedom.
the normal process is setting $S''_0(x_0) = 0, S''_{n-1}(x_{n-1}) = 0$,s.t. we eliminate the rest dimesion of freedom.
solve the reverse of$(n-1) \times (n-1)$ matrix.

Outline

Given P(x), Find Q(x),and without Input Errors

Lete $e^{\int P(x)dx} = m(x)$,then for the linear first order ode,we have

$$m(x)\,y(x) = \int_{x_0}^{x} m(x)\,Q(x)\,dx \tag{10}$$

Equation (10) is an *Volterra Integral Equations of the First Kind*. We can change it as an IFKs.[2]Let

$$\hat{m}(x, x_i) = \begin{cases} m(x) & x \leqslant x_i \\ 0 & x > x_i \end{cases}$$

Then,we have

$$m(x_i)\,y(x_i) = \int_{x_0}^{x_n} \hat{m}(x, x_i)\,Q(x)\,dx \tag{11}$$

[2] Many thanks for prof. Li's help

Given P(x), Find Q(x),and without Input Errors

we use quadrature methods to discretize above system:

$$m(x_i)\,y(x_i) = \sum_{j=1}^{n} \hat{m}(x_j, x_i)\, Q(x_j)\, \Delta x, i = 1, 2, \ldots, n$$

As we can see, we can find that operator G is non-sigular,and have a inverse,for there's no error,so it is well-posed.In fact,

Figure:

$$\begin{pmatrix} z(x_1) \\ z(x_2) \\ .. \\ .. \\ .. \\ z(x_n) \end{pmatrix} = \begin{pmatrix} m(x_1) & 0 & .. & .. & .. & 0 \\ m(x_1) & m(x_2) & 0 & .. & .. & .. \\ .. & .. & .. & 0 & .. & .. \\ .. & .. & .. & m(x_i) & 0 & .. \\ .. & .. & .. & .. & .. & 0 \\ m(x_1) & m(x_2) & .. & .. & m(x_{n-1}) & m(x_n) \end{pmatrix} \begin{pmatrix} Q(x_1) \\ Q(x_2) \\ .. \\ .. \\ .. \\ Q(x_n) \end{pmatrix}$$

where $z(x_i) = \frac{m(x_i)y(x_i)}{\Delta x}$.

Given Q(x), Find P(x),and without Input Errors

Similarly,we can also get

$$m\left(x_i\right) y\left(x_i\right) = \int_{x_0}^{x_n} m\left(x\right) \hat{Q}\left(x_j, x_i\right) dx$$

we use quadrature methods to discretize above system,

$$m\left(x_i\right) y\left(x_i\right) = \sum_{j=1}^{n} \hat{Q}\left(x_j, x_i\right) m\left(x_j\right) \Delta x, i = 1, 2, \ldots, n$$

i.e.

$$\frac{1}{\Delta x} m\left(x_i\right) = \sum_{j=1}^{n} \frac{\hat{Q}\left(x_j, x_i\right)}{y\left(x_i\right)} m\left(x_j\right), i = 1, 2, \ldots, n$$

Let $\frac{\hat{Q}\left(x_j, x_i\right)}{y\left(x_i\right)} = G_{i,j}$,then we have

$$Gm = \frac{1}{\Delta x} m \tag{12}$$

Given Q(x), Find P(x),and without Input Errors

From above equation,we get m is an eigenvector with eigenvalue $\frac{1}{\Delta x}$ of G.It's easy to know that therendifferent eigenvalues of G,but we could only get one solution if we fix Δx.However,Δx is certained before we make our equation,so we make follwing discussion:

- if $\frac{1}{\Delta x}$ is one of the eigenvalues of G,then we can calculate m by it.
- if $\frac{1}{\Delta x}$ is not one of the eigenvalue of G,which is more common.First of all, we choose most closed oneQ_i,that satisfies

$$\min_{i=\{1,2,...,n\}}\left(\frac{1}{\Delta x}-Q_i\right)\stackrel{def}{=}\alpha$$

Now we solve equation

$$(G+\alpha I)\,m_\alpha=\frac{1}{\Delta x}m_\alpha \tag{13}$$

Given P(x), Find Q(x),and Have Input Errors

Similarly with none errors case,we will get

$$y^{\delta}(x_i) = \sum_{j=1}^{n} \frac{\hat{m}(x_j, x_i)\,\Delta x}{m(x_i)} Q(x_j) \tag{14}$$

Let $G_{i,j} = \frac{\hat{m}(x_j,x_i)\Delta x}{m(x_i)}$,then we have

$$y^{\delta} = GQ \tag{15}$$

by Tikhonov Regularization,we have minimum norm solution

$$Q_{\alpha}^{\delta} = \left(G^* G + \alpha^2 I\right)^{-1} G^* y^{\delta} \tag{16}$$

Given Q(x), Find P(x),and Have Input Errors

Sadly, We don't have a proper way in Tikhonov Regularization to solve it, but we have used interpolation method to get a good results.

Outline

Constant Parameter Scenario: Multiple Shooting method

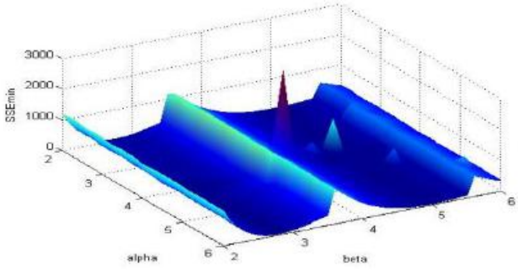
We study a particular differential

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \sin \alpha t + \cos \beta y$$

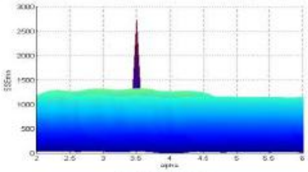
First we study the case that no noise applied,the procedure of solution is:

- determine our dy fuction (destination.m)
- use ODE45 to solve the equation and plot the accurate data(command line)

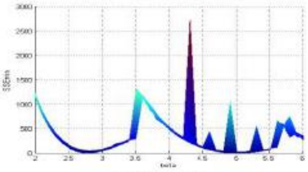
Constant Parameter Scenario: Multiple Shooting method



(a) SSE plot



(b) SSE plot (α)



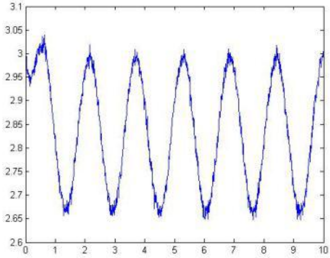
(c) SSE plot (β)

Constant Parameter Scenario: Multiple Shooting method

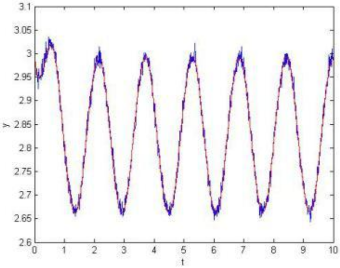
Noises Applied case:

- Add our noise and plot the data again in Figure 2(a) and Figure 2(b):

Figure:



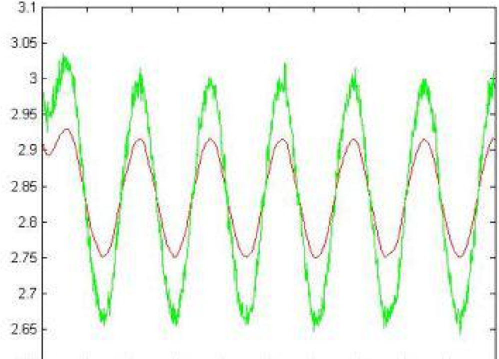
(a) Observed Data



(b) Comparison of data with and without noise

Constant Parameter Scenario: Multiple Shooting method

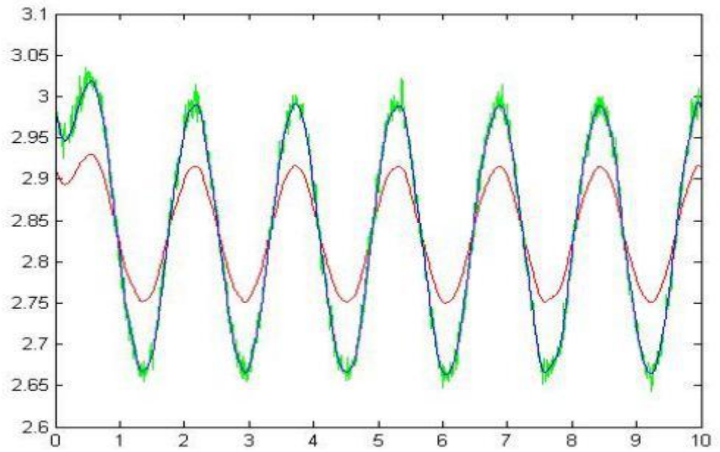
- Eliminate our noise using FFT and IFFT. After FFT and IFFT, the magnitude shrink, so we amplify our data.
- Eliminate high frequency part and do the IFFT. The Figure below, the red is data after IFFT and green is origin.



Constant Parameter Scenario: Multiple Shooting method

- Amplification of our data

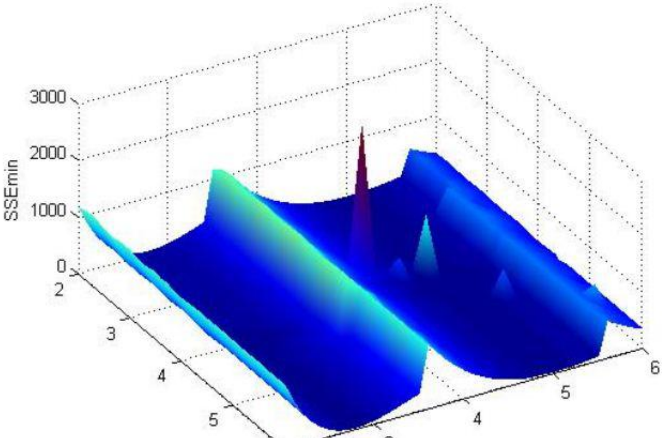
Figure:



Constant Parameter Scenario: Multiple Shooting method

- noise is virtually eliminated.The SSE figures before and after the noise are nearly the same. Repeat fminsearch.

Figure:



Constant Parameter Scenario: Multiple Shooting Method

conclude our answer:

Table 7.1: Conclusion of α and β

	α	β
exact value	4.0000	5.0000
without noise	4.0004	5.0002
with noise	4.0001	4.9987

Figure:

Outline

Continue P(x) fitting

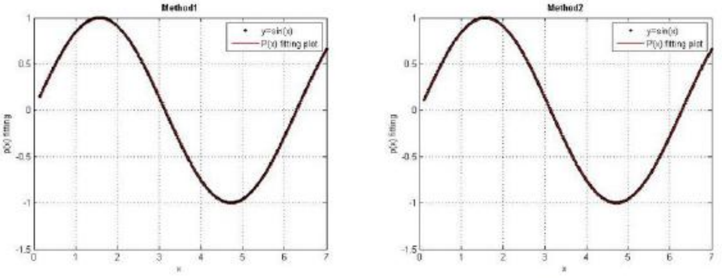
A special case study: $Q(x) = 0$

- determine our P(x) function (expinter.m)
- use spline function to interpolate the original data into a more accurate one
- In order to solve for P(x), the value of dy/dx is obviously necessary.Two ways for do it:
  - one is to get the value of dy/dx every five points (expinter.m)
  - the other one only needs to have one interpolation because it get value of dy/dx every point (expintercubic.m)

Continue P(x) fitting

The results of two methods:

Figure:



(a) Method 1 (b) Method 2

We also do some other interpolation methods,which is omitted here.

Continue P(x) fitting

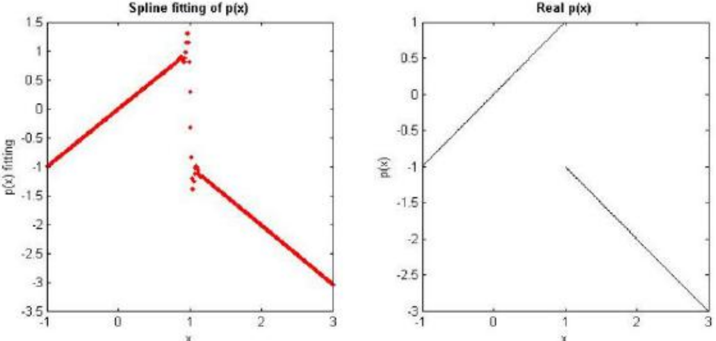
General case: when $Q(x) \neq 0$
For such case it is easy to understand that the question is still easy, because we just need to minus the value of Q(x) from dy/dx , divide the value of y, and then we would get the P(x) at that point.
It's similar to Q(x)=0 case.

Discontinue P(x) fitting

Here also assume Q(x)=0,the procedure is

- determine our P(x) function (uncontinue.m)
- use spline function to interpolate the original data into a more accurate one
- Like before, we need to get the value of dy/dx to solve for P(x), and then plot it out and calculate its SSE.

Figure:



Distribution

- W. Zhang: Discritization,SVD ,Methods of Regularization,Iterative method,Applications of Regularization
- Q. Wu: Multiple Shooting Method with MATLAB,SVD, Redaction
- J. Wang:P(x) Estimation with MATLAB(Cont. and DisCont. case)
- L. Tao:Interpolation Methods