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

Instructor: Thomas Shores
Department of Mathematics

MidTerm Study Notes

Some references:

1. C. Groetsch, *Inverse Problems in the Mathematical Sciences*, Vieweg-Verlag, Braunschweig, Wiesbaden, 1993. (A charmer!)
2. M. Hanke and O. Scherzer, *Inverse Problems Light: Numerical Differentiation*, Amer. Math. Monthly, Vol 108 (2001), 512-521. (Entertaining and gets to the heart of the matter quickly)
3. A. Kirsch, *An Introduction to Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, 1996. (Harder! Definitely a graduate level text)
4. C. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002. (Excellent analysis of computational issues.)
5. A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, 2004. (Very substantial introduction to inverse theory at the graduate level that emphasises statistical concepts.)
6. R. Aster, B. Borchers, C. Thurber, *Estimation and Inverse Problems,* Elsivier, New York, 2005. (And the winner is...)

]

Outline

Contents

1 Brief Introduction to Inverse Theory

1.1 Examples

]

Universal Examples

What are we talking about? A *direct problem* is the sort of thing we traditionally think about in mathematics:

$$\text{Question} \longrightarrow \text{Answer}$$

An *inverse problem* looks like this:

$$\text{Question} \longleftarrow \text{Answer}$$

Actually, this schematic doesn't quite capture the real flavor of inverse problems. It should look more like

$$\text{Question} \longleftarrow \text{(Approximate) Answer}$$

]

Universal Examples

Example 1. (Plato) In the allegory of the cave, unenlightened humans can only see shadows of reality on a dimly lit wall, and from this must reconstruct reality.

Example 2. The game played on TV show "Jeopardy": given the answer, say the question.

]

Math Examples

Matrix Theory: The $m \times n$ *matrix* A, $n \times 1$ *vector* $\mathbf{x}$ *and* $m \times 1$ *vector* $\mathbf{b}$ *satisfy* $A\mathbf{x} = \mathbf{b}$.

- **Direct problem:** given $A, \mathbf{x}$ compute $\mathbf{b}$.
- **Inverse problem:** given $A, \mathbf{b}$, compute $\mathbf{x}$.

Differentiation: given $f(x) \in C[0,1]$ *and* $F(x) = \int_0^x f(t)\,dt$

- **Direct problem:** given $f(x) \in C[0,1]$, find the indefinite integral $F(x)$.
- **Inverse problem:** given $F(0) = 0$ and $F(x) \in C^1[0,1]$, find $f(x) = F'(x)$.

]

Math Examples

Heat Flow in a Rod

Heat flows in a steady state through an insulated inhomogeneous rod with a known heat source and the temperature held at zero at the endpoints. Under modest restrictions, the temperature function $u(x)$ obeys the law

$$-\left(k(x)u'\right)' = f(x), \;\; 0 < x < 1$$

with boundary conditions $u(0) = 0 = u(1)$, thermal conductivity $k(x)$, $0 \leq x \leq 1$ and $f(x)$ determined by the heat source.

Direct Problem: given parameters $k(x), f(x)$, find $u(x) = u(x;k)$.

Inverse Problem: given $f(x)$ and measurement of $u(x)$, find $k(x)$.

1.2 Key Concepts for Inverse Theory

]

Well-Posed Problems

A **well-posed problem** is characterized by three properties:

1. The problem has a solution.
2. The solution is unique.
3. The solution is *stable*, that is, it varies continuously with the given parameters of the problem.

A problem that is not well-posed is called **ill-posed**. In numerical analysis we are frequently cautioned to make sure that a problem is well posed before we design solution algorithms. Another problem with unstable problems: even if exact answers are computable, suppose experimental or numerical error occurs: change in solution could be dramatic!

]

Illustration: a Discrete Inverse Problem

So what's the fuss? The direct problem of computing F from $F = Kf$ is easy and the solution to the inverse problem is $f = K^{-1}F$, right?

Wrong! All of Hadamard's well-posedness requirements fall by the wayside, even for the "simple" inverse problem of solving for $\mathbf{x}$ with $A\mathbf{x} = \mathbf{b}$ a linear system.

1.3 Difficulties and Remedies

]

What Goes Wrong?

1. This linear system $A\mathbf{x} = \mathbf{b}$ has no solution:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

2. This system has infinitely many solutions:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

3. This system has no solution for $\varepsilon \neq 0$ and infinitely many for $\varepsilon = 0$, so solutions do not vary continuously with parameter ε:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \varepsilon \end{bmatrix}$$

]

Some Remedies: Existence

We use an old trick: least squares, which finds the $\mathbf{x}$ that minimizes the size of the residual (squared) $\|\mathbf{b} - A\mathbf{x}\|^2$. This turns out to be equivalent to solving the *normal equations*

$$A^T A\mathbf{x} = A^T\mathbf{b},$$

a system which is guaranteed to have a solution. Further, we can see that if $A\mathbf{x} = \mathbf{b}$ has *any* solution, then every solution to the normal equations is a solution to $A\mathbf{x} = \mathbf{b}$. This trick extends to more abstract linear operators K of equations $Kx = y$ using the concept of "adjoint" operators K^* which play the part of a transpose matrix A^T.

]

Some Remedies: Uniqueness

We "regularize" the problem. We'll illustrate it by one particular kind of regularization, called *Tikhonov* regularization. One introduces a regularization parameter $\alpha > 0$ in such a way that small α give us a problem that is "close" to the original. In the case of the normal equations, one can show that minimizing the modified residual

$$\|\mathbf{b} - A\mathbf{x}\|^2 + \alpha \|\mathbf{x}\|^2$$

leads to the linear system $\left(A^T A + \alpha I\right)\mathbf{x} = A^T\mathbf{b}$, where I is the identity matrix. One can show the coefficient matrix $A^T A + \alpha I$ is always nonsingular. Therefore, the problem has a unique solution.

]

Choice of Regularization Parameter

What should we do about α? This is one of the more fundamental (and intriguing) problems of inverse theory. Let's analyze one of our simple systems for insight, say

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Invariably, our input data for the inverse problem, $(1, 1)$, has error in it, say we have $(1+\delta_1, 1+\delta_2)$ for data instead. Let $\delta = \delta_1 + \delta_2$. The regularized system becomes

$$\begin{bmatrix} 2+\alpha & 2 \\ 2 & 2+\alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2+\delta \\ 2+\delta \end{bmatrix} = (2+\delta) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which has unique solution

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2+\alpha & 2 \\ 2 & 2+\alpha \end{bmatrix}^{-1} (2+\delta) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{2+\delta}{4+\alpha} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

]

Choice of Regularization Parameter

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2+\alpha & 2 \\ 2 & 2+\alpha \end{bmatrix}^{-1} (2+\delta) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{2+\delta}{4+\alpha} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Observe that if the input error δ were 0, all we would have to do is let $\alpha \to 0$ and we would get the valid solution $\frac{1}{2}(1,1)$. But given that the input error is not zero, taking the limit as $\alpha \to 0$ gives us a worse approximation to a solution than we would otherwise get by choosing $\alpha \approx 2\delta$. (Our solutions always satisfy $x_1 = x_2$, so to satisfy $x_1 + x_2 = 1$ we need $x_1 = x_2 = \frac{1}{2}$ or as close as we can get to it.)

There are many questions here, e.g., how do we know in general what the best choice of regularization parameter is, if any? This and other issues are the subject matter of a course in inverse theory.

]

Stability

In this special case, we get stability for free – for each *regularized* problem. We cannot hope to have stability for the unregularized problem $A\mathbf{x} = \mathbf{b}$ since A^{-1} doesn't even exist.

But things are even more complicated: For the general linear problem$Kx = y$, even if K^{-1} is well defined the inverse problem may not be stable (although stability happens in some cases). However, we have to look to infinite dimensional examples such as our differentiation example (operator K is integration), where it can be shown that K^{-1} (differentiation) exists but is not continuous, even though K is.

]

A Continuous Inverse Problem

Let $K : C[0,1] \to C[0,1]$ via the rule $Kf(x) = \int_0^x f(y)\,dy$ This is a one-to-one function. Measure size by the sup norm:

$$\|f\| = \sup_{0 \le x \le 1} |f(x)|$$

so that the "closeness" of $f(x)$ and $g(x)$ is determined by the number $\|f - g\|$. Then one can show that the operator K is continuous in the sense that if $f(x)$ and $g(x)$ are close, then so are $Kf(x)$ and $Kg(x)$.

Let $R = K(C[0,1])$, the range of K. Then $K^{-1} : R \to C[0,1]$ is also one-to-one. But it is not continuous.

]

Failure of Stability

Consider the function

$$g_\varepsilon(x) = \varepsilon \sin\left(\frac{x}{\varepsilon^2}\right),$$

where $\varepsilon > 0$. We have $\|g_\varepsilon\| = \|g_\varepsilon - 0\| \leq \varepsilon$. So for small ε, $g_\varepsilon(x)$ is close to the zero function. Yet,

$$K^{-1}g_\varepsilon(x) = g_\varepsilon(x)' = \frac{\varepsilon}{\varepsilon^2}\cos\left(\frac{x}{\varepsilon^2}\right) = \frac{1}{\varepsilon}\cos\left(\frac{x}{\varepsilon^2}\right)$$

so that $\left\|K^{-1}g_\varepsilon\right\| = \frac{1}{\varepsilon}$, so that $K^{-1}g_\varepsilon$ becomes far from zero as $\varepsilon \to 0$. Hence K^{-1} is not a continuous operator.

]

A General Framework

Forward Problem
consists of

- A model fully specified by (physical) parameters m.
- A known function G that, *ideally*, maps parameters to data d by way of

$$d = G(m).$$

(Pure) Inverse Problem
is to find m given observations d.

We hope (!) that this means to calculate $m = G^{-1}(d)$, but what we really get stuck with is

(Practical) Inverse Problem
is to find m given observations $d = d_{true} + \eta$ so that equation to be inverted is

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

What we are tempted to do is invert the equation

$$d = G(m_{approx})$$

and be happy with m_{approx}. Unfortunately, m_{approx} may be a poor approximation to m_{true}. This makes our job a whole lot tougher – and interesting!

]

Volterra Integral Equations

Our continuous inverse problem example is a special case of this important class of problems:

Definition 3. An equation of the form

$$d(s) = \int_a^s g(s, x, m(x))\, dx$$

is called a *Volterra integral equation of the first kind* (VFK). It is *linear* if

$$g(s, x, m(x)) = g(s, x) \cdot m(x)$$

in which case $g(s, x)$ is the *kernel* of the equation. Otherwise it is a *nonlinear* VFK.

In our example $d(s) = \int_a^s m(x)\, dx$, so $g(s, x) = 1$, $a = 0$.
]

Fredholm Integral Equations of the First Kind (IFK)

Another important class of problems:

Definition 4. An equation of the form

$$d(s) = \int_a^b g(s, x, m(x))\, dx$$

is called a *Fredholm integral equation of the first kind* (IFK). It is *linear* if

$$g(s, x, m(x)) = g(s, x) \cdot m(x)$$

in which case $g(s, x)$ is the *kernel* of the equation. If, further,

$$g(s, x) = g(s - x)$$

the equation is called a *convolution* equation.

Example

Consider our example $d(s) = \int_0^s m(x)\, dx$, again. Define the Heaviside function $H(w)$to be 1 if w is nonnegative and 0 otherwise. Then

$$d(s) = \int_0^s m(x)\, dx = \int_0^\infty H(s - x)\, m(x)\, dx.$$

Thus, this Volterra integral equation can be viewed as a IFK and a convolution equation as well with convolution kernel $g(s, x) = H(s - x)$.
]

Text Example

Gravitational anomaly at ground level due to buried wire mass
where

- Ground level is the x-axis.
- $h(x)$: the depth of the wire at x.
- $\rho(x)$: is the density of the wire at x.
- $d(s)$: measurement of the anomaly at position s, ground level.

This problem leads to linear and (highly) nonlinear inverse problems

2 Chapter 2: Linear Regression

2.1 A Motivating Example

]

The Example

To estimate the mass of a planet of known radius while on the (airless) surface:
Observe a projectile thrown from some point and measure its altitude.

From this we hope to estimate the acceleration a due to gravity and then use Newton's laws of gravitation and motion to obtain from

$$\frac{GMm}{R^2} = ma$$

that

$$M = \frac{aR^2}{G}.$$

Equations of Motion

A Calculus I problem:
What is the vertical position $y(t)$ of the projectile as a function of time?

Just integrate the constant acceleration twice to obtain

$$y(t) = m_1 + m_2 t - \frac{1}{2}at^2.$$

We follow the text and write at time t_k, we have observed value d_k and

$$d_k = y(t_k) = m_1 + t\, m_2 - \frac{1}{2}t^2 m_3$$

where $k = 1, 2, \ldots, m$. This is a system of m equations in 3 unknowns. Here

- m_1 is initial y-displacement
- m_2 is initial velocity
- m_3 is acceleration due to gravity.

Matrix Form

Matrix Form of the System:

(Linear Inverse Problem): $G\mathbf{m} = \mathbf{d}$.

Here $G = \begin{bmatrix} 1 & t_1 & -\frac{t_1^2}{2} \\ 1 & t_2 & \frac{t_2^2}{2} \\ \vdots & \vdots & \vdots \\ 1 & t_m & \frac{t_m^2}{2} \end{bmatrix}$, $\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}$, but we shall examine the problem in the more general setting where G is $m \times n$, $\mathbf{m}$ is $n \times 1$ and $\mathbf{d}$ is $m \times 1$.

A Specific Problem

The exact solution:

$$\mathbf{m} = [10, 100, 9.8]^T = (10, 100, 9.8)\,.$$

Spacial units are meters and time units are seconds. It's easy to simulate an experiment. We will do so assuming an error distribution that is independent and normally distributed with mean $\mu = 0$ and standard deviation $\sigma = 16$.

```
>randn('state',0)
>m = 10
>sigma = 16
>mtrue = [10,100,9.8]'
>G = [ones(m,1), (1:m)', -0.5*(1:m)'.^2]
>datatrue = G*mtrue;
>data = datatrue + sigma*randn(m,1);
```

2.2 Solutions to the System

]

Solution Methods

We could try

- The most naive imaginable: we only need three data points. Let's use them to solve for the three variables. Let's really try it with Matlab and plot the results for the exact data and the simulated data.

- A better idea: We are almost certain to have error. Hence, the full system will be inconsistent, so we try a calculus idea: minimize the sum of the norm of the residuals. This requires development. The basic problem is to find the **least squares solution** $m = m_{L_2}$ such that

$$\text{(Least Squares Problem):} \quad \|\mathbf{d} - G\mathbf{m}_{L_2}\|_2^2 = \min_{\mathbf{m}} \|\mathbf{d} - G\mathbf{m}\|_2^2$$

]

Key Results for Least Squares

Theorem 5. *The least squares problem has a solution for any $m \times n$ matrix G and data $\mathbf{d}$, namely any solution to the* ***normal equations***

$$G^T G\mathbf{m} = G^T\mathbf{d}$$

Proof sketch:

- Show product rule holds for products of matrix functions.
- Note $f(\mathbf{m}) = \|\mathbf{d} - G\mathbf{m}\|^2$ is a nonnegative quadratic function in $\mathbf{m}$, so must have a minimum
- Find the critical points of f by setting $\nabla f(\mathbf{m}) = \mathbf{0}$.

Key Results

Theorem 6. *If $m \times n$ matrix G has full column rank, then the least squares solution is unique, and is given by*

$$\mathbf{m}_{L_2} = \left(G^T G\right)^{-1} G^T\mathbf{d}$$

Proof sketch:

- Show $G^T G$ has zero kernel, hence is invertible.
- Plug into normal equations and solve.

Least Squares Experiments:

Use Matlab to solve our specific problem with experimental data and plot solutions. Then let's see why the theorems are true. There remains:

Problem:
How good is our least squares solution? Can we trust it? Is there a better solution?

2.3 Statistical Aspect of Least Squares

]

Quality of Least Squares

View the ProbStatLectures notes regarding point estimation. Then we see why this fact is true:

Theorem 7. *Suppose that the error of ith coordinate of the residual is normally distributed with mean zero and standard deviation σ_i. Let $W = \mathrm{diag}\,(1/\sigma_1, \ldots, 1/\sigma_m)$ and $G_W = WG$, $\mathbf{d}_W = W\mathbf{d}$. Then the least squares solution to the scaled inverse problem*

$$G_W\mathbf{m} = \mathbf{d}_W$$

*is a **maximum liklihood estimator** to the parameter vector.*

An Example

Let's generate a problem as follows

```
>randn('state',0)
>m = 10
>sigma = blkdiag(8*eye(3),16*eye(3),24*eye(4))
>mtrue = [10,100,9.8]'
>G = [ones(m,1), (1:m)', -0.5*(1:m)'.^2]
>datatrue = G*mtrue;
>data = datatrue + sigma*randn(m,1);
>G = [ones(m,1), (1:m)', -0.5*(1:m)'.^2]
>datatrue = G*mtrue;
>data = datatrue + sigma*randn(m,1);
% compute the least squares solution without
% reference to sigma, then do the scaled least squares
% and compare....also do some graphs
```

Quality of Least Squares

A very nontrivial result which we assume:

Theorem 8. *Let G have full column rank and $\mathbf{m}$ the least squares solution for the scaled inverse problem. The statistic*

$$\|\mathbf{d}_W - G_W\mathbf{m}\|_2^2 = \sum_{i=1}^{m} \left(d_i - (Gm_{L_2})_i\right)^2 / \sigma_i^2$$

in the random variable $\mathbf{d}$ has a chi-square distribution with $\nu = m - n$ degrees of freedom.

This provided us with a statistical assessment (the chi-square test) of the quality of our data. We need the idea of the **p-value** of the test, the probability of obtaining a larger chi-square value than the one actually obtained:

$$p = \int_{\chi^2_{obs}}^{\infty} f_{\chi^2}(x)\, dx.$$

Interpretation of p

As a random variable, the p-value is uniformly distributed between zero and one. This can be very informative:

1. "Normal sized" p: we probably have an acceptable fit
2. Extremely small p: data is very unlikely, so model $G\mathbf{m} = \mathbf{d}$ may be wrong or data may have larger errors than estimated.
3. Extremely large p (i.e., very close to 1): fit to model is almost exact, which may be too good to be true.

]

Uniform Distributions

Reason for uniform distribution:

Theorem 9. *Let X have a continuous c.d.f. $F(x)$ such that $F(x)$is strictly increasing where $0 < x < 1$. Then the r.v. $Y = F(X)$ is uniformly distributed on the interval $(0, 1)$*

Proof sketch:

- Calculate $P(Y \leq y)$ using fact that F has an inverse function F^{-1}.
- Use the fact that $P(X \leq x) = F(x)$ to prove that $P(Y \leq y) = y$.

Application: One can use this to generate random samples for X.

]

An Example

Let's resume our experiment from above. Open the script Lecture8.m and have a look. Then run Matlab on it and resume calculations.

$>$ % now set up for calculating the p-value of the test under both scenarios.

```
>chiobs1 = norm(data - G*mapprox1)^2
>chiobs2 = norm(W*(data - G*mapprox2))^2
>help chis_pdf
>p1 = 1 - chis_cdf(chiobs1,m-n)
>p2 = 1 - chis_cdf(chiobs2,m-n)
% How do we interpret these results?
% Now put the bad estimate to the real test
```

How do we interpret these results?

]

More Conceptual Tools

Examine and use the MVN theorems of ProbStatLectures to compute the expectation and variance of the r.v. $\mathbf{m}$, where $\mathbf{m}$ is the modified least squares solution, G has full column rank and $\mathbf{d}$ is a vector of independent r.v.'s.

- Each entry of $\mathbf{m}$ is a linear combination of independent normally distributed variables, since

$$\mathbf{m} = \left(G_W^T G_W\right)^{-1} G_W^T \mathbf{d}_W.$$

- The weighted data $\mathbf{d}_W = W\mathbf{d}$ has covariance matrix I.
- Deduce that $\mathrm{Cov}(\mathbf{m}) = \left(G_W^T G_W\right)^{-1}$.
- Note simplification if variances are constant: $\mathrm{Cov}(\mathbf{m}) = \sigma^2 \left(G^T G\right)^{-1}$.

Conceptual Tools

Next examine the mean of $\mathbf{m}$ and deduce from the facts that

$$E[\mathbf{d}_W] = W\mathbf{d}_{true} \text{ and } G_W\mathbf{m}_{true} = \mathbf{d}_{true}$$

and MVN facts that

- $E[\mathbf{m}] = \mathbf{m}_{true}$
- Hence, modified least squares solution is an **unbiased estimator** of $\mathbf{m}_{true}$.
- Hence we can construct a confidence interval for our experiment:

$$\mathbf{m} \pm 1.96 \cdot \mathrm{diag}\left(\mathrm{Cov}(\mathbf{m})\right)^{1/2}$$

- What if the (constant) variance is unknown? Student's t to the rescue!

How do we interpret these results?

]

Outliers

These are discordant data, possibly due to other error or simply bad luck. What to do?

- Use statistical estimation to discard the outliers.
- Use a different norm from $\|\cdot\|_2$. The 1-norm is an alternative, but this makes matters much more complicated! Consider the optimization problem

$$\|\mathbf{d} - G\mathbf{m}_{L_2}\|_1 = \min_{\mathbf{m}} \|\mathbf{d} - G\mathbf{m}\|_1$$

How do we interpret these results?

3 Chapter 3: Discretizing Continuous Inverse Problems

3.1 Motivating Example

]

A Motivating Example: Integral Equations

Contanimant 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

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

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

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

The Inverse Problem

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

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

to estimate

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

Some Methods

More generally

Problem:
Given the IFK

$$d(s) = \int_a^b g(x,s)\, m(x)\, dx$$

and a finite sample of values $d(s_i)$, $i = 1, 2, \ldots, m$, to estimate parameter $m(x)$.

Methods we discuss at the board:

1. Quadrature
2. Representers
3. Other Choices of Trial Functions

3.2 Quadrature Methods

]

Quadrature
Basic Ideas:
Approximate the integrals

$$d_i \approx d(s_i) = \int_a^b g(s_i, x)\, m(x)\ dx \equiv \int_a^b g_i(x)\, m(x)\ dx,\ i = 1, 2, \ldots, m$$

(where the representers or data kernels $g_i(x) = g_i(s_i, x)$) by

- Selecting a set of collocation points x_j, $j = 1, 2, \ldots, n$. (It might be wise to ensure $n < m$.)
- Select an integration approximation method based on the collocation points.
- Use the integration approximations to obtain a linear system $G\mathbf{m} = \mathbf{d}$ in terms of the unknowns $m_j \equiv m(x_j)$, $j = 1, 2, \ldots, n$.

3.3 Representer Method

]

Representers

Rather than focusing on the value of m at individual points, take a global view that $m(x)$ lives in a function space which is spanned by the **representer** functions $g_1(x), g_2(x), \ldots, g_n(x), \ldots$

Basic Ideas:

- Make a selection of the basis functions $g_1(x), g_2(x), \ldots, g_n(x)$ to approximate $m(x)$, say
$$m(x) \approx \sum_{j=1}^{n} \alpha_j g_j(x)$$
- Derive a system $\Gamma \mathbf{m} = \mathbf{d}$ with a Gramian coefficient matrix
$$\Gamma_{i,j} = \langle g_i, g_j \rangle = \int_a^b g_i(x)\, g_j(x)\, dx$$

Example

The Most Famous Gramian of Them All:

- Suppose the basis functions turn out to be $g_i(x) = x^{i-1}$, $i = 1, 2, \ldots, m$, on the interval $[0, 1]$.
- Exhibit the infamous Hilbert matrix.

3.4 Generalizations

]

Other Choices of Trial Functions

Take a still more global view that $m(x)$ lives in a function space spanned by a spanning set which may *not* be the representers!

Basic Ideas:

- Make a selection of the basis functions $h_1(x), h_2(x), \ldots, h_n(x)$ with linear span H_n (called "trial functions" in the weighted residual literature) to approximate $m(x)$, say
$$m(x) \approx \sum_{j=1}^{n} \alpha_j h_j(x)$$
- Derive a system $G\alpha = \mathbf{d}$ with a coefficient matrix
$$G_{i,j} = \langle g_i, h_j \rangle = \int_a^b g_i(x)\, h_j(x)\, dx$$

Trial Functions

Orthogonal Idea:
An appealing choice of basis vectors is an orthonormal (o.n.) set of nonzero vectors. If we do so:

- $\|m(x)\| = \sum_{j=1}^{n} \alpha_j^2$
- $\text{Proj}_{H_n}(g_i(x)) = \sum_{j=1}^{n} \langle g_i, h_j \rangle h_j(x)$, $i = 1, \ldots, m$.
- Meaning of ith equation: $\langle \text{Proj}_{H_n}(g_i), m \rangle = d_i$

3.5 Method of Backus and Gilbert

]

Backus-Gilbert Method

Problem: we want to estimate $m(x)$ at a single point $\widehat{x}$ using the available data, and do it well. How to proceed?

Basic Ideas:

- Write $m(\widehat{x}) \approx \widehat{m} = \sum_{j=1}^{m} c_j d_j$ and $d_j = \int_a^b g_j(x) m(x) dx$.
- Reduce the integral conditions to $\widehat{m} = \int_a^b A(x) m(x) dx$ with $A(x) = \sum_{j=1}^{m} c_j g_j(x)$.
- Ideally $A(x) = \delta(x - \widehat{x})$. What's the next best thing?

Backus-Gilbert Equations

Constraints on the averaging kernel $A(x)$:

- First, an area constraint: total area $\int_a^b A(x) dx = 1$. Set $q_j = \int_a^b g_j(x) dx$ and get $\mathbf{q}^T \mathbf{c} = 1$.
- Secondly, minimize second moment $\int_a^b A(x)^2 (x - \widehat{x})^2 dx$.
- This becomes a quadratic programming problem: objective function quadratic and constraints linear.
- In fact, it is convex, i.e., objective function matrix is positive definite. We have a tool for solving this: `quad_prog.m`.

- One could constrain the variance of the estimate $\widehat{m}$, say $\sum_{i=1}^{m} c_i^2 \sigma_i^2 \leq \Delta$, where σ_i is the known variance of d_i. This is a more complicated optimization problem.

]

A Case Study for the EPA

The Problem:
A factory on a river bank has recently been polluting a previously unpolluted river with unaccepable levels of polychlorinated biphenyls (PCBs). We have discovered a plume of PCB and want to estimate its size to assess damage and fines, as well as confirm or deny claims about the amounts by the company owning the factory.

- We control measurements but have an upper bound on the number of samples we can handle, that is, at most 100.
- Measurements may be taken at different times, but at most 20 per time at different locales.
- How would we design a testing procedure that accounts for and reasonably estimates this pollution dumping using the contaminant transport equation as our model?

4 Chapter 4: Rank Deficiency and Ill-Conditioning

4.1 Properties of the SVD

]

Basic Theory of SVD

Theorem 10. *(Singular Value Decomposition) Let G be an $m \times n$ real matrix. Then there exist $m \times m$ orthogonal matrix U, $n \times n$ orthogonal matrix V and $m \times n$ diagonal matrix S with diagonal entries $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_q$, with $q = \min\{m, n\}$, such that $U^T G V = S$. Moreover, numbers $\sigma_1, \sigma_2, \ldots, \sigma_q$ are uniquely determined by G.*

Definition 11. With notation as in the SVD Theorem, and U_p, V_p the matrices consisting of the first p columns of U, V, respectively, and S_p the first p rows and columns of S, where σ_p is the last nonzero singular value, then the **Moore-Penrose pseudoinverse** of G is $G^\dagger = V_p S_p^{-1} U_p^T \equiv \sum_{j=1}^{p} \frac{1}{\sigma_j} \mathbf{V}_j \mathbf{U}_j^T$.

Matlab Knows It

Carry out these calculations in Matlab:

```
> n = 6
> G = hilb(n);
> svd(G)
>[U,S,V] = svd(G);
>U'*G*V - S
>[U,S,V] = svd(G,'econ');
> % try again with n=16 and then G=G(1:8)
> % what are the nonzero singular values of G?
]
```

Applications of the SVD

Use notation above and recall that the null space and column space (range) of matrix G are $N(G) = \{\mathbf{x} \in \mathbb{R}^n \,|\, G\mathbf{x} = \mathbf{0}\}$ and

$$R(G) = \{\mathbf{y} \in \mathbb{R}^m \,|\, \mathbf{y} = G\mathbf{x},\, \mathbf{x} \in \mathbb{R}^n\} = \text{span}\,\{\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_n\}$$

Theorem 12. *(1)* $\text{rank}\,(G) = p$ *and* $G = \sum_{j=1}^{p} \sigma_j \mathbf{U}_j \mathbf{V}_j^T$ *(2)*$N(G) = \text{span}\,\{\mathbf{V}_{p+1}, \mathbf{V}_{p+2}, \ldots, \mathbf{V}_n\}$,$R(G) = \text{span}\,\{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_p\}$ *(3)*$N\left(G^T\right) = \text{span}\,\{\mathbf{U}_{p+1}, \mathbf{U}_{p+2}, \ldots, \mathbf{U}_m\}$,$R(G) = \text{span}\,\{\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_p\}$ *(4)* $\mathbf{m}_\dagger = G^\dagger \mathbf{d}$ *is the least squares solution to* $G\mathbf{m} = \mathbf{d}$ *of minimum* 2*-norm.*

Heart of the Difficulties with Least Squares Solutions

Use the previous notation, so that G is $m \times n$ with rank p and SVD, etc as above. By **data space** we mean the vector space $\mathbb{R}^m$ and by **model space** we mean $\mathbb{R}^n$.

No Rank Deficiency:
This means that $p = m = n$. Comments:

- This means that null space of both G and G^T are trivial (both $\{0\}$).
- Then there is a perfect correspondence between vectors in data space and model space:
$$G\mathbf{m} = \mathbf{d},\ \mathbf{m} = G^{-1}\mathbf{d} = G^\dagger \mathbf{d}.$$
- This is the ideal. But are we out of the woods?
- No, we still have to deal with data error and ill-conditioning of the coefficient matrix (remember Hilbert?).

Heart of the Difficulties with Least Squares Solutions

Use the notation $\mathbf{m}_\dagger = G^\dagger \mathbf{d}$.

Row Rank Deficiency:
This means that $d = n < m$. Comments:

- This means that null space of G is trivial, but that of G^T is not.
- Here $\mathbf{m}_\dagger$ is the unique least squares solution.
- And $\mathbf{m}_\dagger$ is the exact solution to $G\mathbf{m} = \mathbf{d}$ exactly if $\mathbf{d}$ is in the range of G.
- But $\mathbf{m}$ is insensitive to any translation $\mathbf{d} + \mathbf{d}_0$ with $\mathbf{d}_0 \in N\left(G^\dagger\right)$

Heart of the Difficulties with Least Squares Solutions

Column Rank Deficiency:
This means $p = m < n$. Comments:

- This means that null space of G^T is trivial, but that of G is not.
- Here $\mathbf{m}_\dagger$ is a solution of minimum 2-norm.
- And $\mathbf{m}_\dagger + \mathbf{m}_0$ is also a solution to $G\mathbf{m} = \mathbf{d}$ for any $\mathbf{m}_0 \in N(G)$.
- So $\mathbf{d}$ is insensitive to any translation $\mathbf{m}_\dagger + \mathbf{m}_0$ with $\mathbf{m}_0 \in N(G)$.

Heart of the Difficulties with Least Squares Solutions

Row and Column Rank Deficiency:
This means $p < \min\{m, n\}$. Comments:

- This means that null space of both G and G^T are nontrivial.
- Here $\mathbf{m}_\dagger$ is a least squares solution.
- We have trouble in both directions.

4.2 Covariance and Resolution of the Generalized Inverse Solution

Covariance and Resolution

Definition 13. The **model resolution matrix** for the problem $G\mathbf{m} = \mathbf{d}$ is $R_{\mathbf{m}} = G^{\dagger}G$.

Consequences:

- $R_{\mathbf{m}} = V_p V_p^T$, which is just I_n if G has full column rank.
- If $G\mathbf{m_{true}} = \mathbf{d}$, then $E[\mathbf{m}_{\dagger}] = R_m \mathbf{m}_{true}$
- Thus, the bias in the gereralized inverse solution is $E[\mathbf{m}_{\dagger}] - \mathbf{m}_{true} = (R_{\mathbf{m}} - I)\mathbf{m}_{true} = -V_0 V_0^T \mathbf{m}_{true}$ with $V = [V_p V_0]$.
- Similarly, in the case of identically distributed data with variance σ^2, the covariance matrix is $\mathrm{Cov}(\mathbf{m}_{\dagger}) = \sigma^2 G^{\dagger}\left(G^{\dagger}\right)^T = \sigma^2 \sum_{i=1}^{p} \frac{\mathbf{V}_i \mathbf{V}_i^T}{\sigma_i^2}$.
- From expected values we obtain a **resolution test**: if a diagonal entry are close to 1, we claim good resolution of that coordinate, otherwise not.

4.3 Instability of Generalized Inverse Solutions

]

Instability of Generalized Inverse Solution

The key results:

- For $n \times n$ square matrix G $\mathrm{cond}_2(G) = \|G\|_2 \left\|G^{-1}\right\|_2 = \sigma_1/\sigma_n$.
- This inspires the definition: the condition number of an $m \times n$ matrix G is σ_1/σ_q where $q = \min\{m, n\}$.
- Note: if $\sigma_q = 0$, the condition number is infinity. Is this notion useful?
- If data $\mathbf{d}$ vector is perturbed to $\mathbf{d}'$, resulting in a perturbation of the generalized inverse solution $\mathbf{m}_{\dagger}$ to $\mathbf{m}'_{\dagger}$, then $\frac{\left\|\mathbf{m}'_{\dagger} - \mathbf{m}_{\dagger}\right\|_2}{\|\mathbf{m}_{\dagger}\|_2} \leq \mathrm{cond}(G) \frac{\left\|\mathbf{d}' - \mathbf{d}\right\|_2}{\|\mathbf{d}\|_2}$.

Stability Issues

How these facts affect stability:

- If $\text{cond}\,(G)$ is not too large, then the solution is stable to perturbations in data.
- If $\sigma_1 \gg \sigma_p$, there is a potential for instability. It is diminished if the data itself has small components in the direction of singular vectors corresponding to small singular values.
- If $\sigma_1 \gg \sigma_p$, and there is a clear delineation between "small" singular values and the rest, we simple discard the small singular values and treat the problem as one of smaller rank with "good" singular values.
- If $\sigma_1 \gg \sigma_p$, and there is no clear delineation between "small" singular values and the rest, we have to discard some of them, but which ones? This leads to regularization issues. In any case, any method that discards small singular values produces a **truncated SVD** (TSVD) solution.

4.4 An Example of a Rank-Deficient Problem

]

Linear Tomography Models

Note: Rank deficient problems are *automatically* ill-posed.

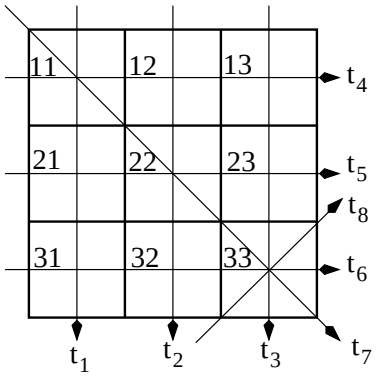
Basic Idea:
A ray emanates from one known point to another along a known path ℓ, with a detectable property which is observable data. These data are used to estimate a travel property of the medium. For example, let the property be travel time, so that:

- Travel time is given by $t = \int_\ell \frac{dt}{dx} dx = \int_\ell \frac{1}{v\,(x)} dx$
- We can linearize by making paths straight lines.
- Discretize by embedding the medium in a square (cube) and subdividing it into regular subsquares (cubes) in which we assume "slowness" (parameter of the problem) is constant.
- Transmit the ray along specified paths and collect temporal data to be used in estimating "slowness".

]

Example 1.6 and 4.1

The figure for this experiment (assume each subsquare has sides of length 1, so the size of the large square is 3×3):



Example 1.6 and 4.1

Corresponding matrix of distances G (rows of G represent distances along corresponding path, columns the ray distances across each subblock) and resulting system:

$$G\mathbf{m} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ \sqrt{2} & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} s_{11} \\ s_{12} \\ s_{13} \\ s_{21} \\ s_{22} \\ s_{23} \\ s_{31} \\ s_{32} \\ s_{33} \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \\ t_8 \end{bmatrix} = \mathbf{d}$$

Observe: in this Example $m = 8$ and $n = 9$, so this is rank deficient. Now run the example file for this example. We need to fix the path. Assuming we are in the directory MatlabTools, do the following:

```
>addpath('Examples/chap4/examp1')
>path
```

4.5 Discrete Ill-Posed Problems

]

What Are They?

These problems arise due to ill-conditioning of G, *as opposed to a rank deficiency problem. Theoretically*, they are not ill-posed, like the Hilbert matrix. But practically speaking, they behave like ill-posed problems. Authors present a

hierarchy of sorts for a problem with system $G\mathbf{m} = \mathbf{d}$. These order expressions are valid as $j \to \infty$.

- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $0 < \alpha \leq 1$, the problem is **mildly** ill-posed.
- $\mathcal{O}\left(\frac{1}{j^\alpha}\right)$ with $\alpha > 1$, the problem is **moderately** ill-posed.
- $\mathcal{O}\left(e^{-\alpha j}\right)$ with $0 < \alpha$, the problem is **severely** ill-posed.

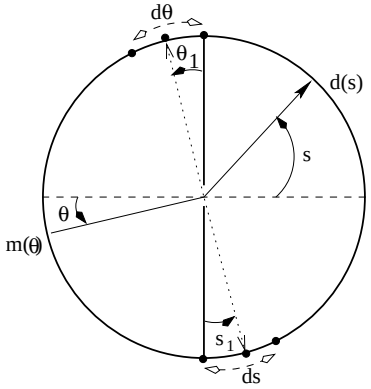
]

A Severly Ill-Posed Problem

The Shaw Problem:
An optics experiment is performed by dividing a circle using a vertical transversal with a slit in the middle. A variable intensity light source is placed around the left half of the circle and rays pass through the slit, where they are measured at points on the right half of the circle.

- Measure angles counterclockwise from the x-axis, using $-\pi/2 \leq \theta \leq \pi/2$ for the source intensity $m(\theta)$, and $-\pi/2 \leq s \leq \pi/2$ for destination intensity $d(s)$.
- The model for this problem comes from diffraction theory: $d(s) = \int_{-\pi/2}^{\pi/2} (\cos(s) + \cos(\theta))^2 \left(\frac{\sin(\pi(\sin(s)$

The Shaw Problem



Two Problems:

- The forward problem: given source intensity $m(\theta)$, compute the destination intensity $d(s)$.

- The inverse problem: given destination intensity $d(s)$, compute the source intensity $m(\theta)$.
- It can be shown that the inverse problem is severly ill-posed.

The Shaw Problem

How To Discretize The Problem:

- Discretize the parameter domain $-\pi/2 \leq \theta \leq \pi/2$ and the data domain $-\pi/2 \leq s \leq \pi/2$ into n subintervals of equal size $\Delta s = \Delta\theta = \pi/n$.
- Therefore, and let s_i, θ_i be the midpoints of the i-th subintervals:

$$s_i = \theta_i = -\frac{\pi}{2} + \frac{(i - 0.5)\,\pi}{n},\ i = 1, 2, \ldots, n.$$

- Define

$$G_{i,j} = (\cos(s_i) + \cos(\theta_j))^2 \left(\frac{\sin(\pi(\sin(s_i) + \sin(\theta_j)))}{\pi(\sin(s_i) + \sin(\theta_j))} \right)^2 \Delta\theta$$

- Thus if $m_j \approx m(\theta_j)$, $d_i \approx d(s_i)$, $\mathbf{m} = (m_1, m_2, \ldots, m_n)$ and $\mathbf{d} = (d_1, d_2, \ldots, d_n)$, then discretization and the midpoint rule give $G\mathbf{m} = \mathbf{d}$, as in Chapter 3.

The Shaw Problem

Now we can examine the example files on the text CD for this problem. This file lives in 'MatlabTools/Examples/chap4/examp1'. First add the correctd path, then open the example file `examp.m` for editing. However, here's an easy way to build the matrix G without loops. Basically, these tools were designed to help with 3-D plotting.

```
> n = 20
> ds = pi/n
> s = linspace(ds/2, pi - ds/2,n)
> theta = s;
> [S, Theta] = meshgrid(s,theta);
>G = (cos(S) + cos(Theta)).^2 .* (sin(pi*(sin(S) + ...
sin(Theta))./(pi*(sin(S) + sin(Theta))).^2*ds;
> % want to see G(s,theta)?
> mesh(S,Theta,G)
> cond(G)
> svd(G)
> rank(G)
```

5 Chapter 5: Tikhonov Regularization

5.1 Tikhonov Regularization and Implementation via SVD

]

Basics

Regularization:
This means "turn an ill-posed problem into a well-posed 'near by' problem". Most common method is Tikhonov regularization, which is motivated in context of our possibly ill-posed $G\mathbf{m} = \mathbf{d}$, i.e., minimize $\|G\mathbf{m} - \mathbf{d}\|_2$, problem by:

- Problem: minimize $\|\mathbf{m}\|_2$ subject to $\|\mathbf{Gm} - \mathbf{d}\|_\mathbf{2} \leq \delta$
- Problem: minimize $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to $\|\mathbf{m}\|_2 \leq \epsilon$
- Problem: (**damped least squares**) minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{m}\|_2^2$. This is the **Tikhonov regularization** of the original problem.
- Problem: find minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$.e function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$.

Basics

Regularization:
All of the above problems are equivalent under mild restrictions thanks to the principle of Lagrange multipliers:

- The minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$ must occur at the stationary points of function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$ (so we could write $\lambda = \alpha^2$ to emphasize non-negativity.)
- We can see why this is true in the case of a two dimensional $\mathbf{x}$ by examining contour curves.
- Square the terms in the first two problems and we see that the associated Lagrangians are related if we take reciprocals of α.
- Various values of α give a trade-off between the instability of the unmodified least squares problem and loss of accuracy of the smoothed problem. This can be understood by tracking the value of the minimized function in the form of a path depending on δ, ϵ or α.

6 Chapter 5: Tikhonov Regularization

6.1 Tikhonov Regularization and Implementation via SVD

]

Basics

Regularization:
This means "turn an ill-posed problem into a well-posed 'near by' problem". Most common method is Tikhonov regularization, which is motivated in context of our possibly ill-posed $G\mathbf{m} = \mathbf{d}$, i.e., minimize $\|G\mathbf{m} - \mathbf{d}\|_2$, problem by:

- Problem: minimize $\|\mathbf{m}\|_2$ subject to $\|\mathbf{Gm} - \mathbf{d}\|_\mathbf{2} \leq \delta$
- Problem: minimize $\|G\mathbf{m} - \mathbf{d}\|_2$ subject to $\|\mathbf{m}\|_2 \leq \epsilon$
- Problem: (**damped least squares**) minimize $\|G\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{m}\|_2^2$. This is the **Tikhonov regularization** of the original problem.
- Problem: find minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$.e function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$.

Basics

Regularization:
All of the above problems are equivalent under mild restrictions thanks to the principle of Lagrange multipliers:

- Minima of $f(\mathbf{x})$ occur at **stationary points** of $f(x)$ ($\nabla f = 0$.)
- Minima of $f(\mathbf{x})$ subject to constraint $g(\mathbf{x}) \leq c$ must occur at stationary points of function $L = f(\mathbf{x}) + \lambda g(\mathbf{x})$, for some $\lambda \geq 0$ (we can write $\lambda = \alpha^2$ to emphasize non-negativity.)
- We can see why this is true in the case of a two dimensional $\mathbf{x}$ by examining contour curves.
- Square the terms in the first two problems and we see that the associated Lagrangians are related if we take reciprocals of α.
- Various values of α give a trade-off between the instability of the unmodified least squares problem and loss of accuracy of the smoothed problem. This can be understood by tracking the value of the minimized function in the form of a path depending on δ, ϵ or α.