CHAPTER 1: INTRODUCTION

1.1 Inverse Theory: What It Is and What It Does

Inverse theory, at least as I choose to define it, is the fine art of estimating model parameters from data. It requires a knowledge of the forward model capable of predicting data if the model parameters were, in fact, already known. Anyone who attempts to solve a problem in the sciences is probably using inverse theory, whether or not he or she is aware of it. Inverse theory, however, is capable (at least when properly applied) of doing much more than just estimating model parameters. It can be used to estimate the "quality" of the predicted model parameters. It can be used to determine which model parameters, or which combinations of model parameters, are best determined. It can be used to determine which data are most important in constraining the estimated model parameters. It can determine the effects of noisy data on the stability of the solution. Furthermore, it can help in experimental design by determining where, what kind, and how precise data must be to determine model parameters.

Inverse theory is, however, inherently mathematical and as such does have its limitations. It is best suited to estimating the numerical values of, and perhaps some statistics about, model parameters for some *known* or *assumed* mathematical model. It is less well suited to provide the fundamental mathematics or physics of the model itself. I like the example Albert Tarantola gives in the introduction of his classic book[1] on inverse theory. He says, ". . . you can always measure the captain's age (for instance by picking his passport), but there are few chances for this measurement to carry much information on the number of masts of the boat." You must have a good idea of the applicable forward model in order to take advantage of inverse theory. Sooner or later, however, most practitioners become rather fanatical about the benefits of a particular approach to inverse theory. Consider the following as an example of how, or how not, to apply inverse theory. The existence or nonexistence of a God is an interesting question. Inverse theory, however, is poorly suited to address this question. However, if one assumes that there is a God and that She makes angels of a certain size, then inverse theory might well be appropriate to determine the number of angels that could fit on the head of a pin. Now, who said practitioners of inverse theory tend toward the fanatical?

In the rest of this chapter, I will give some useful definitions of terms that will come up time and again in inverse theory, and give some examples, mostly from Menke's book, of how to set up forward problems in an attempt to clearly identify model parameters from data.

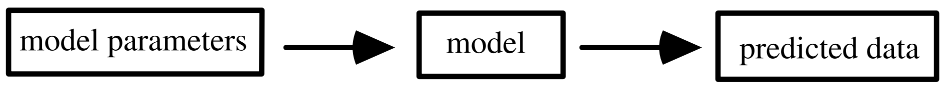
[1] *Inverse Problem Theory*, by Albert Tarantola, Elsevier Scientific Publishing Company, 1987.

1.2 Useful Definitions

Let us begin with some definitions of things like *forward* and *inverse* theory, *models* and *model parameters*, *data*, etc.

Forward Theory: The (mathematical) process of predicting data based on some physical or mathematical model with a given set of model parameters (and perhaps some other appropriate information, such as geometry, etc.).

Schematically, one might represent this as follows:



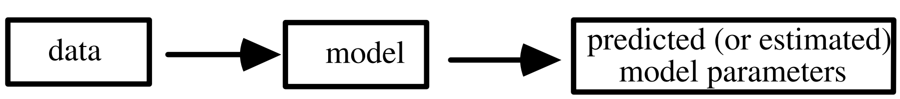
As an example, consider the two-way vertical travel time t of a seismic wave through M layers of thickness h_i and velocity v_i. Then t is given by

$$t = 2\sum_{i=1}^{M} \frac{h_i}{v_i} \tag{1.1}$$

The forward problem consists of predicting data (travel time) based on a (mathematical) model of how seismic waves travel. Suppose that for some reason thickness was known for each layer (perhaps from drilling). Then only the M velocities would be considered model parameters. One would obtain a particular travel time t for each set of model parameters one chooses.

Inverse Theory: The (mathematical) process of predicting (or estimating) the numerical values (and associated statistics) of a set of model parameters of an assumed model based on a set of data or observations.

Schematically, one might represent this as follows:



As an example, one might invert the travel time t above to determine the layer velocities. Note that one needs to know the (mathematical) model relating travel time to layer thickness and velocity information. Inverse theory should not be expected to provide the model itself.

Model: The model is the (mathematical) relationship between model parameters (and other auxiliary information, such as the layer thickness information in the previous example) and the data. It may be linear or nonlinear, etc.

Model Parameters: The model parameters are the numerical quantities, or unknowns, that one is attempting to estimate. The choice of model parameters is usually problem dependent, and quite often arbitrary. For example, in the case of travel times cited earlier, layer thickness is *not* considered a model parameter, while layer velocity is. There is nothing sacred about these choices. As a further example, one might choose to cast the previous example in terms of slowness s_i, where:

$$s_i = 1 / v_i \tag{1.2}$$

Travel time t is a nonlinear function of layer velocities but a linear function of layer slowness. As you might expect, it is much easier to solve linear than nonlinear inverse problems. A more serious problem, however, is that linear and nonlinear formulations may result in different estimates of velocity if the data contain any noise. The point I am trying to impress on you now is that there is quite a bit of freedom in the way model parameters are chosen, and it can affect the answers you get!

Data: Data are simply the observations or measurements one makes in an attempt to constrain the solution of some problem of interest. Travel time in the example above is an example of data. There are, of course, many other examples of data.

Some examples of inverse problems (mostly from Menke) follow:

- Medical tomography
- Earthquake location
- Earthquake moment tensor inversion
- Earth structure from surface or body wave inversion
- Plate velocities (kinematics)
- Image enhancement
- Curve fitting
- Satellite navigation
- Factor analysis

1.3 Possible Goals of an Inverse Analysis

Now let us turn our attention to some of the possible goals of an inverse analysis. These might include:

1. Estimates of a set of model parameters (obvious).
2. Bounds on the range of acceptable model parameters.
3. Estimates of the formal uncertainties in the model parameters.
4. How sensitive is the solution to noise (or small changes) in the data?
5. Where, and what kind, of data are best suited to determine a set of model parameters?
6. Is the fit between predicted and observed data adequate?

7. Is a more complicated (i.e., more model parameters) model significantly better than a more simple model?

Not all of these are completely independent goals. It is important to realize, as early as possible, that there is much more to inverse theory than simply a set of estimated model parameters. Also, it is important to realize that there is very often not a single "correct" answer. Unlike a mathematical inverse, which either exists or does not exist, there are many possible approximate inverses. These may give different answers. Part of the goal of an inverse analysis is to determine if the "answer" you have obtained is reasonable, valid, acceptable, etc. This takes experience, of course, but you have begun the process.

Before going on with how to formulate the mathematical methods of inverse theory, I should mention that there are two basic branches of inverse theory. In the first, the model parameters and data are discrete quantities. In the second, they are continuous functions. An example of the first might occur with the model parameters we seek being given by the moments of inertia of the planets:

$$\text{model parameters} = I_1, I_2, I_3, \ldots, I_{10} \tag{1.3}$$

and the data being given by the perturbations in the orbital periods of satellites:

$$\text{data} = T_1, T_2, T_3, \ldots, T_N \tag{1.4}$$

An example of a continuous function type of problem might be given by velocity as a function of depth:

$$\text{“model parameters”} = v(z) \tag{1.5}$$

and the data given by a seismogram of ground motion

$$\text{“data”} = d(t) \tag{1.6}$$

Separate strategies have been developed for discrete and continuous inverse theory. There is, of course, a fair bit of overlap between the two. In addition, it is often possible to approximate continuous functions with a discrete set of values. There are potential problems (aliasing, for example) with this approach, but it often makes otherwise intractable problems tractable. Menke's book deals exclusively with the discrete case. This course will certainly emphasize discrete inverse theory, but I will also give you a little of the continuous inverse theory at the end of the semester.

1.4 Nomenclature

Now let us introduce some nomenclature. In these notes, vectors will be denoted by boldface lowercase letters, and matrices will be denoted by boldface uppercase letters.

Suppose one makes N measurements in a particular experiment. We are trying to determine the values of M model parameters. Our nomenclature for data and model parameters will be

$$\text{data: } \mathbf{d} = [d_1, d_2, d_3, \ldots, d_N]^T \tag{1.7}$$

$$\text{model parameters: } \mathbf{m} = [m_1, m_2, m_3, \ldots, m_M]^T \tag{1.8}$$

where **d** and **m** are N and M dimensional column vectors, respectively, and T denotes transpose.

The model, or relationship between **d** and **m**, can have many forms. These can generally be classified as either *explicit* or *implicit,* and either *linear* or *nonlinear*.

Explicit means that the data and model parameters *can* be separated onto different sides of the equal sign. For example,

$$d_1 = 2m_1 + 4m_2 \tag{1.9}$$

and

$$d_1 = 2m_1 + 4m_1^2 m_2 \tag{1.10}$$

are two explicit equations.

Implicit means that the data *cannot* be separated on one side of an equal sign with model parameters on the other side. For example,

$$d_1(m_1 + m_2) = 0 \tag{1.11}$$

and

$$d_1(m_1 + m_1^2 m_2) = 0 \tag{1. 12}$$

are two implicit equations. In each example above, the first represents a *linear* relationship between the data and model parameters, and the second represents a *nonlinear* relationship.

In this course we will deal exclusively with *explicit* type equations, and predominantly with *linear* relationships. Then, the *explicit linear* case takes the form

$$\mathbf{d} = \mathbf{Gm} \tag{1.13}$$

where **d** is an N-dimensional data vector, **m** is an M-dimensional model parameter vector, and **G** is an $N \times M$ matrix containing only constant coefficients.

The matrix **G** is sometimes called the *kernel* or *data kernel* or even the Green's function because of the analogy with the continuous function case:

$$\mathbf{d}(x) = \int \mathbf{G}(x,\, t)\mathbf{m}(t)\, dt \tag{1. 14}$$

Consider the following discrete case example with two observations (N = 2) and three model parameters (M = 3):

$$\begin{aligned} d_1 &= 2m_1 + 0m_2 - 4m_3 \\ d_2 &= m_1 + 2m_2 + 3m_3 \end{aligned} \tag{1.15}$$

which may be written as

$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -4 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} \tag{1.16}$$

or simply

$$\mathbf{d} = \mathbf{Gm} \tag{1.13}$$

where

$$\mathbf{d} = [d_1, d_2]^{\mathrm{T}}$$

$$\mathbf{m} = [m_1, m_2, m_3]^{\mathrm{T}}$$

and

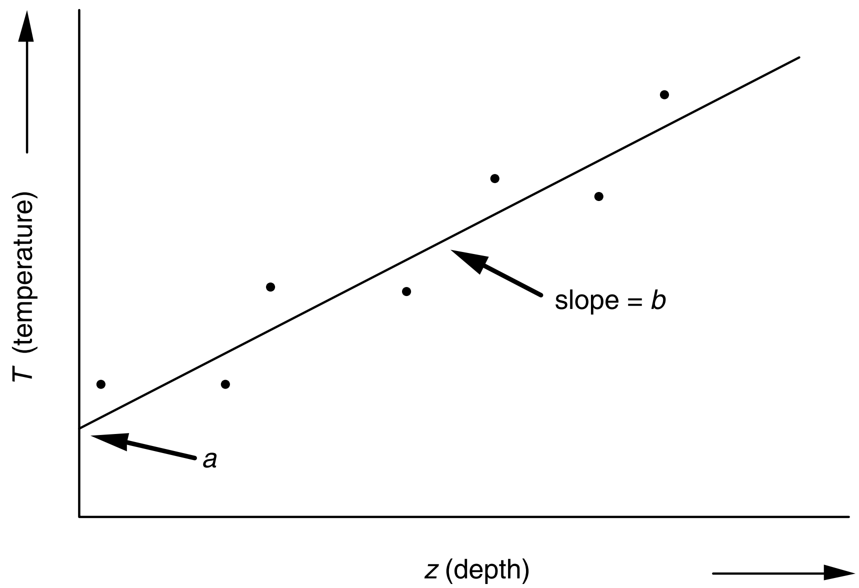
$$\mathbf{G} = \begin{bmatrix} 2 & 0 & -4 \\ 1 & 2 & 3 \end{bmatrix} \tag{1.17}$$

Then **d** and **m** are 2 × 1 and 3 × 1 column vectors, respectively, and **G** is a 2 × 3 matrix with constant coefficients.

On the following pages I will give some examples of how forward problems are set up using matrix notation. See pages 10–16 of Menke for these and other examples.

1.5 Examples of Forward Problems

1.5.1 Example 1: Fitting a Straight Line (See Page 10 of Menke)



Suppose that N temperature measurements T_i are made at depths z_i in the earth. The data are then a vector **d** of N measurements of temperature, where $\mathbf{d} = [T_1, T_2, T_3, \ldots, T_N]^T$. The depths z_i are not data. Instead, they provide some auxiliary information that describes the geometry of the experiment. This distinction will be further clarified below.

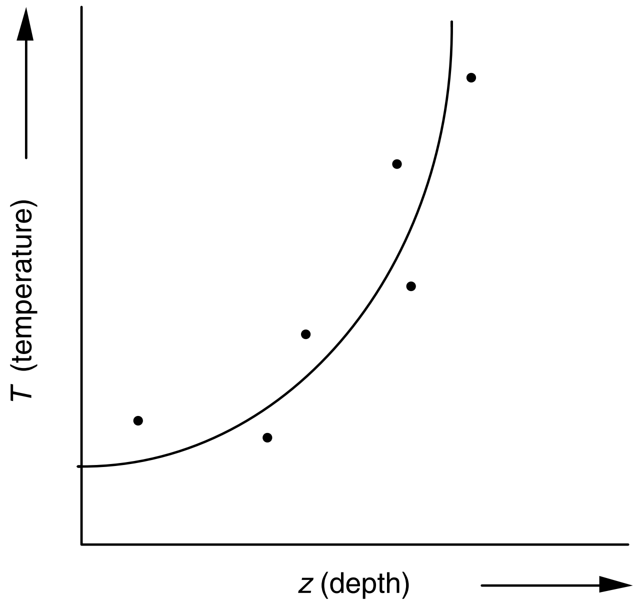
Suppose that we assume a model in which temperature is a linear function of depth: $T = a + bz$. The intercept a and slope b then form the two model parameters of the problem, $\mathbf{m} = [a, b]^T$. According to the model, each temperature observation must satisfy $T = a + zb$:

$$T_1 = a + bz_1$$
$$T_2 = a + bz_2$$
$$\vdots$$
$$T_N = \mathrm{a} + bz_N$$

These equations can be arranged as the matrix equation $\mathbf{Gm} = \mathbf{d}$:

$$\begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_N \end{bmatrix} = \begin{bmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_N \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

1.5.2 Example 2: Fitting a Parabola (See Page 11 of Menke)



If the model in example 1 is changed to assume a quadratic variation of temperature with depth of the form $T = a + bz + cz^2$, then a new model parameter is added to the problem, $\mathbf{m} = [a, b, c]^T$. The number of model parameters is now $M = 3$. The data are supposed to satisfy

$$\begin{aligned} T_1 &= a + bz_1 + cz_1^2 \\ T_2 &= a + bz_2 + cz_2^2 \\ &\vdots \\ T_N &= \text{a} + bz_N + cz_N^2 \end{aligned}$$

These equations can be arranged into the matrix equation

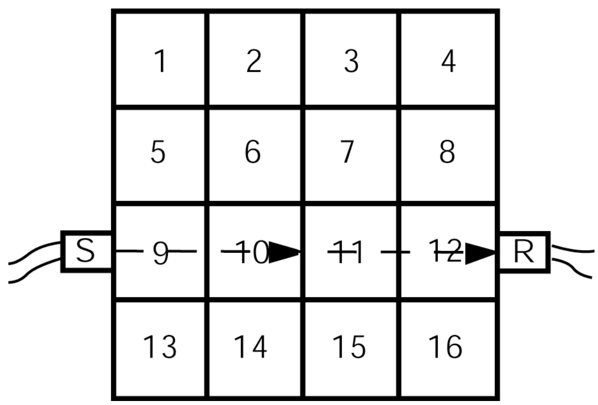
$$\begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_N \end{bmatrix} = \begin{bmatrix} 1 & z_1 & z_1^2 \\ 1 & z_2 & z_2^2 \\ \vdots & \vdots & \vdots \\ 1 & z_N & z_N^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

This matrix equation has the explicit linear form $\mathbf{Gm} = \mathbf{d}$. Note that, although the equation is linear in the data and model parameters, it is not linear in the auxiliary variable z.

The equation has a very similar form to the equation of the previous example, which brings out one of the underlying reasons for employing matrix notation: it can often emphasize similarities between superficially different problems.

1.5.3 Example 3: Acoustic Tomography (See Pages 12–13 of Menke)

Suppose that a wall is assembled from a rectangular array of bricks (Figure 1.1 from Menke, below) and that each brick is composed of a different type of clay. If the acoustic velocities of the different clays differ, one might attempt to distinguish the different kinds of bricks by measuring the travel time of sound across the various rows and columns of bricks, in the wall. The data in this problem are $N = 8$ measurements of travel times, $\mathbf{d} = [T_1, T_2, T_3, \ldots, T_8]^T$. The model assumes that each brick is composed of a uniform material and that the travel time of sound across each brick is proportional to the width and height of the brick. The proportionality factor is the brick's *slowness* s_i, thus giving $M = 16$ model parameters, $\mathbf{m} = [s_1, s_2, s_3, \ldots, s_{16}]^T$, where the ordering is according to the numbering scheme of the figure as



The travel time of acoustic waves (dashed lines) through the rows and columns of a square array of bricks is measured with the acoustic source S and receiver R placed on the edges of the square. The inverse problem is to infer the acoustic properties of the bricks (which are assumed to be homogeneous).

$$
\begin{array}{ll}
\text{row 1:} & T_1 = hs_1 + hs_2 + hs_3 + hs_4 \\
\text{row 2:} & T_2 = hs_5 + hs_6 + hs_7 + hs_8 \\
\vdots & \vdots \\
\text{column 4:} & T_8 = hs_4 + hs_8 + hs_{12} + hs_{16}
\end{array}
$$

and the matrix equation is

$$
\begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_8 \end{bmatrix} = h \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{16} \end{bmatrix}
$$

Here the bricks are assumed to be of width *and* height h.

1.5.4 Example 4: Seismic Tomography

An example of the impact of inverse methods in the geosciences: Northern California

- A large amount of data is available, much of it redundant.
- Patterns in the data can be interpreted qualitatively.
- Inversion results quantify the patterns.
- Perhaps, more importantly, inverse methods provide quantitative information on the resolution, standard error, and "goodness of fit."
- We cannot overemphasize the "impact" of colorful graphics, for both good and bad.
- Inverse theory is not a magic bullet. Bad data will still give bad results, and, interpretation of even good results requires breadth of understanding in the field.
- Inverse theory does provide quantitative information on how well the model is "determined," importance of data, and model errors.
- Another example: improvements in "imaging" subduction zones.

1.5.5 Example 5: Convolution

Convolution is widely significant as a physical concept and offers an advantageous starting point for many theoretical developments. One way to think about convolution is that it describes the action of an observing instrument when it takes a weighted mean of some physical quantity over a narrow range of some variable. All physical observations are limited in this way, and for this reason alone convolution is ubiquitous (paraphrased from Bracewell, *The Fourier Transform and Its Applications*, 1964). It is widely used in time series analysis as well to represent physical processes.

The convolution of two functions f(x) and g(x) represented as $f(x)*g(x)$ is

$$\int_{-\infty}^{\infty} f(u)\, g(x-u)\, du \tag{1.18}$$

For discrete finite functions with common sampling intervals, the convolution is

$$h_k = \sum_{i=0}^{m} f_i g_{k-i} \qquad 0 < k < m+n \tag{1. 19}$$

A FORTRAN computer program for convolution would look something like:

```
      L=M+N-1
      DO 10 I=1,L
10    H(I)=0
      DO 20 I=1,M
      DO 20 J=1,N
20    H(I+J-1)=H(I+J-1)+G(I)*F(J)
```

Convolution may also be written using matrix notation as

$$\begin{bmatrix} f_1 & 0 & \cdot & \cdot & \cdot & 0 \\ f_2 & f_1 & \cdot & \cdot & \cdot & 0 \\ \cdot & f_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & f_1 \\ f_n & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & f_n & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & f_n \end{bmatrix} \cdot \begin{bmatrix} g_1 \\ g_2 \\ \cdot \\ \cdot \\ \cdot \\ g_m \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ \cdot \\ \cdot \\ \cdot \\ h_{n+m-1} \end{bmatrix} \tag{1. 20}$$

In the matrix form, we recognize our familiar equation $\mathbf{Gm} = \mathbf{d}$ (ignoring the confusing notation differences between fields, when, for example, g_1 above would be m_1), and we can define deconvolution as the inverse problem of finding $\mathbf{m} = \mathbf{G}^{-1}\mathbf{d}$. Alternatively, we can also reformulate the problem as $\mathbf{G}^T\mathbf{Gm} = \mathbf{G}^T\mathbf{d}$ and find the solution as $\mathbf{m} = [\mathbf{G}^T\mathbf{G}]^{-1}\,[\mathbf{G}^T\mathbf{d}]$.

1.6 Final Comments

The purpose of the previous examples has been to help you formulate forward problems in matrix notation. It helps you to clearly differentiate model parameters from other information needed to calculate "predicted" data. It also helps you separate data from everything else. Getting the forward problem set up in matrix notation is essential before you can invert the system.

The logical next step is to take the forward problem given by

$$\mathbf{d} = \mathbf{Gm} \tag{1.13}$$

and invert it for an estimate of the model parameters $\mathbf{m}^{\text{est}}$ as

$$\mathbf{m}^{\text{est}} = \mathbf{G}^{\text{“inverse”}}\,\mathbf{d} \tag{1.21}$$

We will spend a lot of effort determining just what $\mathbf{G}^{\text{”inverse”}}$ means when the inverse does not exist in the mathematical sense of

$$\mathbf{G}\mathbf{G}^{\text{“inverse”}} = \mathbf{G}^{\text{“inverse”}}\mathbf{G} = \mathbf{I} \tag{1.22}$$

where **I** is the identity matrix.

The next order of business, however, is to shift our attention to a review of the basics of *matrices and linear algebra* as well as *probability and statistics* in order to take full advantage of the power of inverse theory.