Inverse Problems in Geophysics

What is an inverse problem?

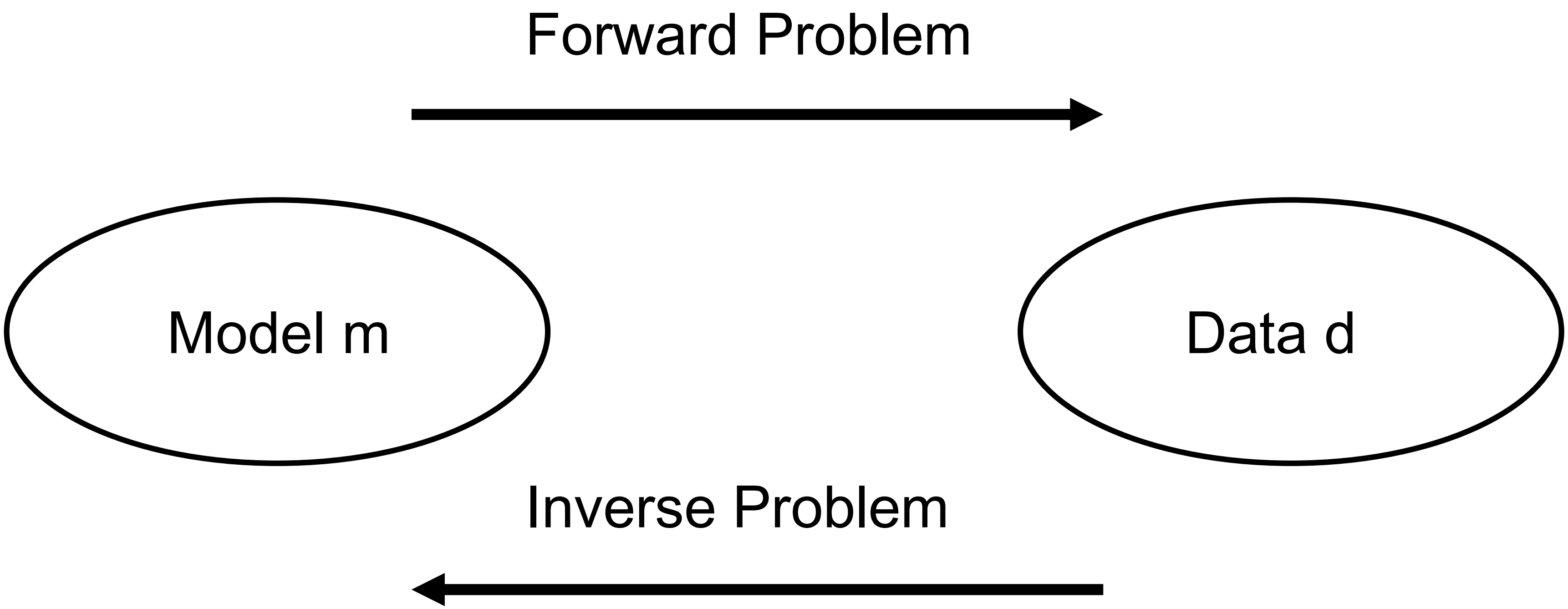
- ➢ - Illustrative Example
- ➢ - Exact inverse problems
- ➢ - Nonlinear inverse problems

Examples in Geophysics

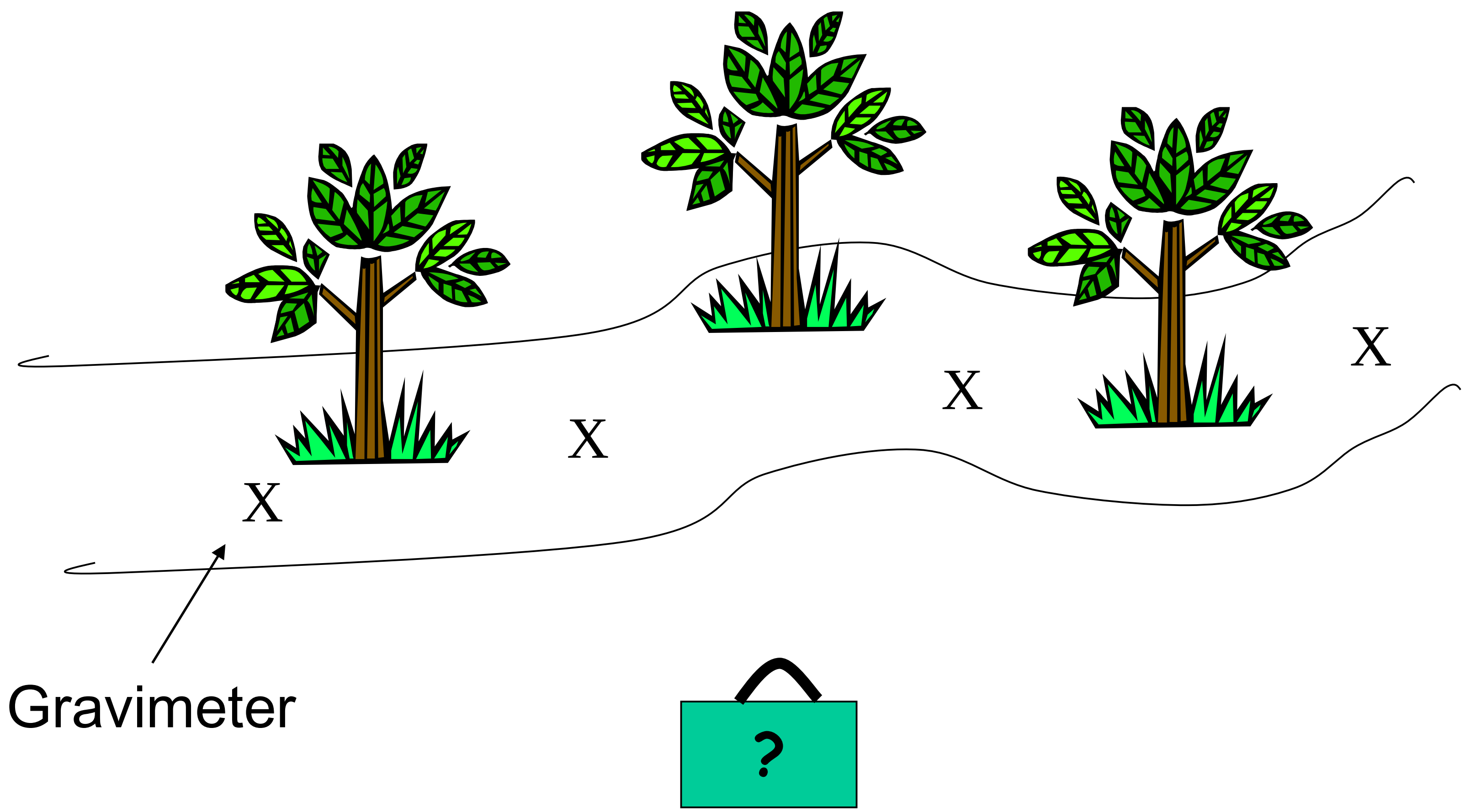
- ➢ - Traveltime inverse problems
- ➢ - Seismic Tomography
- ➢ - Location of Earthquakes
- ➢ - Global Electromagnetics
- ➢ - Reflection Seismology

Scope: Understand the concepts of data fitting and inverse problems and the associated problems. Simple mathematical formulation as linear (-ized) systems.

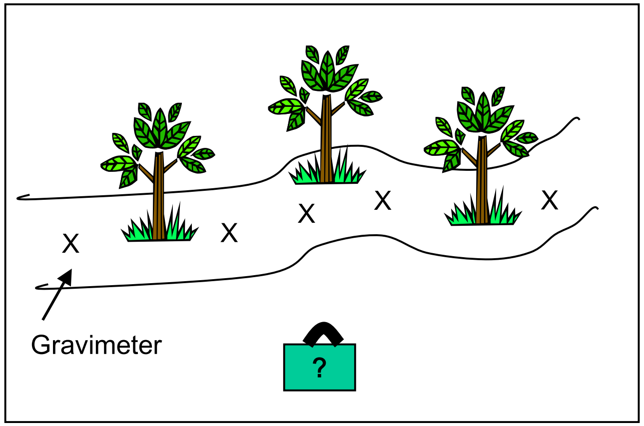
What is an inverse problem?



Treasure Hunt



Treasure Hunt – Forward Problem



We have observed some values:

10, 23, 35, 45, 56 μgals

How can we relate the observed gravity values to the subsurface properties?

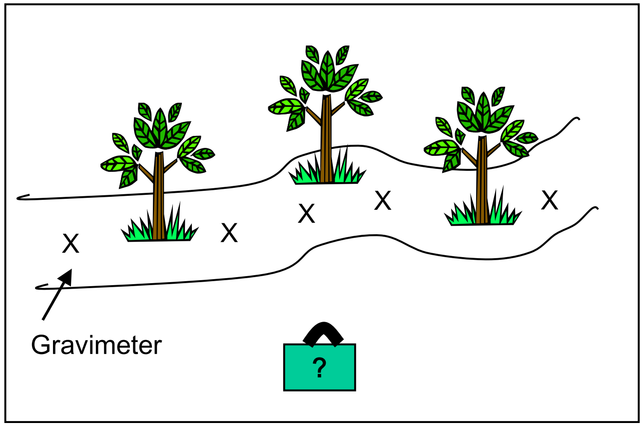
We know how to do the *forward* problem:

$$\Phi(r) = \int \frac{G\rho(r')}{|r - r'|} dV'$$

This equation relates the (observed) gravitational potential to the subsurface density.

-> given a density model we can predict the gravity field at the surface!

Treasure Hunt – Trial and Error



What else do we know?

Density sand: 2,2 g/cm^3
Density gold: 19,3 g/cm^3

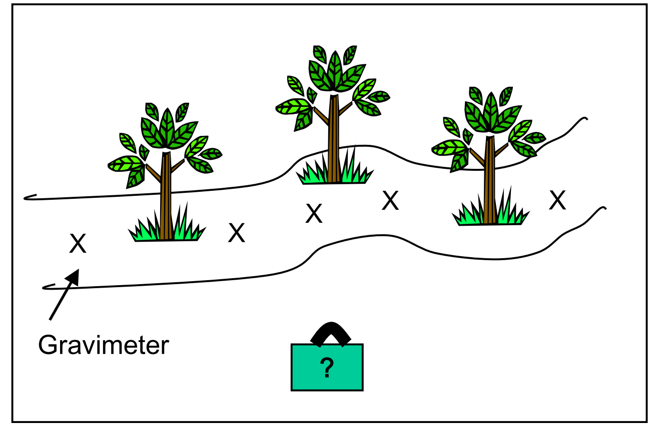
Do we know these values *exactly*?
How can we find out whether and if so where is the box with gold?

One approach:

Use the *forward* solution to calculate many models for a rectangular box situated somewhere in the ground and compare the *theoretical* (*synthetic*) data to the observations.

->Trial and error method

Treasure Hunt – Model Space

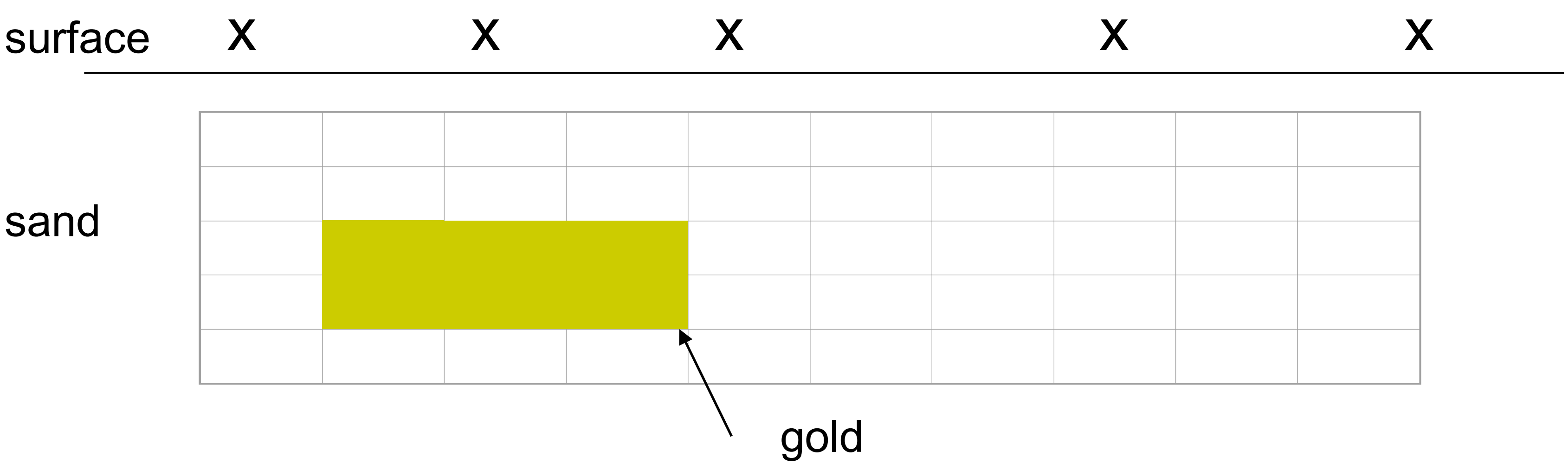


But ...

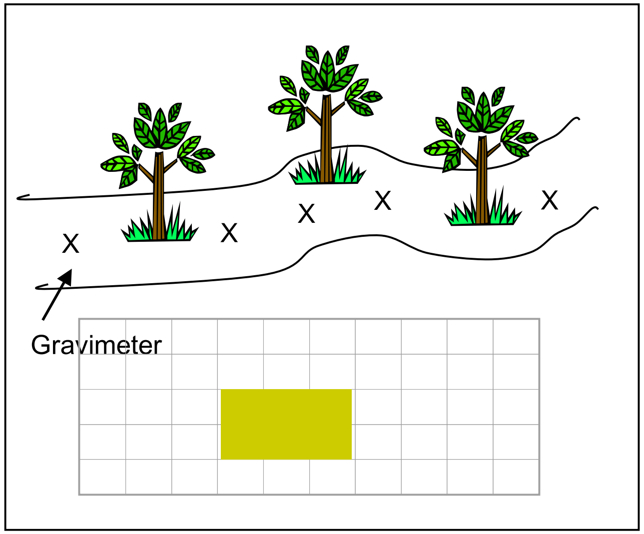
... we have to define *plausible* models for the beach. We have to somehow describe the model geometrically.

-> Let us

- divide the subsurface into a rectangles with variable density
- Let us assume a flat surface



Treasure Hunt – Non-uniqueness



Could we go through all possible models and compare the synthetic data with the observations?

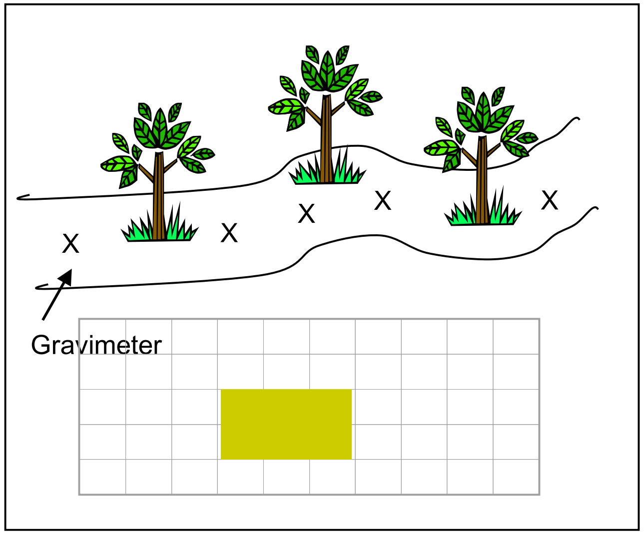
- at every rectangle two possibilities (sand or gold)
- $2^{50} \sim 10^{15}$ possible models

- **Too many models!**

- We have 10^{15} possible models but only 5 observations!
- It is likely that two or more models will fit the data (possibly perfectly well)

-> **Nonuniqueness of the problem!**

Treasure Hunt – A priori information

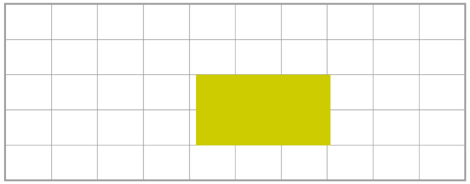


Is there anything we know about the treasure?

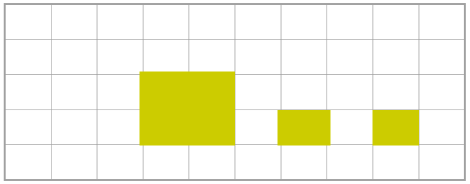
- How large is the box?
- Is it still intact?
- Has it possibly disintegrated?
- What was the shape of the box?
- Has someone already found it?

This is independent information that we may have which is as important and relevant as the observed data. This is colled *a priori* (or prior) information.
It will allow us to define plausible, possible, and unlikely models:

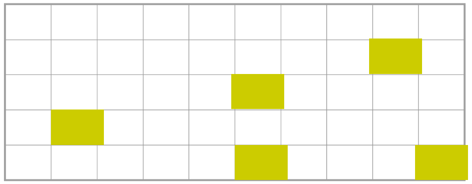
plausible



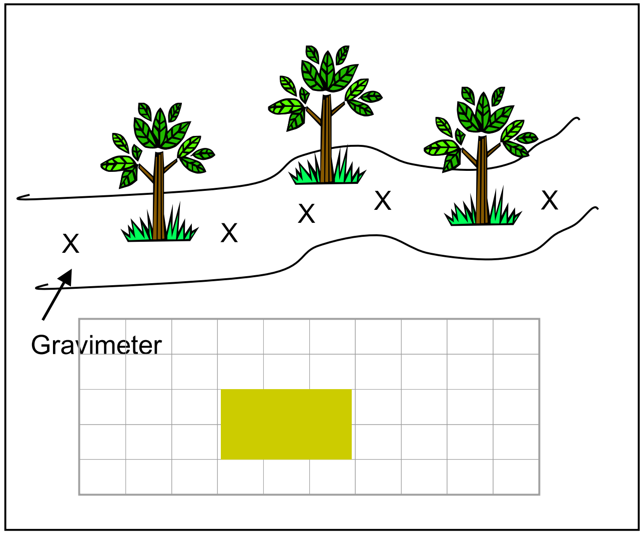
possible



unlikely



Treasure Hunt – Uncertainties (Errors)



Do we have errors in the data?

- Did the instruments work correctly?
- Do we have to correct for anything?
  (e.g. topography, tides, ...)

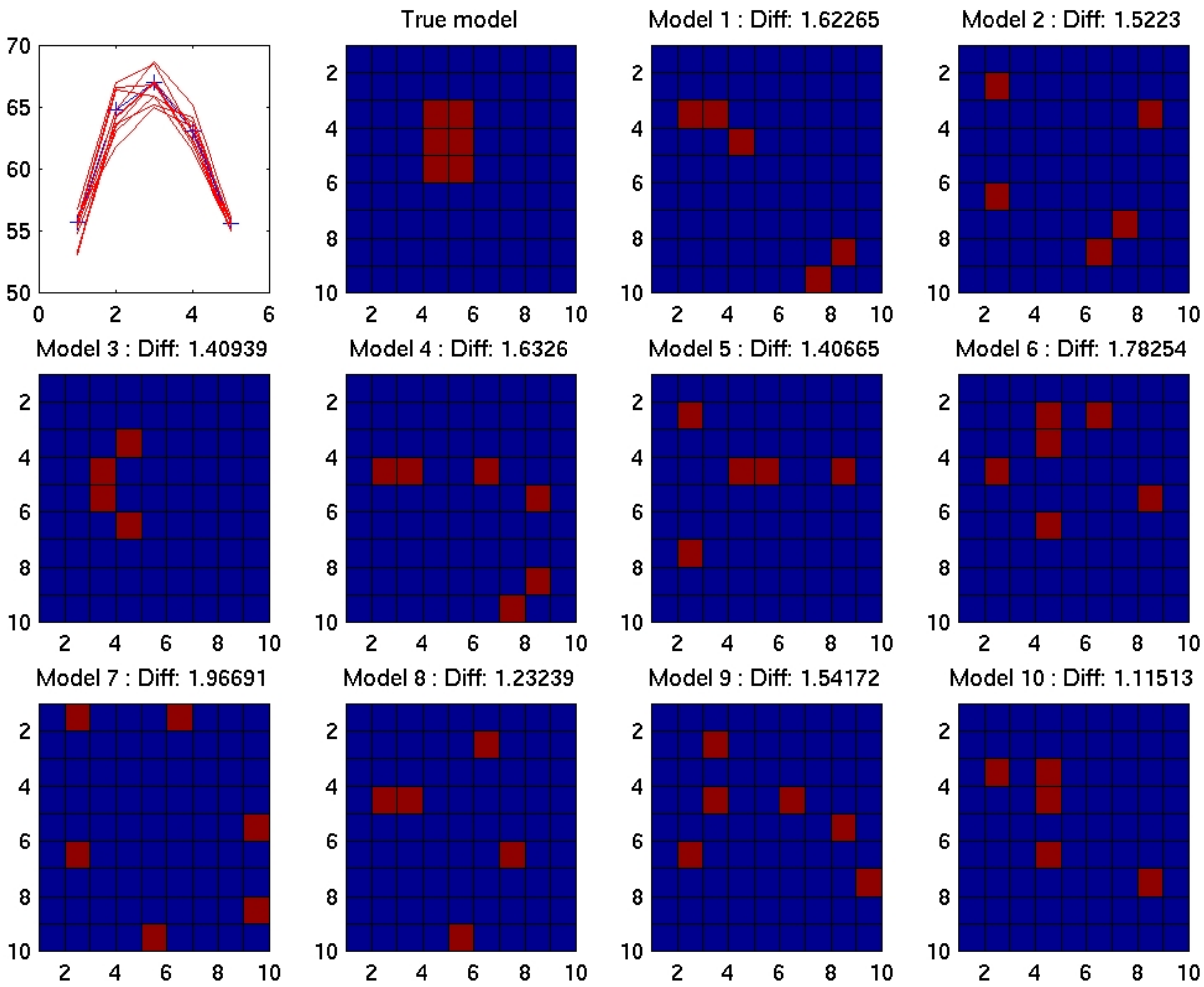
Are we using the right theory?

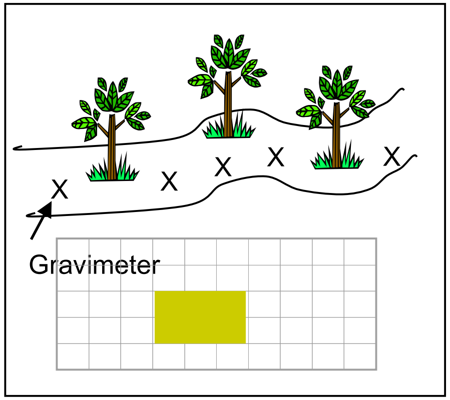
- Do we have to use 3-D models?
- Do we need to include the topography?
- Are there other materials in the ground apart from gold and sand?
- Are there adjacent masses which could influence the observations?

How (on Earth) can we quantify these problems?

Treasure Hunt - Example

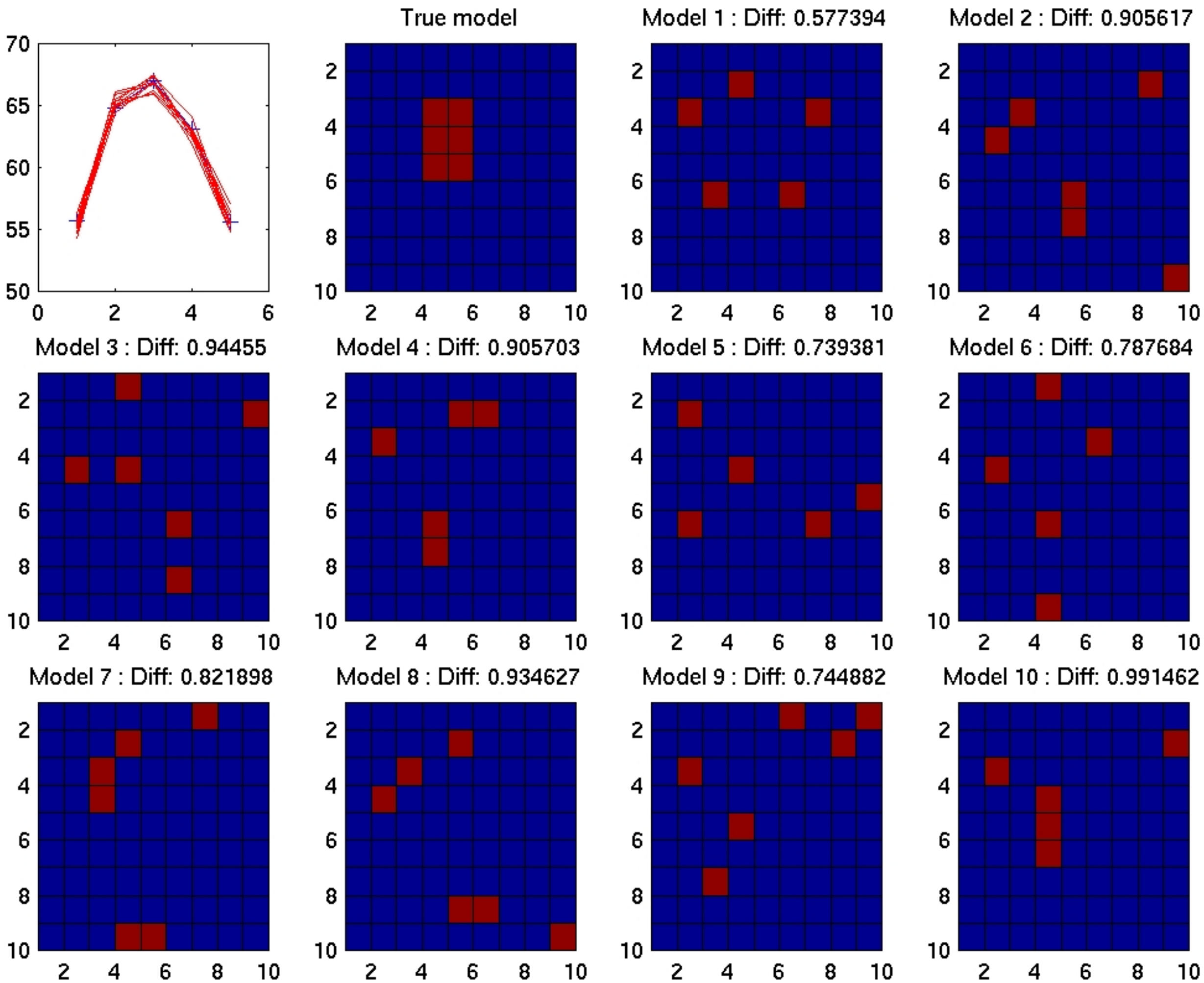
Models with less than 2% error.

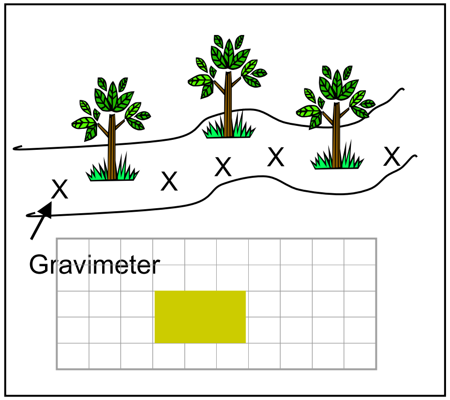




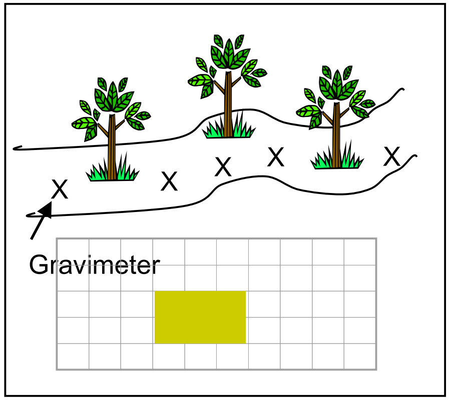
Treasure Hunt - Example

Models with less than 1% error.





Inverse Problems - Summary



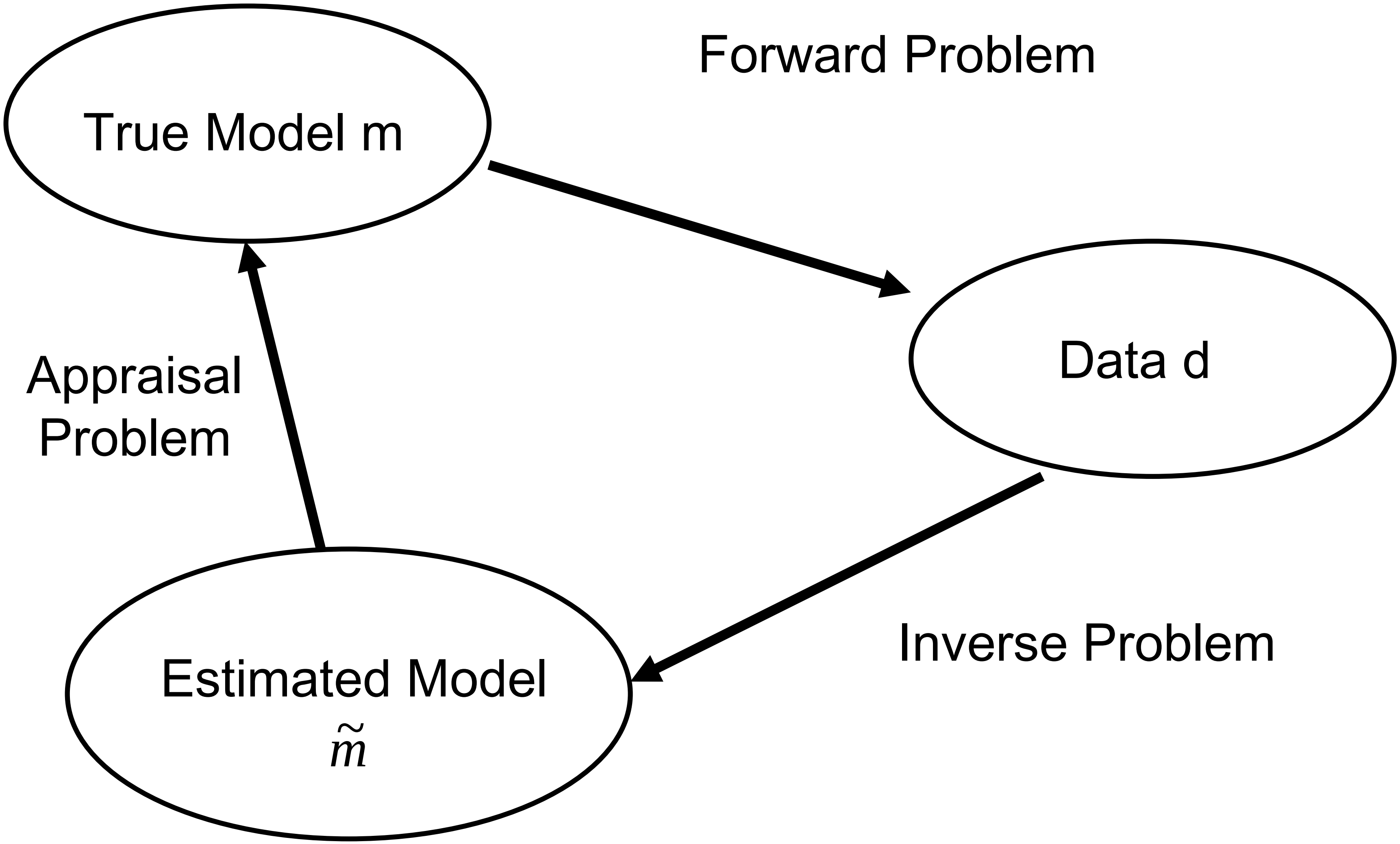
Inverse problems – inference about physical systems from data

- Data usually contain errors (data uncertainties)
- Physical theories are continuous
- infinitely many models will fit the data (non-uniqueness)
- Our physical theory may be inaccurate (theoretical uncertainties)
- Our forward problem may be highly nonlinear
- We always have a finite amount of data

The fundamental questions are:

How accurate are our data?
How well can we solve the forward problem?
What independent information do we have on the model space (a priori information)?

Corrected scheme for the real world



Linear(ized) Inverse Problems

Let us try and formulate the inverse problem mathematically:
Our goal is to determine the parameters of a (discrete) model m_i, i=1,...,m from a set of observed data d_j j=1,...,n. Model and data are functionally related (physical theory) such that

$$
\begin{aligned}
d_1 &= g_1(m_1,...,m_m) \\
d_2 &= g_2(m_1,...,m_m) \\
&\vdots \\
d_n &= g_n(m_1,...,m_m)
\end{aligned}
$$

This is the nonlinear formulation.

Note that m_i need not be model parameters at particular points in space but they could also be expansion coefficients of orthogonal functions (e.g. Fourier coefficients, Chebyshev coefficients etc.).

Linear(ized) Inverse Problems

If the functions $g_i(m_j)$ between model and data are linear we obtain

$$d_i = G_{ij} m_j$$

or

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

in matrix form. If the functions $A_i(m_j)$ between model and data are mildly non-linear we can consider the behavior of the system around some known (e.g. initial) model m_j^0:

$$d_i = G_l(m_j^0) + \left.\frac{\partial G_i}{\partial m_j}\right|_{m_j^0} \Delta m_j + \ldots$$

Linear(ized) Inverse Problems

$$d_i = G_l(m_j^0) + \left.\frac{\partial G_i}{\partial m_j}\right|_{m_j^0} \Delta m_j + \ldots$$

We will now make the following definitions:

$$d_i = G_i(m_j^0) + \Delta d_i$$

$$\Delta d_i = d_i - G_i(m_j^0)$$

Then we can write a linear(ized) problem for the nonlinear forward problem around some (e.g. initial) model m_0 neglecting higher order terms:

$$\Delta d_i = \left.\frac{\partial G_i}{\partial m_j}\right|_{m_j^0} \Delta m_j \qquad \Delta d_i = G_{ij} \Delta m_j \qquad G_{ij} = \left.\frac{\partial G_i}{\partial m_j}\right|_{m_j^0}$$

→ $$\mathbf{\Delta d = G \Delta m}$$

Linear(ized) Inverse Problems

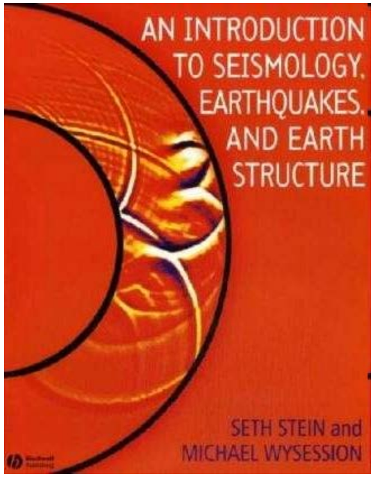
$$\Delta d = G\Delta m$$

Interpretation of this result:

1. m_0 may be an initial guess for our physical model
2. We may calculate (e.g. in a nonlinear way) the synthetic data $d=f(m_0)$.
3. We can now calculate the data misfit, $\Delta d=d-d_0$, where d_0 are the observed data.
4. Using some formal inverse operator A^{-1} we can calculate the corresponding model perturbation Dm. This is also called the gradient of the misfit function.
5. We can now calculate a new model $m=m_0+$ Dm which will – by definition – is a better fit to the data. We can start the procedure again in an iterative way.

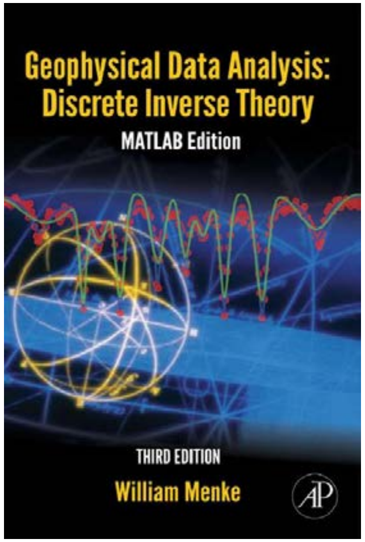
Literature

Stein and Wysession: Introduction to seismology, Chapter 7



Aki and Richards: Theoretical Seismology (1s edition) Chapter 12.3

Shearer: Introduction to seismology, Chapter 5



Menke, Discrete Inverse Problems

http://www.ldeo.columbia.edu/users/menke/gdadit/index.htm

Full ppt files and matlab routines

Formulation

Linear(-ized) inverse problems can be formulated in the following way:

$$d_i = G_{ij} m_j$$

(summation convention applies)

i=1,2,...,N	number of data
j=1,2,...,M	number of model parameters
G_{ij}	known (mxn)

We observe:

- The inverse problem has a unique solution if N=M and det(G)≠0, i.e. the data are linearly independent
- the problem is overdetermined if N>M
- the problem is underdetermined if M>N

Illustration – Unique Case

In this case N=M, and det(G) ≠0. Let us consider an example

$$\begin{aligned} 1 &= d_1 = 3m_1 + 2m_2 \\ 2 &= d_2 = m_1 + 4m_2 \end{aligned} \Rightarrow \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}$$

$$\mathrm{d} = \mathrm{Gm}$$

Let us check the determinant of this system: det(G)=10

$$\mathrm{G}^{-1}\mathrm{d} = \mathrm{G}^{-1}\mathrm{Gm} \Rightarrow \mathrm{m} = \mathrm{G}^{-1}\mathrm{d}$$

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} 0.4 & -0.2 \\ -0.1 & 0.3 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.5 \end{pmatrix}$$

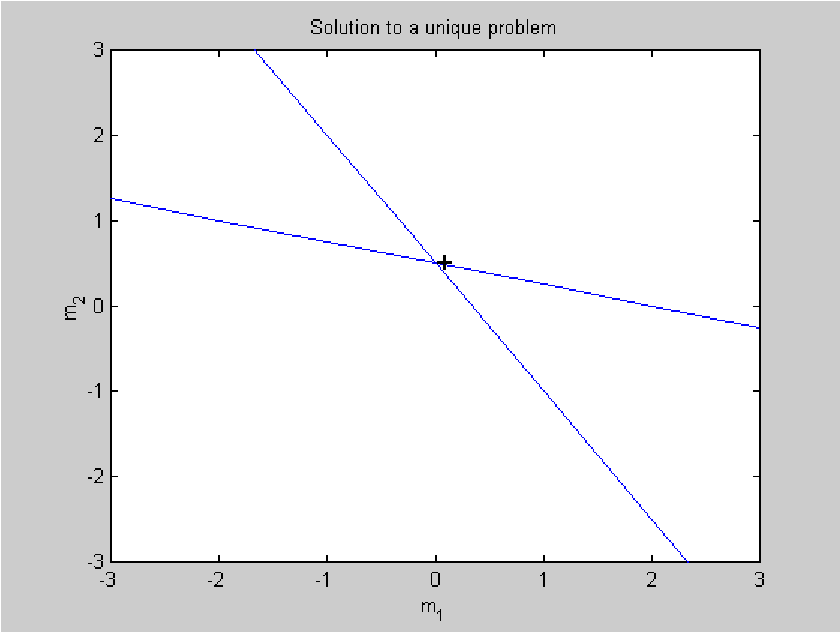


Illustration – Overdetermined Case

In this case N>M, there are more data than model parameters.
Let us consider examples with M=2, an overdetermined system would exist if N=3.

$$1 = d_1 = m_1$$
$$2 = d_2 = m_2$$
$$2 = d_3 = m_1 + m_2$$

A physical experiment which could result in these data: Individual Weight measurement of two masses m_1 and m_2 leading to the data d_1 and d_2 and weighing both together leads to d_3. In matrix form:

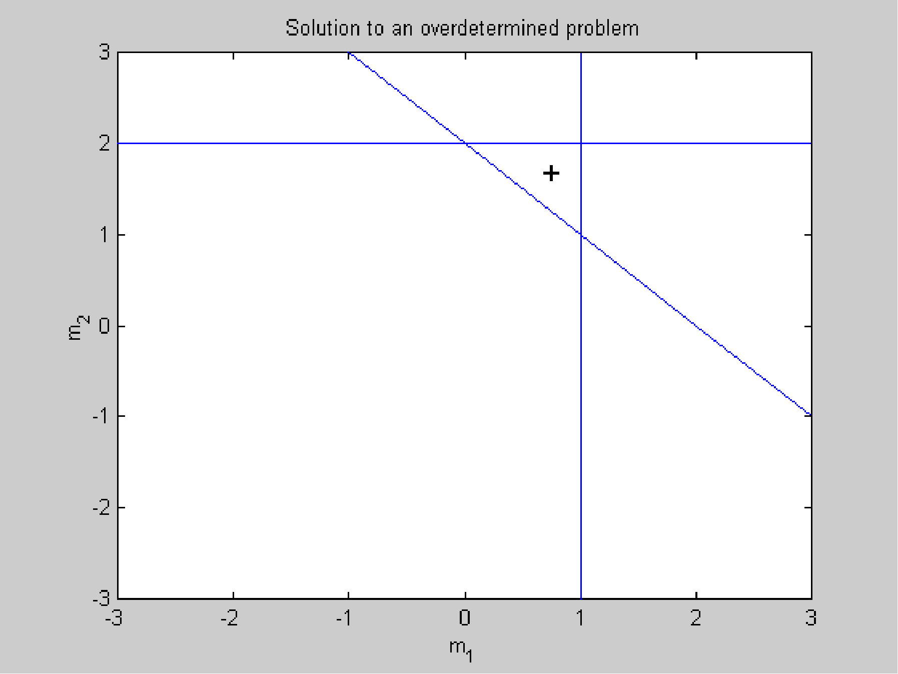
$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}$$

$$\mathrm{d} = \mathrm{Gm}$$

Illustration – Overdetermined Case

Let us consider this problem graphically

$$1 = m_1$$
$$2 = m_2$$
$$2 = m_1 + m_2$$



A common way to solve this problem is to minimize the difference between data vector d and the predicted data for some model m such that

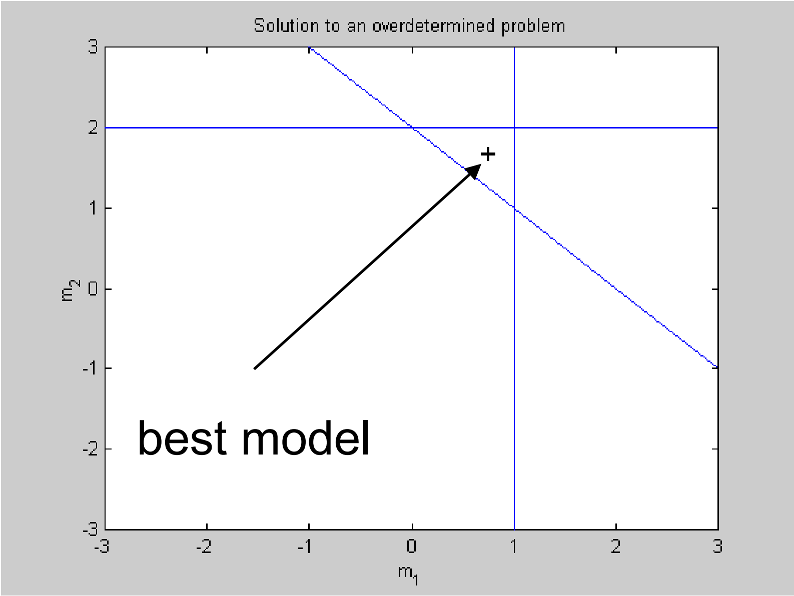
$$S = \|\mathrm{d} - \mathrm{Gm}\|^2$$

is minimal.

Illustration – Overdetermined Case

Using the L_2-norm leads us to the *least-squares* formulation of the problem. The solution to the minimization (and thus the inverse problem) is given as:

$$\tilde{m} = (G^T G)^{-1} G^T d$$



In our example the resulting (best) model estimation is:

$$\tilde{m} = \begin{pmatrix} 2/3 \\ 5/3 \end{pmatrix}$$

and is the model with the minimal distance to all three lines in the plot.

Illustration – Underdetermined Case

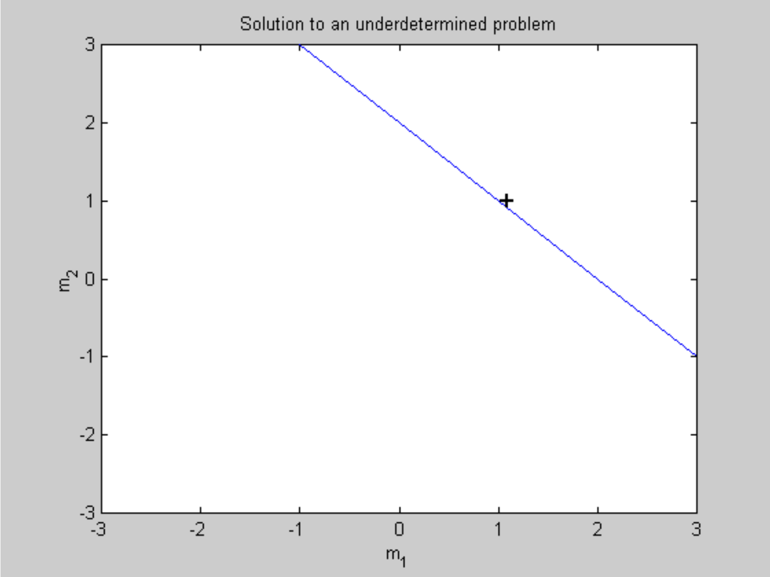
Let us assume we made one measurement of the combined weight of two masses:

$$m_1 + m_2 = d = 2$$

Clearly there are infinitely many solutions to this problem. A model estimate can be defined by choosing a model that fits the data exactly Am=d and has the smallest l_2 norm ||m||. Using Lagrange multipliers one can show that the minimum norm solution is given by

$$\tilde{m} = G^T (GG^T)^{-1} d$$

$$\tilde{m} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



Nonlinear Inverse Problems

Assume we have a wildly nonlinear functional relationship between model and data

$$\mathbf{d} = g(\mathbf{m})$$

The only option we have here is to try and go – in a sensible way – through the whole model space and calculate the misfit function

$$\mathbf{L} = \left\| \mathbf{d} - g(\mathbf{m}) \right\|$$

and find the model(s) which have the minimal misfit.

Model Search

The way how to explore a model space is a science itself!
Some key methods are:

1. Monte Carlo Method: Search in a random way through the model space and collect models with good fit.

2. Simulated Annealing. In analogy to a heat bath, or the generation of crystal one optimizes the quality (improves the misfit) of an ensemble of models. Decreasing the temperature would be equivalent to reducing the misfit (energy).

3. Genetic Algorithms. A pool of models recombines and combines information, every generation only the fittest survive and give on the successful properties.

4. Evolutionary Programming. A formal generalization of the ideas of genetic algorithms.

Inversion: the probabilistic approach

The misfit function

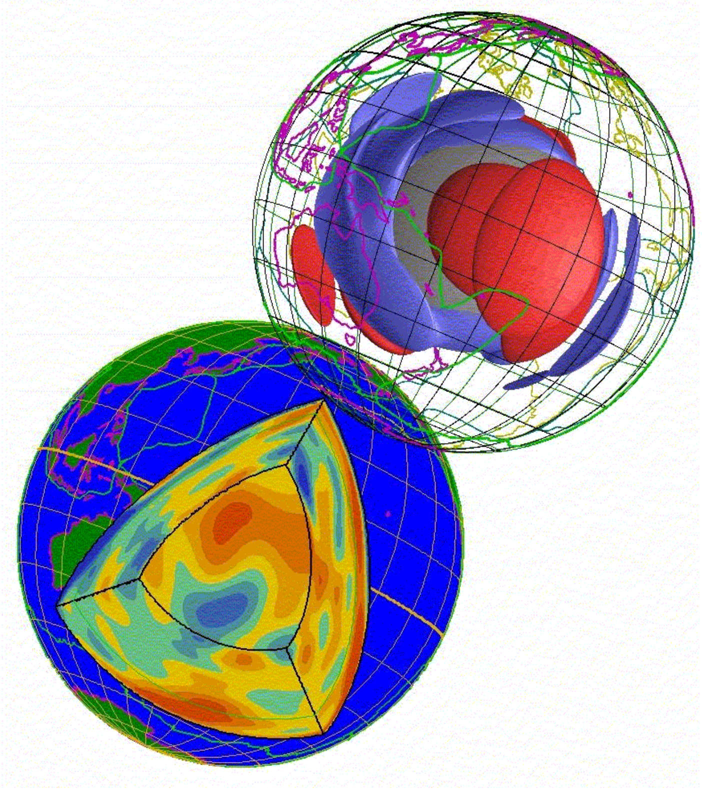
$$\mathbf{S}(\mathbf{m}) = (\mathbf{d} - g(\mathbf{m}))^{\mathsf{T}} (\mathbf{d} - g(\mathbf{m}))$$

can also be interpreted as a
likelihood function:

$$\sigma(\mathbf{m}) = e^{-\left[(\mathbf{d}-g(\mathbf{m}))^{\mathsf{T}}(\mathbf{d}-g(\mathbf{m}))\right]}$$

describing a probability density function (pdf) defined over the whole model space (assuming exact data and theory). This pdf is also called the a posteriori probability. In the probabilistic sense the a posteriori pdf is THE solution to the inverse problem.

Examples: Seismic Tomography



Data vector d:

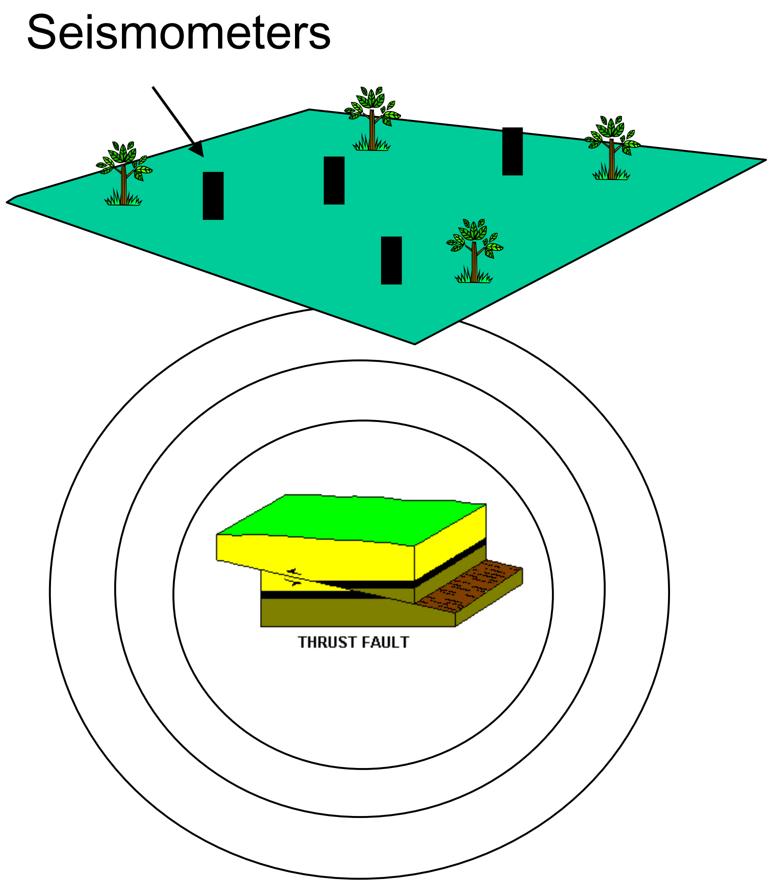
Traveltimes of phases observed at stations of the world wide seismograph network

Model m:

3-D seismic velocity model in the Earth's mantle. Discretization using splines, spherical harmonics, Chebyshev polynomials or simply blocks.

Sometimes 100000s of travel times and a large number of model blocks: underdetermined system

Examples: Earthquake location



Data vector d:

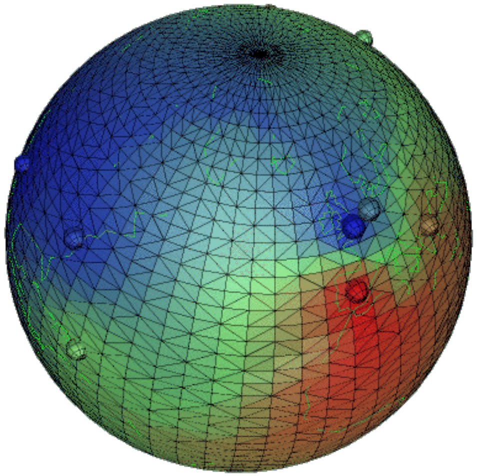
Traveltimes observed at various (at least 3) stations above the earthquake

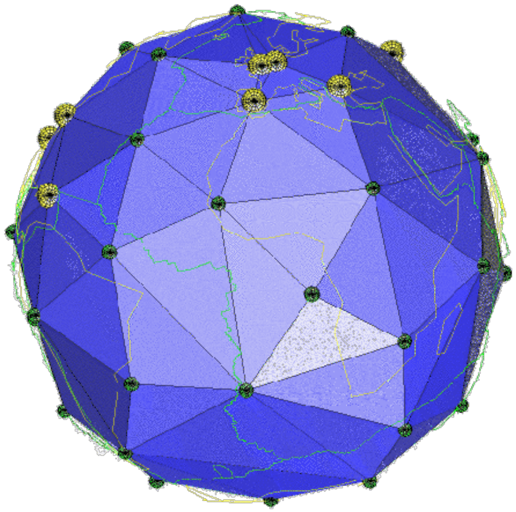
Model m:

3 coordinates of the earthquake location (x,y,z).

Usually much more data than unknowns: overdetermined system

Examples: Global Electromagnetism





Data vector d:

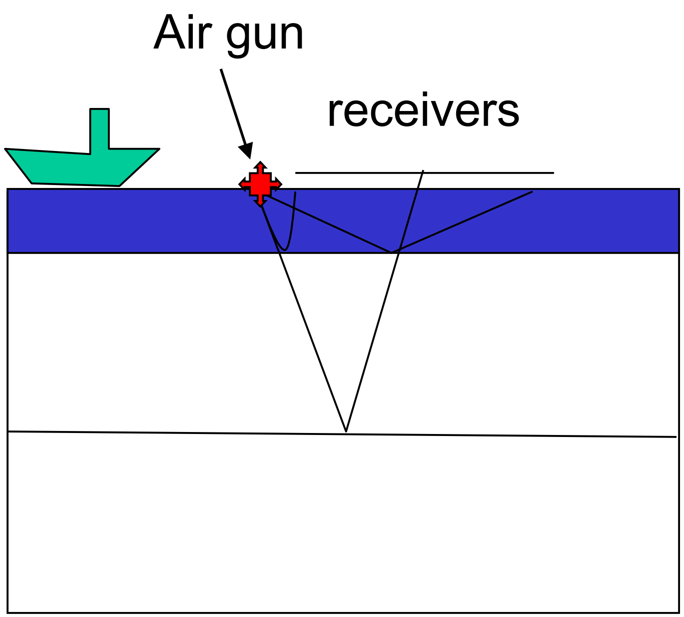
Amplitude and Phase of magnetic field as a function of frequency

Model m:

conductivity in the Earth’s mantle

Usually much more unknowns than data: underdetermined system

Examples: Reflection Seismology



Data vector d:

ns seismograms with nt samples

-> vector length ns*nt

Model m:

the seismic velocities of the subsurface, impedances, Poisson's ratio, density, reflection coefficients, etc.

Inversion: Summary

We need to develop formal ways of

1. calculating an inverse operator for
   $$d=Gm \;\; \rightarrow \;\; m=G^{-1}d$$
   (linear or linearized problems)

2. describing errors in the data and theory (linear and nonlinear problems)

3. searching a huge model space for good models (nonlinear inverse problems)

4. describing the quality of good models with respect to the real world (appraisal).