12.510 Introduction to Seismology
Spring 2008

✧ Go back to the observations again and look at the deviation from the model.

Objective: we need to find a model that minimize δt

T_{obs}= 3D structure + error in source location + noise (instrument, measurement)

(error in source location) → Hypocenter: to,x,y,z (or θ,φ,r)

$\Rightarrow \delta t = \delta t_{\mathbf{3D}} + \delta t_{\mathbf{mislocation}} + \delta t_{\mathbf{noise}}$

Errors caused by the model and the source location are usually combined together, and the noise is assumed to be white, with a Gaussian distribution.

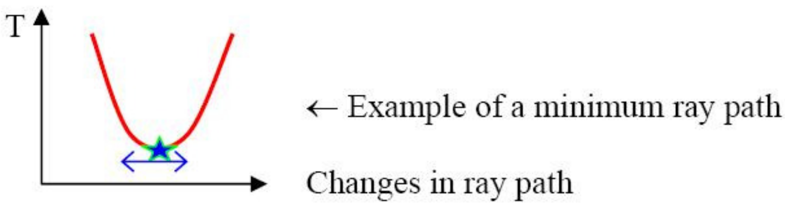
If we know the 1D model, we can apply **Snell's law** to estimate the geometry:

$$T_{obs} = \int_{\substack{3D \\ Ray \\ Path}} \frac{1}{c(\underline{x})} dl$$

Note that

$$\nabla T(x) = \frac{1}{c(x)} \underline{k}$$

If c change, the ray path changes. We end up with a nonlinear problem. Thus, we should try to linearize the inversion, using **the Fermat's Principle**.



If we change a little bit the ray path around the optimum, we'll end up with a small change in travel time. We have two kinds of "deviations" from the reference ray:

1. First contribution: effect of changes in velocity Δc.

2. Second contribution: effect of change in the ray path. Fermat's principle says that we can ignore it.

Linearization of the travel time

Travel time residual:

$$Observation = \delta t_{3D} = T_{obs} - T_{ref} = \int_{\substack{true\\3D\\structure}} \frac{1}{c(\underline{x})} dl - \int_{\substack{reference\\3D\\path}} \frac{1}{c_0(\underline{x})} dl_0$$

$$\overset{Fermat's\ \text{Principle}}{\simeq} \int_{\substack{reference\\3D\\path}} \frac{1}{c(\underline{x})} dl_0 - \int_{\substack{reference\\3D\\path}} \frac{1}{c_0(\underline{x})} dl_0$$

$$= \int_{\substack{reference\\3D\\path}} \frac{\Delta c}{c_0^2} dl_0 = \int_{\substack{reference\\3D\\path}} \left(s(\underline{x}) - s_0(\underline{x})\right) dl_0$$

$$= \int_{\substack{reference\\3D\\path}} \left(\Delta s(\underline{x})\right) dl_0$$

In linearizing the problem, we get rid of the unknown ray. We can do our calculation in a reference earth model.

- The travel time tomography is an iterative process:
- Create 1D model
- Ray tracing and get new rays in the model
- Update ray geometry
- Get the reference ray related to the 3D

$$\delta t = T_{obs} - T_{ref}\ (3D)$$

(The reference model does not have to be a 1D model.)

Linearization of the hypocenter mislocation

$$\delta t_{misloc} = \frac{\partial T}{\partial t_o} \delta t_o + \frac{\partial T}{\partial x} \delta x + \frac{\partial T}{\partial y} \delta y + \frac{\partial T}{\partial z} \delta z$$

$$\delta \otimes = \frac{\partial \otimes}{\partial \oplus} \delta \oplus$$

with t_0 the origin time, (x, y, z) the location of the earthquake.

Then we try to solve for Δs, δt, δx, δy, δz.

Inverse Problem

observations **model perturbation (m-m$_o$)**

$\delta t = T - T_o$

$$d_i = \int_{volume} G_i(\underline{r})\mu(\underline{r})d\underline{r}$$

sensitivity kernel/ Frechet derivatives

$$d\otimes = \frac{\partial \otimes}{\partial \oplus}\delta\oplus$$

Partial derivatives / Frechet derivatives

First, we need to discretize the problem, i.e. **parameterization**:

$$\mu(\underline{r}) = \sum_{k=1}^{M} \gamma_k h_k(\underline{r})$$

γ_k: weight

h_k: basis function

For example, plane wave summation:

$$\phi = \iiint \Phi\left(x, y, z, \omega\right) e^{i(\underline{k}\cdot\underline{x}-\omega t)} d\underline{k} d\omega$$

Then, we inject μ into the equation:

$$d_i = \sum_{k=1}^{M} \left\{ \int G_i(\underline{r}) h_k(\underline{r}) d\underline{r} \right\} \gamma_k$$

kernel projected on the basis function h

$$d_i = \sum_{k=1}^{M} A_{ik}\gamma_k \quad \text{for } i=1,2,\ldots, N \text{ data}$$

sensitivity matrix

kernel projected on the basis functions

projector

$$\underline{d} = \underline{\underline{A}}\,\underline{m} \qquad m = (\gamma_1, \gamma_2, \ldots, \gamma_M)$$

model (includes the values of the weight)

data

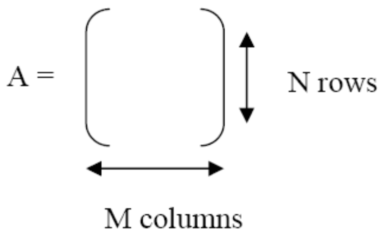
where

G_i : **Green functions**, solution of a point source. We need to do a convolution with a point perturbation in order to get the observations.

M : Number of model parameters.

N : Number of observations.

In general, $M \neq N$. As a consequence, the matrix $\underline{\underline{A}}$ is not square.



Multiplying the equation by A^T , we can get the solution:

$$\rightarrow \boxed{\hat{m} = (A^T A)^{-1} A^T \underline{d}} \ \textbf{Generalized Least Squares inversion}$$

$$\| \hat{m} - (A^T A)^{-1} A^T \underline{d} \| = \varepsilon \ \rightarrow \text{Goal is to } \textbf{minimize } \varepsilon.$$

Back to our specific inverse problem:

$$\delta t = \int \Delta s . dl \ \rightarrow \text{parametrize } \Delta s$$

$$\Delta s = \sum_{k=1}^{M} \gamma_k h_k$$

One way is to take h_k as a series of cells/blocks, with a value for $\underline{x}$ inside the cell k and zero otherwise. We have

$$\delta t_i = \int \Delta s dl = \sum_{k=1}^{M\ cell} \Delta s_k \left(dl \right)_{ik}$$

where

i : event-station pair.

$\left(dl \right)_{ik}$: path length.

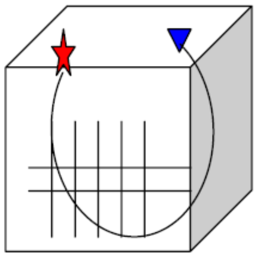
Rewrite the equation in the matrix form:

$$A_{ik} \cdot m = \begin{bmatrix} \Delta l_{11} & \dots & \Delta l_{1M} \\ \vdots & \vdots & \vdots \\ \Delta l_{N1} & \dots & \Delta l_{NM} \end{bmatrix} \begin{bmatrix} \Delta s_1 \\ \vdots \\ \Delta s_M \end{bmatrix} = \begin{bmatrix} \delta t_1 \\ \vdots \\ \delta t_N \end{bmatrix}$$

where each ray gives a row in the matrix.

We have an average wavespeed along the ray. In order to construct a model vector, we need to get data from different rays crossing each other.

A is a sparse matrix. If we look at one ray:



M~300,000
~ 20 layers
the ray samples ~ 100 cells

A will have only ~100 elements non-zero. The good thing about sparse matrix is that $A^T A$ is approximately diagonal. The problem is that there are many singularities, which make the inversion unstable (in that case, we need to add a damping factor or regularize the problem). One possibility is to not use cells of the same size. Consequently, it reduces the number of cells; the inverse matrix is less singular. Nevertheless, the computation time increases.

Another way is take h_k as spherical harmonics (in global seismology).