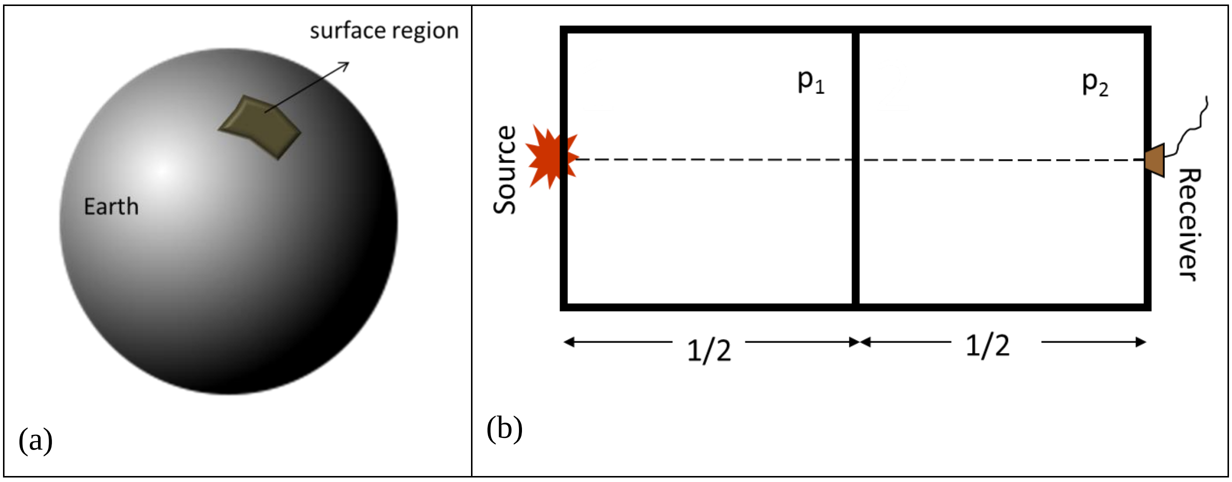
Computational Exercise #4

Type of Activity: The least squares solution in the matrix form when the forward mode does not provide enough information to determine uniquely all the parameters. **The problem is said to be underdetermined.** Notice that this computational exercise resembles the item 4.1 of the theoretical exercise 1.

Geologic problem:

Let us assume a surface region on the Earth (Figure 1a) with unknown physical-property distribution. We wish estimate the physical-property distribution in this region from a simplified tomography (Aster et al., 2012). Let's assume that geometric ray theory (essentially the high-frequency limiting case of the wave equation) is valid. Hence, in this simplified tomography the wave energy traveling between a source and receiver can be considered to be propagating along infinitesimally narrow ray paths. In seismic tomography, if the slowness at a point $\boldsymbol{x}$ is $s(\boldsymbol{x})$ and the ray path ℓ is known, then the travel time for seismic energy transiting along that ray path is given by the line integral along ℓ:

$$t = \int_{\ell} s(\boldsymbol{x}(\ell))d\ell \qquad (4.1)$$



From the travel time measurements, we wish to estimate the slowness distribution of seismic waves on the surface region (Figure 1a). A common way of discretizing the model in a tomographic problem is as uniform blocks (Figure 1b). This approach is essentially applying the midpoint rule to the travel-time forward problem. Consider that the surface region in Figure 1a can be discretized in two blocks.

Simplifications:

a) We assume that the two blocks shown in the Figure 1b will be sufficient to describe the slowness variations on the surface region

b) We assume that the distribution of the physical property (slowness) is constant within a block. However, slowness variations among blocks are allowed.

c) The ray refraction effects are disregarded

Forward model

Consider that the two blocks (Figure 1b) with sides of 1/2 length and unknown slowness (p_1 and p_2) are crossed by one (1) ray path.

The total travel time to the wave energy travel between the source and receiver (Figure 1b) will be the sum of the travel time spent by the wave to cross each block found in its path. In the case illustrated in Figure 1b above we have:

$$t_1 = \frac{1}{2}p_1 + \frac{1}{2}p_2 \qquad (4.2)$$

Mathematical problem: We want to estimate p_1 and p_2 from ONE observed travel time t_1. Here, the data is a single measure (N=1) of the travel time (t_1).

The forward model (equation 4.2) can be expressed, in matrix notation, as

$$\bar{\mathbf{y}} = \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}} \quad , \qquad (4.3)$$

where $\bar{\mathbf{y}}$ is an N-dimensional vector of the exact data (N = 1), $\bar{\mathbf{p}}$ is a M-dimensional vector of parameters (M=2) given by

$$\bar{\mathbf{p}} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}, \qquad (4.4)$$

where p_1 and p_2 are the slownesses (Figure 1b), and $\bar{\bar{\mathbf{A}}}$ is the N × M sensitivity matrix given by

$$\bar{\bar{\mathbf{A}}} = [1/2 \quad 1/2]. \qquad (4.5)$$

We wish to estimate the unknown slownesses.

When the forward mode (equation 4.3) does not provide enough information to determine uniquely all the parameters, the problem is said to be **underdetermined**. Underdetermined problems occur when there are more unknown parameters than data, that is, M > N.

To estimate the parameter vector $\bar{\mathbf{p}}$**,** if we have M > N = rank ($\bar{\bar{\mathbf{A}}}$), we can solve the following constrained problem of

$$\begin{cases} \text{minimizing} \quad \bar{\mathbf{p}}^{\mathrm{T}}\, \bar{\mathbf{p}} \\ \text{subject to:} \quad \bar{\bar{\mathbf{A}}}\, \bar{\mathbf{p}} = \bar{\mathbf{y}}^{\mathbf{o}} \quad , \end{cases} \tag{4.6}$$

where $\bar{\boldsymbol{y}}^{\boldsymbol{o}}$ is an N-dimensional vector of the noise-corrupted data (observed data set).

To solve this constrained problem, we employ the Lagrange multipliers method and minimize the unconstrained objective function:

$$\Gamma = \bar{\mathbf{p}}^{\mathrm{T}}\, \bar{\mathbf{p}} + \left(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\bar{\mathbf{A}}}\, \bar{\mathbf{p}}\right)^{\mathrm{T}} \bar{\boldsymbol{\lambda}}, \tag{4.7}$$

where $\bar{\boldsymbol{\lambda}}$ is a N-dimensional vector of the Lagrange multipliers.

The solution of minimizing the unconstrained function Γ with respect to the unknown parameters $\bar{\mathbf{p}}$ is:

$$\hat{\mathbf{p}} = \bar{\bar{\mathbf{A}}}^{\boldsymbol{T}} \left(\overline{\overline{\mathbf{A}\bar{\bar{\mathbf{A}}}^{\boldsymbol{T}}}}\right)^{-\mathbf{1}} \bar{\boldsymbol{y}}^{\boldsymbol{o}} \quad , \tag{4.8}$$

I usually call this estimator (equation 4.8) as Least Square solution of an underdetermined problem (M > N = rank($\bar{\bar{\mathbf{A}}}$)). Menke (1989) calls this estimator as the solution of purely underdetermined problem.

(1) Generate a synthetic test:

(1.1) Assign numeric values to the slownesses ($p_1$and p_2). These values are the true parameters

(1.2) Compute a noise-free data (exact data) by using equation 4.3 and the true parameters.

(1.3) Generate the noise-corrupted data (observed data). Generate one pseudorandom sequence of noise with Gaussian distribution, null mean and standard deviation of 1% of the noise-free data.

(1.4) Estimate the underdetermined least squares solution for the model parameters (p_1 and p_2) by using equation (4.8).

(1.5) Compute the predicted data

(1.6) Compute the residual vector given by the differences between observed data and the predicted data.

(2) Visualization of the nonuniqueness

(2.1) Plot in the space $p_1$versus p_2 the following function

$$Q = \|\bar{\boldsymbol{\varepsilon}}\|_2^2 = \bar{\boldsymbol{\varepsilon}}^{\boldsymbol{T}}\bar{\boldsymbol{\varepsilon}}$$

where $\bar{\boldsymbol{\varepsilon}}$ is an N-dimensional residual vector (N=1): $\bar{\boldsymbol{\varepsilon}} = \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}} - \overline{\boldsymbol{y}}^{\boldsymbol{o}}.$ In this graphic, plot the true and estimated parameters.

(2.2) Plot the true parameters in the plot of the function Q (item 2.1)

(2.3) The function Q (item 2.1) has multiples minima, showing that any pair of parameters that falls on the minima of function Q satisfies exactly the equation $\overline{\mathbf{y}^{\boldsymbol{o}}} = \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}$. The estimated parameter vector by using equation (4.8) (see item 1.4) is a particular solution close to the minimum of the function

$$\phi = \|\bar{\mathbf{p}}\|_2^2 = \bar{\mathbf{p}}^{\boldsymbol{T}}\bar{\mathbf{p}}.$$

Plot the function ϕ in the space $p_1$versus p_2 . In this graphic, plot the minimum of the function $\phi,$ the true and estimated parameters.

(2.4) Is the solution (estimated parameter vector) stable? Justify your response.

HINT: To justify your response, you can apply the practical procedure to analyze the stability of the solution in the presence of noise. This practical procedure was presented in the Computational Exercise #2.

(2.5) Is the estimated parameter vector close to the true one? If the estimated parameter vector is not close to the true one, repeat the test by using the exact data. Can you explain these results?

(2.6) Can the least-square solution (equation 4.8) retrieve the true parameters? If it is true, can you explain, what is the relationship between the true and estimated parameters?