Computational Exercise #5

Type of Activity: We focus our attention on inversion methods that stabilize the solution by transforming the ill-posed problem into a well-posed one via the Tikhonov regularization method (Tikhonov and Arsenin, 1977). Here, we address two regularizing functions. The first one is named the **zeroth-order Tikhonov regularization** and the second regularizing function is the **first-order Tikhonov regularization**.

Geologic problem:

Let us assume a surface region on the Earth (Figure 1) 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:

$$t = \int_{\ell} s(\boldsymbol{x}(\ell))d\ell \qquad (5.1)$$

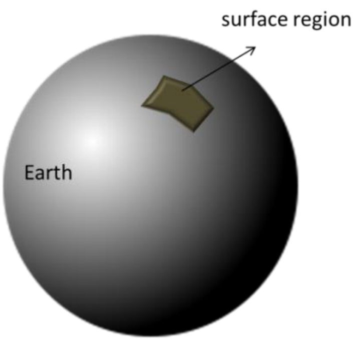

Figure 1 – Surface region on the Earth with unknown physical-property distribution.

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

Simplifications:

a) We assume that the two blocks 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 - By assuming that the two blocks (Figure 2) will be sufficient to describe the slowness variations on the surface region shown in Figure 1. Consider that these two blocks have unknown slowness (p_1 and p_2) are crossed by two ray paths.

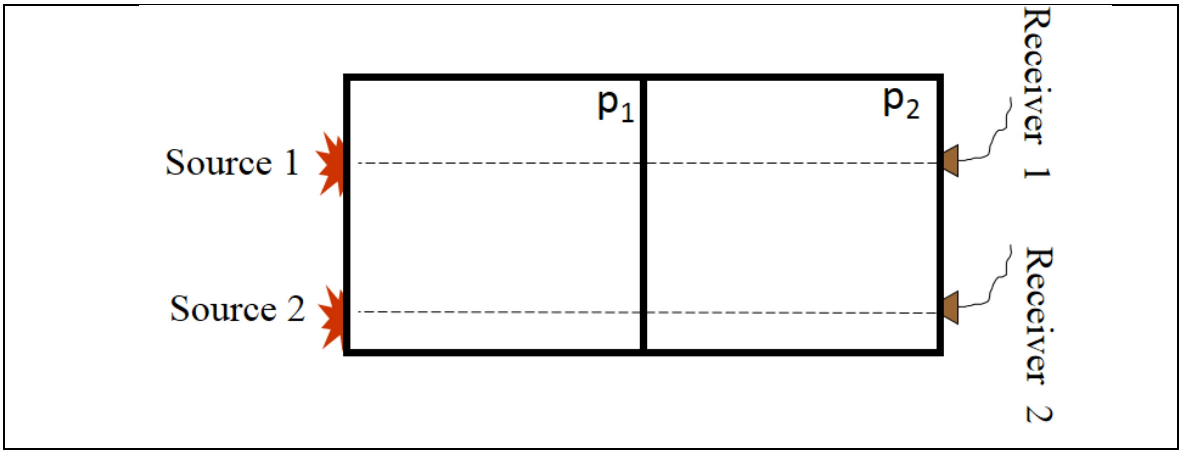

Figure 2 – **Forward model**. The two blocks will be sufficient to describe the slowness (p_1 and p_2) . The first total travel time to the wave energy travel between the source 1 and receiver 1 is written by equation 5.2a. The second total travel time to the wave energy travel between the source 2 and receiver 2 is written by equation 5.2b.

The first total travel time to the wave energy travel between the source 1 and receiver 1 (Figure 2) 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 2, the first total travel time is:

$$t_1 = 2.0\,p_1 + 1.001\,p_2. \qquad (5.2a)$$

The second total travel time to the wave energy travel between the source 2 and receiver 2 (Figure 2) 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 2, the second total travel time is:

$$t_2 = 2.0\, p_1 + 1.0\, p_2. \tag{5.2b}$$

Mathematical problem: We want to estimate p_1 and p_2 from two observed travel times (t_1 and t_2). The forward model (equations 5.2a and 5.2b) can be expressed, in matrix notation, as

$$\bar{\mathbf{y}} = \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}} \quad , \tag{5.3}$$

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

$$\bar{\mathbf{p}} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}, \tag{5.4}$$

where p_1 and p_2 are the slownesses (Figure 2), and $\bar{\bar{\mathbf{A}}}$ is the N × M sensitivity matrix given by

$$\bar{\bar{\mathbf{A}}} = \begin{bmatrix} 2.0 & 1.001 \\ 2.0 & 1.0 \end{bmatrix}. \tag{5.5}$$

We wish to estimate the unknown slownesses.

By solving the constrained problem of

$$\begin{cases} \text{Minimizing} \quad \phi(\bar{\mathbf{p}}) \\ \text{subject to} \quad \|\bar{\boldsymbol{\varepsilon}}\|_2^2 = \delta \end{cases} \tag{5.6}$$

where δ is the expected the expected mean square of the noise realizations in the data, $\bar{\boldsymbol{\varepsilon}}$ is an N-dimensional residual vector defined as the difference between the observed data and the predicted data, i.e.,

$$\bar{\boldsymbol{\varepsilon}} = \bar{\boldsymbol{y}}^{\boldsymbol{o}} - \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}, \tag{5.7}$$

and $\phi(\bar{\mathbf{p}})$ is the regularizing function which is defined in the parameter space (model space) that imposes physical or geological attributes on a solution.

To solve this constrained problem (equation 5.6), we employ the Lagrange multipliers method and minimize the unconstrained objective function:

$$\Gamma(\bar{\mathbf{p}}) = \mu\, \phi(\bar{\mathbf{p}}) + \bar{\boldsymbol{\varepsilon}}^T \bar{\boldsymbol{\varepsilon}} \tag{5.8}$$

where μ is the regularizing parameter that balances the relative importance between the data-misfit measure and the regularizing function.

Two regularizing functions commonly used in geophysics are the zeroth- and first-order Tikhonov regularizations. Both regularizing functions impose a smooth character on the estimated parameter.

The zeroth-order Tikhonov regularization is given by

$$\phi(\bar{\mathbf{p}}) = \bar{\mathbf{p}}^T \bar{\mathbf{p}} \qquad (5.9a)$$

In the literature, this constraint is also known as minimum Euclidean norm of the parameters or ridge regression method (Hoerl and Kennard, 1970). The zeroth-order Tikhonov regularization (equation 5.9a) imposes that all estimated parameters must be as close as possible to zero.

The first-order Tikhonov regularization is given by

$$\phi(\bar{\mathbf{p}}) = \bar{\mathbf{p}}^T \bar{\bar{\mathbf{B}}}^T \bar{\bar{\mathbf{B}}} \bar{\mathbf{p}} \qquad (5.9b)$$

where the L x M matrix $\bar{\bar{\mathbf{B}}}$ represents the first-order discrete differential operator (Twomey, 1996; Constable et al., 1987; Aster et al., 2004) whose rows contain only two nonnull elements, 1 and -1. Notice that the L-dimensional vector $\bar{\bar{\mathbf{B}}}\bar{\mathbf{p}}$ is a finite-difference approximation that is proportional to the first derivative of $\bar{\mathbf{p}}$. The first-order Tikhonov regularization will favor solutions that are relatively flat (Aster et al., 2004). In the literature, this constraint is also known as minimum structure solution, smoothness solution or flatness solution. The first-order Tikhonov regularization (equation 5.9b) imposes a smoothing character on the solution.

The zeroth-order Tikhonov regularization solution is given by

$$\hat{\mathbf{p}} = \left(\bar{\bar{\mathbf{A}}}^T \bar{\bar{\mathbf{A}}} + \boldsymbol{\mu} \bar{\bar{\mathbf{I}}}\right)^{-1} \bar{\bar{\mathbf{A}}}^T \ \bar{\boldsymbol{y}}^{\boldsymbol{o}} \qquad (5.10a).$$

The first-order Tikhonov regularization solution is given by

$$\hat{\mathbf{p}} = \left(\bar{\bar{\mathbf{A}}}^T \bar{\bar{\mathbf{A}}} + \boldsymbol{\mu} \bar{\bar{\mathbf{B}}}^T \bar{\bar{\mathbf{B}}}\right)^{-1} \bar{\bar{\mathbf{A}}}^T \ \bar{\boldsymbol{y}}^{\boldsymbol{o}} \qquad (5.10b)$$

Here, the matrix $\bar{\bar{\mathbf{B}}} = [1 \quad -1]$.

(A) Generate a synthetic test in the case of **the forward model (Figure 2):**

(A.1) Assign numeric values to the slownesses ($p_1$and p_2). These values are the true parameters

(A.2) Compute a noise-free data (exact data)

(A.3) Generate the noise-corrupted data (observed data). Generate pseudorandom sequence of noise with Gaussian distribution, null mean and standard deviation of 1% of the absolute value of the average of the noise-free data.

(B) Compute the L-curve (Hansen, 1992)

(B.1) Plot the zeroth-order Tikhonov regularization L-curve to pick the regularization parameter by a visual inspection.

(B.2) Plot the first-order Tikhonov regularization L-curve to pick the regularization parameter by a visual inspection.

(C) Compute the solution

(C.1) By using the chosen regularization parameter in the item B.1, compute the zeroth-order Tikhonov regularization solution (equation 5.10a). Discuss your result.

(C.2) By using the chosen regularization parameter in the item B.2, compute the first-order Tikhonov regularization solution (equation 5.10b). Discuss your result

(**D**) Plot, in the space $p_1$versus p_2 , the following graphics:

(D.1) the data-misfit function, the true and estimated parameters by using zeroth-order Tikhonov regularization

(D.2) the zeroth-order Tikhonov regularization function ($\phi(\bar{\mathbf{p}})$, given by equation 5.9a), the true and estimated parameters obtained using zeroth-order Tikhonov regularization.

(D.3) the unconstrained objective function $\Gamma(\bar{\mathbf{p}})$ (equation 5.8) with the zeroth-order Tikhonov regularization function ($\phi(\bar{\mathbf{p}})$, given by equation 5.9a), the true and estimated parameters obtained using zeroth-order Tikhonov regularization.

(**E**) Repeat the item D changing the regularizing function by the first-order Tikhonov regularization.

References:

Aster, R. C., B. Borchers, and C. H. Thurber, 2004, Parameter estimation and inverse problems: Elsevier Academic Press.

Constable, S. C.,R.L. Parker, and C. G. Constable, 1987, Occam's inversion: Apractical algorithm for generating smooth models from electromagnetic sounding data: Geophysics, 52, 289–300.

Hansen, P. C, 1992, Analysis of discrete ill-posed problems by means of the L-curve. SIAM Review, 34(4):561–580.

Hoerl, A. E., and Kennard, R. W., 1970, Ridge regression: Biased estimation for nonorthogonal problems: Technometrics, 12, 55–67.

Tikhonov, A. N. and Arsenin V. Y. Solutions of Ill-Posed Problems. Halsted Press, New York, 1977.

Twomey S, 1996, Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements. Dover, Mineola, NY.