PROBABILISTIC INVERSE THEORY. ASSIGNMENT 2 – GRAVITY PROFILE ACROSS A GLACIER

J. Alejandro Duran, ETH November 2017

1 Objective

Use a metropolis algorithm in order to solve the inverse problem of finding the thickness of a glacier from measurements of the gravitational anomaly it produces.

2 The inverse problem

We want to estimate the thickness of the glacier $h(x)$ at a given $x-$point. First we discrete the problem. We already know the solution of the forward problem:

$$\Delta g(x_j) = G\Delta\rho(x)\int_0^a \ln[\frac{(x-x_j)^2+h^2(x)}{(x-x_j)^2}]dx. \tag{1}$$

This equation gives the solution of the gravity anomaly at any of the measurement points. They are a set of 12 points X_j over the glacier where the measurements were performed.

For each measurement x_j we need to integrate along the x-axis. We need then to sample this axis. Let's name N as the number of samples I want to use, therefore each sample will be separated a distance $\Delta x = \frac{x_{max}-x_0}{N-1}$. Each point in my sampling is defined as $x_{i+1} = x_i + \Delta x$

When I include this discretization in equation 1 I end up with the following expression:

$$\Delta g(x_j) = G\Delta\rho(x_i)\sum_{i=1}^{N} \ln[\frac{(x_i-x_j)^2+h^2(x_i)}{(x_i-x_j)^2}]\Delta x. \tag{2}$$

An important issue is the treatment of the denominator in the last equation. In reality, we must satisfy $x_i - x_j \neq 0\ \forall i,j$. We can do it by including a small factor ϵ in the denominator and avoid the singularity.

The problem to solve can be described as $\mathbf{d} = \mathbf{G(m)}$. Where $\mathbf{d}$ is our data (gravity anomaly) and $\mathbf{G(m)}$ is given by equation 2. The inverse problem consist in finding the model $\mathbf{m}$ (a set of N values $h(x_i)$) which satisfies our data. The problem is however non linear, the logarithm does not allow a linear relation between data and model parameters.

3 Setting a metropolis algorithm

In our current case, the metropolis algorithm will consist essentially in the next steps:

- Generating several solutions $\mathbf{m}_k$
- For each solution I will compute the prior $\rho(\mathbf{m}_k)$ and the likelihood $L(\mathbf{m}_k)$
- I compute the posterior $\sigma(\mathbf{m}_k) = \frac{\rho(\mathbf{m}_k)L(\mathbf{m}_k)}{\mu(\mathbf{m}_k)}$

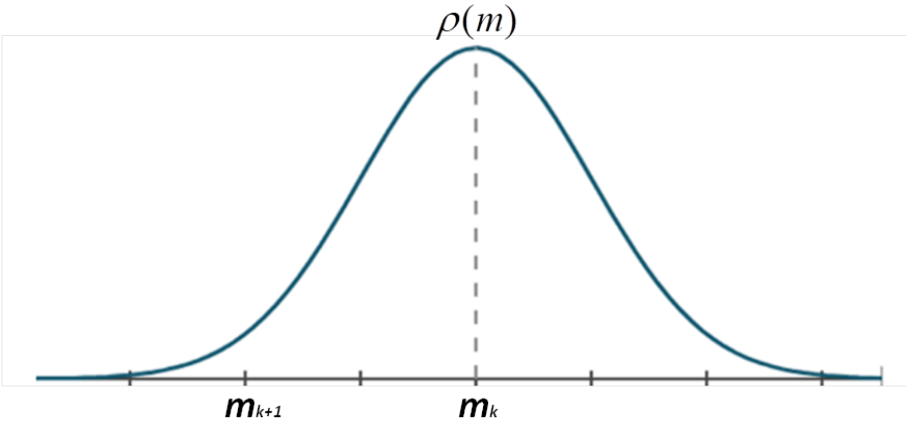

Figure 1: New model $\mathbf{m}_{k+1}$ selected based on a probability density $\rho(\mathbf{m})$ centered at the current model $\mathbf{m}_k$.

- The model $\mathbf{m}_k$ is comparing with the previous model. It is accepted only if it satisfies:

$$\alpha < \frac{\sigma(\mathbf{m}_k)}{\sigma(\mathbf{m}_{k-1})} \tag{3}$$

where α is a random number such that $\alpha \in [0, 1)$

3.1 The initial model

The initial model $\mathbf{m}_0$ is set by using the gravity response of the Bouguer plate. For a thickness h, it is defined as:

$$\Delta g_B = 2\pi\rho G h, \tag{4}$$

where G is the gravitational constant and ρ the mass density.

For a given measurement $\Delta g(x_j)$ (which are known values) we can obtain the first model (set of points $h(x_j)$) by operating equation 4.

3.2 Generating models

After starting with the first model, we will use the metropolis criteria to evaluate if a given model is admitted or rejected. The question to answer now is how to generate new models. If the models are absolutely random, the time to converge to the adequate solution could be too large. Then, a better approach is to test each new model based on a probability distribution around the current model. This is represented graphically in figure 1 (the figure used an arbitrary Gaussian distribution just for didactic proposes). The new model $\mathbf{m}_{k+1}$ is selected based on the distribution $\rho(\mathbf{m})$ centered at the current model $\mathbf{m}_k$.

Given the initial model, we should compute our prior. Then, as mentioned in the exercise, the *a priori* variances should be set up. This variance is in practice the square of the standard deviation that each value $h(x_i)$ is allow to have. However, the thickness between the interval $x \in (0, a)$ should be non zero. Therefore, I established the following constraint for the prior:

$$h(x) - \sigma_{prior} > 0 \quad \forall x \in (0, a) \tag{5}$$

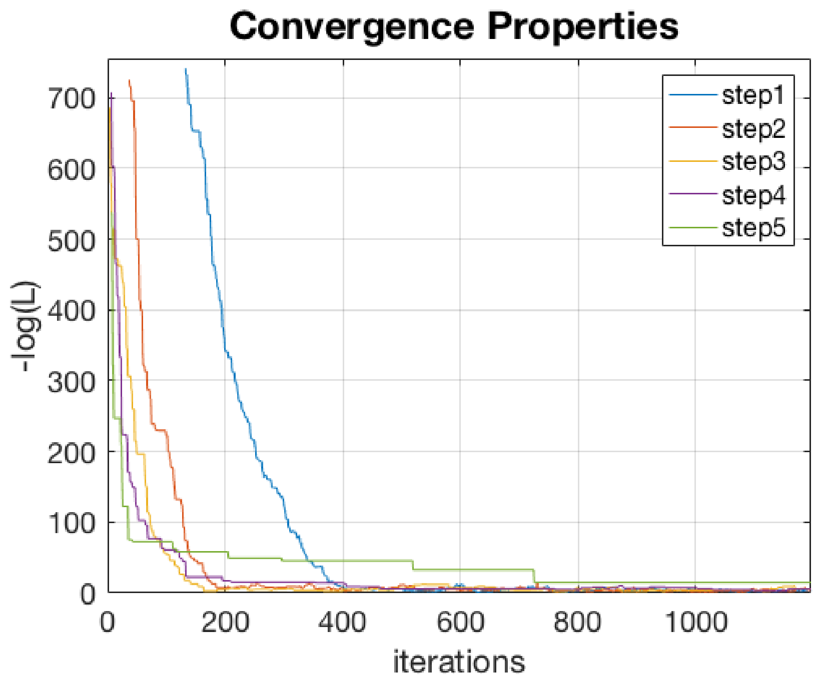

Figure 2: Analysis of the convergence by increasing the step size (standard deviation of the prior) from which a new model is proposed. Steps 1 to 5 go in an increasing order.

The convergence of the models depends on the step size I use. In figure 1 for instance, the new model $\mathbf{m}_{k+1}$ is selected for a given step size. This can be seen as the standard deviation of the prior. If a choose an extreme large step, the proposed model can be very far from the current one, and therefore many of the proposed models will not full-fill the metropolis criteria (step 5 in figure 2). On the other hand, if the step is too small, the algorithm will quickly find a model which full-fill the criteria but I will have to wait for a long time to get the best model which minimize the misfit:

$$-\log(L(\mathbf{m})) \propto (\mathbf{d}_{obs} - \mathbf{G}(\mathbf{m}))^2$$

In other words, the convergence is obtained but after more number of iterations (step 1 in figure 2). These behaviours is observed in figure 2, where the convergence is studied for different steps sizes.

Figure 3 shows a comparison with the observed data and the data obtained from the model which maximizes the posterior, i.e., the best model found such that the misfit between the observed and synthetic data is reached. The number of points are different (the synthetic data had 20 points and the observed only 12) but both plots agree. This increase our confidence that the solution of the inverse problem makes sense.

Several models are found that satisfied the posterior distribution. Figure 4 shows the average solution of the model which maximizes it. For a given model parameter, it shows a histogram of all possible values which full-filled the metropolis criteria. Each histogram can be interpreted as a probability distribution, both will be the same when the sampling goes to infinity.

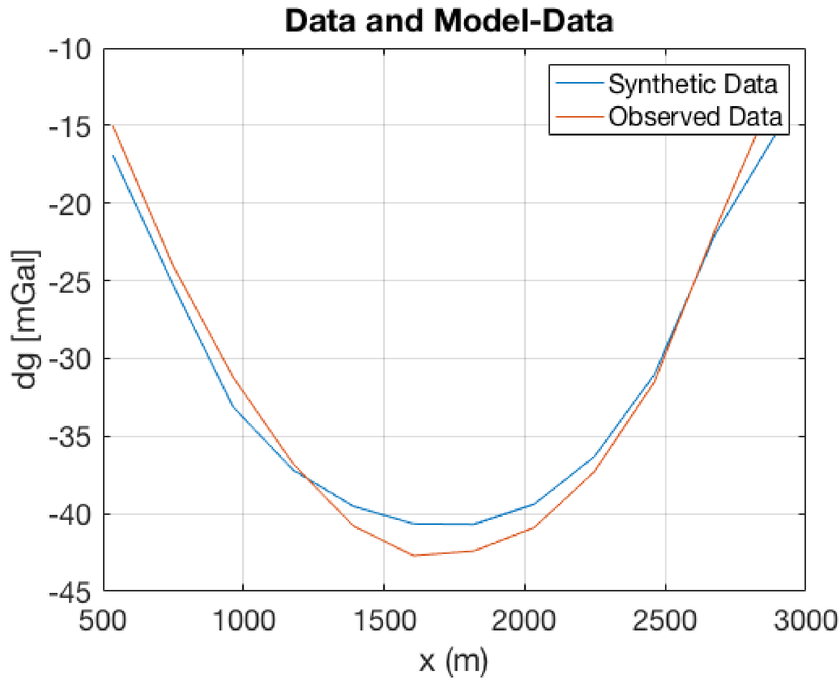

Figure 3: Comparison of the observed data and the synthetic data computed from the inverse model.

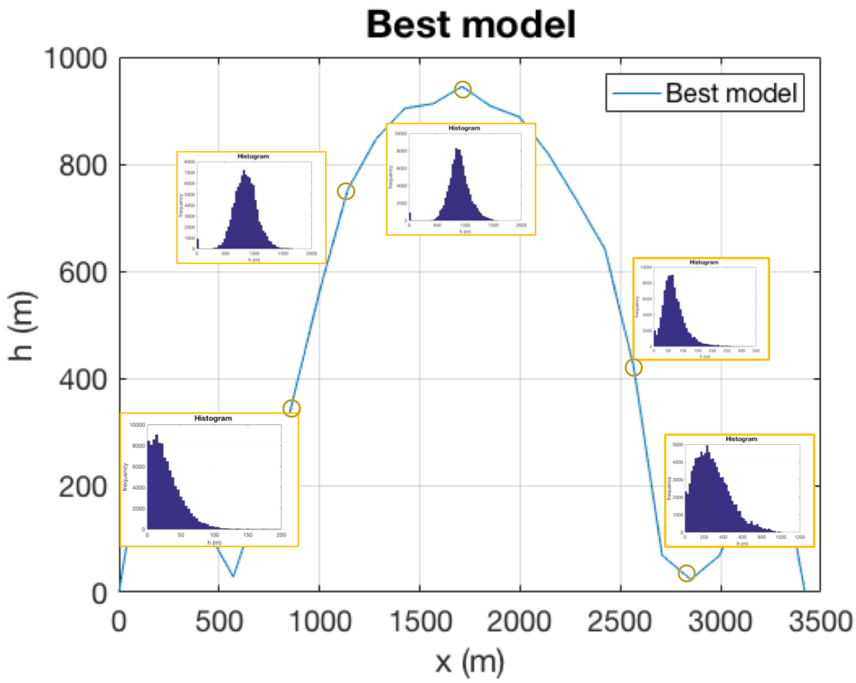

Figure 4: Average of the best model and histograms of showing the distribution of various model parameters.