GPGN 410 / 510

Due Friday, April 11th, 2025

Homework 2

Gauss-Newton & Nonlinear Optimization

Please feel free to abandon this specific prompt and be creative. Just show me that your creative spin utilizes nonlinear optimization in some manner. For example… Are you doing research? Show me an application of nonlinear inversion in your research! Are you interested in speeding up GN? Show me Gauss-Newton Conjugate Gradient! Are you interested in large scale problems? Use run the experiments with inexact Newton methods. Using complicated model objective functions? I want to see.

You are welcome to follow the prompt below; however, I would encourage you to write up your research/personal interests instead. As a grader, I would much rather read about something you enjoy doing. Just show me that what you are interested in is relevant to nonlinear optimization.

# Introduction

Gauss Newton optimization is an example of a first order optimization method – the Hessian is approximated via first derivatives, and second order derivatives are never explicitly calculated. If you need a review, I suggest beginning at slide 114 of the class material.

As an example, problem, we will invert a profile of observed gravity data for basin depth. Not interested in gravity? *Pick your own problem*.

As in the last assignment, I am not interested in your favorite computational linear algebra library, your favorite programming language, or your favorite operating system. Please demonstrate your understanding with words and pictures.

# Motivation: Geophysical Investigation of a Rift Zone

Today the West Africa rift zone is inactive, buried beneath twenty kilometers of sediment. A piece by Fairhead and Okereke ([1987](https://www.sciencedirect.com/science/article/pii/0040195187900849)) has captured your attention – after measuring gravity across the rift zone, Figure 10 of their work is a rudimentary density inversion (Figure 1 of this document). You are at once intrigued, and question, “Could I verify these results?”

This area is ripe with geophysical information – local seismic profiles tell us the depth to the top of the rift to 23 kilometers below the surface. Off the sides of the rift zone, the same people who ran the seismic refraction lines tell us the depth to bedrock is 34 kilometers. Fairhead & Okereke used this information to conclude that the density contrast in the rift zone is 0.17 grams per cubic centimeter. They remark this density contrast is anomalously low. You will challenge their findings (in a good natured, scientific manner) and determine whether or not this is so.

A diagram of a graph

AI-generated content may be incorrect.

Figure : This is figure 10 from Fairhead and Okereke, 1987, the motivation for our work today. On top is a measured gravity anomaly (labeled “OBSERVED”) and the forward modeled, or predicted, gravity (labeled “COMPUTED”) that the authors computed using the simple model in the lower diagram. The rift system is highlighted along the profile, stretching from about 175 kilometers to 250 kilometers.

# The Data, Model, and the Forward Operator

To get you started, I extracted the measured gravity profile from their paper. This is your data. There is no true model to generate the data from, because we are working with real data – the truth is unknown. There is probably noise in the data, but it is unknown how much noise there is.

I will get you started inverting for the rift zone geometry. We will define some prisms at a depth of 34 kilometers and increase their height such that the prisms grow towards the surface. Our hope is that the prisms will mimic the behavior of the geologic rift system. The heights of the individual prisms will be your model. I will describe this more below.

To map from the model space to the data space, we use a forward operator. In this case, we will use rectangular prisms of infinite horizontal extent to map density into vertical gravity anomalies. LaFehr & Nabighian ([2012](https://library.seg.org/doi/book/10.1190/1.9781560803058)) give an expression for the forward operator of our problem on page 38 of their wonderful little book “Fundamentals of Gravity Exploration”:

,

labeled equation 22 by the authors. I will verbosely write all the parameters below but take a glance at Figure 2 for an intuitive view of what is going on. The constant is the fundamental gravity constant ( in SI units) and is the density (170 kg/m3 in SI units). The horizontal edges of the prism are given by and , with , in SI units (meters). Finally, the vertical extent is determined by and , with , positive “down”, in SI units (meters). For computational purposes, I recommend you use the [two argument arctangent function](https://en.wikipedia.org/wiki/Atan2) when implementing this.

A diagram of a square with lines and arrows

AI-generated content may be incorrect.

Figure : The visual description of the forward model. Infinite horizontal prism. A picture of Figure 14 of "Fundamentals of Gravity Exploration", taken with Nate’s phone.

So, the model space is the difference between and . The larger the density contrasts, the larger the prism. Because prism height and gravity interact in a nonlinear fashion, we will implement a Gauss Newton inversion to solve for the rift system.

# Setting up the Optimization

Now let us prepare the data! I have digitalized one of the profiles from Fairhead and Okereke ([1987](https://www.sciencedirect.com/science/article/pii/0040195187900849)), displayed in Figure 1. The file “profile.txt” contains the data I digitalized for you. In the first column is distance along the transect in kilometers. In the second column we have the complete Bouguer anomaly (minus the regional trend) in mGals. Although these are customary units, do not invert with them! Transfer all the units into the SI system: kilometers to meters, and mGals to meters/s2. Compare your results to my Figure 3.

A graph of a graph

AI-generated content may be incorrect.

Figure : The raw data, in SI units. Notice I have not used traditional units for the gravity anomaly and distance along the profile (mGals and kilometers respectively) so that you can compare your unit conversions to mine. When you invert the data, you must use a consistent system of units to get physically meaningful answers!

Now let us set up the model. You can be creative with this, but I am going to keep things simple for now. We will line up prisms in a row under the gravity profile, like carts in a train, or links in a sausage. We will invert for the height of each of these prisms, or carts, or sausages… Ok, enough with the analogies. Let the model be a thousand prisms wide, stretching from 0 meters to 462,500 meters. So, each prism will have a width of 462.5 () meters. We will place thjjjjjhe row of prisms at a depth () of 34 kilometers. We will invert for the height of the prism, .

Because gravity fields are linear, we can add up the contribution of each model cell independently. Therefore, to compute full forward model, simply plug in the height of each individual prism, compute the individual prisms individual gravity field, and add up the gravity fields. The sum of all the gravity fields is equivalent to the gravity field of all the prisms. Linearity of the sources makes life easy!

# Gauss-Newton Optimization

Once you define the forward operator , we have enough information to begin computing the Jacobian , or sensitivity, or whatever the hell you want to call this operator. It’s just a derivative of the forward model with respect to the model parameters. There are many was to calculate this – analytic derivations, finite differences, ever automatic differentiation. We will stick to finite difference approximations for simplicity:

Here, is the ’th column of your Jacobian, and the forward model is a function of the prism height . The critical term in the finite difference approximation is the , which we can call the perturbation. You can try different values, but I suggest something small – maybe 100 meters, or 200 meters?

Dr. Darrh, a previous student of Professor Li’s, called this the “brute force” method to compute the Jacobian. You can do nicer things in your own work. Now let’s set up the rest of the Gauss-Newton iteration.

The gradient of the data misfit term, the gradient of the model objective function, and the gradient of the total objective function are given respectively as

Here the represents your model of prism heights. The above is not the gravity – it is the gradient of the objective function with respect to the model . The Jacobian has a subscript to indicate that it is the Jacobian at the current model iteration. Every time you update , you also must update . That can get expensive!

Now you can solve the linearized Gauss-Newton system for an update to your model :