Faculty of Arts and Sciences
Department of Computer Science
CMPS 351 – Numerical Optimization
Assignment 4 – Due Feb 18, 2019

Reading Material:

- Section 10.2, Chapter 11

1. Cross-well Tomography. Inverse problems are problems where we seek to measure the parameters of a physical system from measurements of its response to one or more inputs. This is in contrast to forward problems where, given a physical system, we seek to determine its response to a given input.

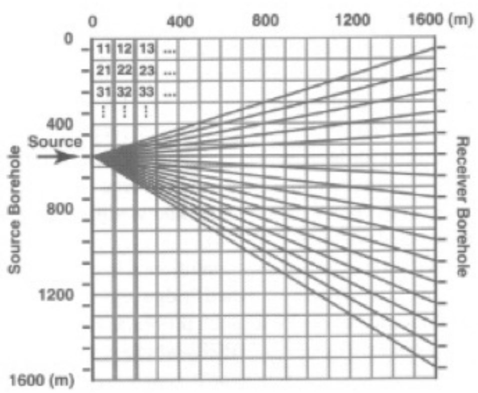
Consider, as an example, the following cross-well tomography problem used in petroleum exploration. Two vertical wells are located 1600 m apart. A seismic source is inserted in one of the wells at depths of $50, 150, \cdots, 1550$ m. A string of receivers is inserted in the other well at depths of $50, 150, \cdots, 1550$ m as shown in the figure below. For each source-receiver pair a travel time is recorded (such measurements, while accurate, have an error on the order of 0.5 ms). There are 256 ray paths and 256 corresponding measurements. We wish to determine the velocity structure in the two-dimensional plane between the two wells.

Discretizing the problem into a 16×16 grid ($100m \times 100m$ blocks) gives 256 model parameters. The coefficient matrix and noisy measurement data for a problem are provided in data files.
Use `d = np.load('d.npy')` and `G = np.load('G.npy')` to load the d vector and G matrix respectively.

- Confirm that the coefficient matrix is ill-conditioned and hence cannot be inverted to find the solution of $Gx = d$.
- Formulate the problem as a (linear) least squares problem, and confirm that this formulation cannot recover the problem parameters.
- Formulate the problem as a regularized linear least squares problem, that minimizes the sum of the misfit and a regularization term

$$\underset{x}{\text{minimize}} \quad \tfrac{1}{2}\|Gx - d\|^2 + \tfrac{\alpha}{2}\|x\|^2$$

  for various values of alpha $10^{-6} \le \alpha \le 10^{6}$. What is the best α value to choose? Show the corresponding solution.
- Another regularization strategy used with ill-conditioned least squares problems relies on $A = U\Sigma V^t$, the singular value decomposition of the coefficient matrix to avoid computing $(A^tA)\backslash(A^tb)$. Derive the optimal solution as $x^* = \sum_{i=1}^{r} \frac{u_i^t b}{\sigma_i} v_i$, and use an appropriate r for solving the problem.



2. Systems of Nonlinear Equations. The solution of a system of n nonlinear equations in n unknowns appears in a variety of applications. Techniques for solving nonlinear equations overlap in their motivation, analysis, and implementation with unconstrained optimization techniques we've been discussing. In both optimization and nonlinear equations, Newton's method lies at the heart of many algorithms. Consider for example the following the following small system whose solution is $x^* = [0\ 0]^t$.

$$\frac{2x_1 + x_2}{(1 + (2x_1 + x_2)^2)^{1/2}} = 0 \qquad \frac{2x_1 - x_2}{(1 + (2x_1 - x_2)^2)^{1/2}} = 0$$

- Implement a basic Newton's method (with unit step length) and use it to solve this system to a tolerance $\|f\| \leq 10^{-6}$ starting at the point $x^0 = [0.3\ 0.3]^t$
- Verify the quadratic convergence of the algorithm both in the iterates and in the residuals
- When the initial guess is not near the solution, the basic Newton method will not converge. Verify that the starting point $x^0 = [0.5\ 0.5]^t$ leads to a divergent behavior.
- In order to "globalize" Newton's method, a line-search with an appropriate scalar merit function may be used at every iteration. Implement such a strategy with the merit function $m(x) = \frac{1}{2}f(x)^t f(x)$ and verify that it indeed convergences starting from $x^0 = [0.5\ 0.5]^t$.

3. Catenary Equation. In this problem we will solve a set of nonlinear algebraic equations arising from the discretization of a nonlinear differential equation. The equation describes the form of a chain or cable hanging by its own weight between two posts. The shape is not a parabola but is the solution of the following equation, where $u(x)$ is the vertical height of the chain above some reference:

$$\frac{d^2u}{dx^2} = c\sqrt{1 + \left(\frac{du}{dx}\right)^2}$$

One way to solve this equation numerically is to discretize the horizontal x-axis between the two end poles using a grid of N cells of size h, and find the values of $u(x)$ at the interior $n = (N-1)$ grid points. By using the second-order approximations of the derivatives below, the differential equation may be replaced by a set of n nonlinear algebraic equations—one equation at each of the interior grid points.

$$\frac{d^2u}{dx^2} = \frac{-1}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) \qquad \text{and} \qquad \frac{du}{dx} = \frac{1}{2h}(u_{i+1} - u_{i-1})$$

The set of equations may be written in canonical form as $f(u) = 0$, where u is the $n \times 1$ vector we seek, f is a vector function that returns the values of the n equations, and the right hand side is the zero-vector. Solve the problem with the following data: $c = 0.5$ (c depends on the weight and tension in the chain); horizontal distance between the poles is 5; height of the left pole is 2; height of the right pole is 3; $N = 100$.

- Write a function that returns the vector f
- Write a function that returns the Jacobian of f
- Pick a starting point and solve the problem using Newton's method.