SI 211: Numerical Analysis
Homework 7

Prof. Boris Houska

Deadline: Dec 2 , 2020

1. Provide short answers to the following questions

   (a) What is the main idea of Newton's method?
   Solution:
   Newton's method can be summarized as a method of approximating the zero point of a given function $f(x)$ by iteration, where each iteration can be written as $x_{k+1} = x_k - M(x_k)^{-1} \cdot f(x_k)$, with $M(x_k) = f'(x_k)$. It also includes some Newton-like algorithms derived therefrom. The main difference is that for the approximation of $M(x_k)$, for specific problems, the computational complexity of the algorithm can be greatly reduced by approximation.

   (b) What is the main idea of Gauss-Newton methods?
   Solution:
   The Gauss-Newton algorithm is a kind of Newton-like algorithm, which is only suitable for solving the least square problem

$$\min_x \frac{1}{2}\|f(x)\|_2^2.$$

   If Newton's method is used to solve the optimal problem, it often involves the inverse matrix of the Hessian matrix of the original function, which is usually accompanied by great computational complexity, while the Gauss-Newton algorithm approximates the inverse matrix of the Hessian matrix with $M(x_k) = f'(x_k)^T(x_k)$. The method avoids this problem, greatly reduces the computational complexity, and also ensures convergence.

(c) What is the local convergence rate of the exact Newton method?

Solution:

For Newton type method, we have

$$\|x_{k+1} - x^*\| \leq \kappa \|x_k - x^*\| + \frac{\omega}{2} \|x_k - x^*\|_2^2$$

where $\left\|I - M(x_k)^{-1} J(x_k)\right\| \leq \kappa$. when $M(x_k) = f'(x_k)$, we have the exact Newton method. so that $k = 0$, and then we have $\|x_{k+1} - x^*\| \leq \frac{\omega}{2} \|x_k - x^*\|_2^2$. In this case, the convergence rate is called quadratic.

(d) Under which conditions converges Newton's method in one step?

Solution:

For optimization problems, if the Hessian matrix obtained is a zero matrix, which is $f''(x_k) = 0$. Or when the gradient of the given function $J(x_k)^T$ is linear, we can find the convergence point in one step.

(e) What is the Armijo line search condition and what is it good for?

Solution:

The Armijo line search conditionthe Armijo line search condition can be represented by

$$F(x_k + \alpha_k \Delta x_k) \leq F(x_k) - c\alpha_k F'(x_k) M(x_k)^{-1} F'(x_k)^T$$

Through armijo line search, the search step size of Newton's method can be adjusted to improve the robustness of the method.

2. Consider the function $f(x) = x^3 - 2x + 2$. What happens if you apply Newton's method to the equation $f(x) = 0$ starting with $x_0 = 0$? Work out the iterates of Newton's method explicitly.

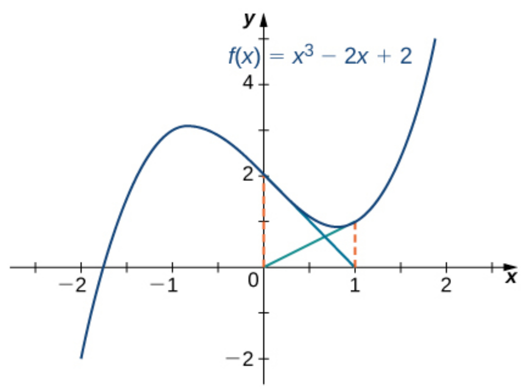

Figure 1: The approximations never approach the root of f.

Solution:
For $f(x) = x^3 - 2x + 2$, the derivative is $f'(x) = 3x^2 - 2$. Therefore,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 0 - \frac{f(0)}{f'(0)} = -\frac{2}{-2} = 1$$

In the next step,

$$x_2 = x_1 - \frac{f(x_1)}{f\prime(x_1)} = 1 - \frac{f(1)}{f'(1)} = 1 - \frac{1}{1} = 0$$

Consequently, the numbers x_0, x_1, x_2, continue to bounce back and forth between 0 and 1 and never get closer to the root of f . Fortunately, if we choose an initial approximation x_0 closer to the actual root, we can avoid this situation.

3. The goal of this exercise is to implement and exact Newton methods with Gauss-Newton methods by minimizing the nonlinear function

$$f(x,y) = \frac{1}{2}(x-1)^2 + \frac{1}{2}\left(10\left(y - x^2\right)\right)^2 + \frac{1}{2}y^2 \tag{1}$$

   (a) Derive the gradient and Hessian matrix of the function in (1). Then, re-write it in the form $f(x,y) = \frac{1}{2}\|R(x,y)\|_2^2$, where $R : \mathbb{R}^2 \to \mathbb{R}^3$ is the residual function. Derive the Gauss-Newton Hessian approximation and compare it with the exact one. When do

the two matrices coincide?
Solution:

$$\nabla f(x,y) = \begin{bmatrix} 200x^3 - 200xy + x - 1 \\ 101y - 100x^2 \end{bmatrix}$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 1 - 200y + 600x^2 & -200x \\ -200x & 101 \end{bmatrix}$$

$$R(x,y) := \begin{bmatrix} x-1 & 10\left(y - x^2\right) & y \end{bmatrix}^\top$$

$$\nabla R(x,y)^\top = J(x,y) = \begin{bmatrix} 1 & 0 \\ -20x & 10 \\ 0 & 1 \end{bmatrix}$$

$$B_{\mathrm{GN}}(x,y) := J(x,y)^\top J(x,y) = \begin{bmatrix} 1 + 400x^2 & -200x \\ -200x & 101 \end{bmatrix}$$

$$y = x^2 \Rightarrow \nabla^2 f(x,y) = B_{\mathrm{GN}}(x,y)$$

(b) Implement your own Newton's Method with exact Hessian information and full steps. Start from the initial point $(x_0, y_0) = (-1,-1)$ and use as termination condition $\|\nabla f\,(x_k, y_k)\|_2 \leq 10^{-3}$. Keep track of the iterates (x_k, y_k) and plot the results.
Solution:

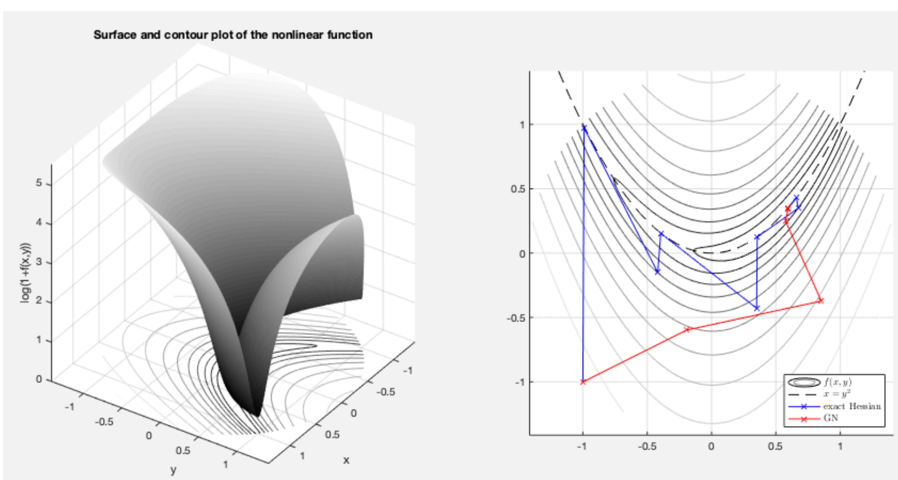

Figure 2: Newton Type algorithms.

(c) Update your code to use the Gauss-Newton Hessian approximation instead. Plot the difference between exact and approximate Hessian as a function of the iterations.

Solution:

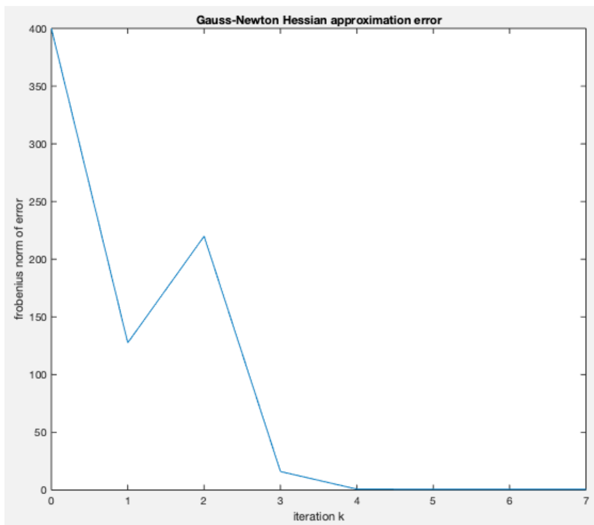

Figure 3: GN error.

(d) Compare the performance of the implemented methods. Consider the iteration path (x_k, y_k), the number of iterations and the run time. *Solution*:

Method	#iter	t in s
exact hessian	10	3.962e-03
Gauss-Newton	7	4.404e-04

Gauss Newton is much faster in computing time and also guarantees the consistency of the final result. Therefore, for the least squares problem, Using the Gauss Newton method is a method that can not only meet the accuracy requirements, but also greatly reduce the computing time.