**Unconstrained Optimization**

**Tappan Ajmera**

**Chin-yi Chu**

1. ***Results:***
2. 1. features of the function:
      1. convex
      2. local minimum is global minimum
   2. optimization solution results (*see figures 1-3*)
      1. For input x(1:100) = 1, delta = 0
      2. Gradient descent method: 0 with 23498 iterations and b = 4.9407e-324
      3. Newton method: 0 with 2 iteration & b = 0
      4. Quasi-Newton method: 0 with 100 iteration & b is 8.64748425175212e-14
   3. Since it is a convex function, Newton method works very well when Hessian matrix is positive semi-definite and non-singular. But Gradient Descent suffers due to small steps taken at every iteration. It converges slowly due to its nature. However, Quasi Newton converges at a faster rate than Gradient Descent.
3. , where m=500 & n=100
   1. features of the function: It is a complex function and hence it is not convex
   2. optimization solution results
      1. For input x = zeros(1,100), delta = 0
      2. Gradient descent method: b is -1.6967e+03 with 50 iterations
      3. Newton method: b is -1.9799e+03 with 51 iterations
      4. Quasi-Newton method: b is -1.7092e+03 with 50 iterations
   3. All three methods work almost similar for the given test function. Complex functions cannot be used for optimization, this is evident from our result as all the methods work similarly for the input function.
   4. features of the function:
      1. non-convex function
      2. local minimum is not global minimum
   5. optimization solution results
      1. For input x = [-1.2;1] , delta = 0 the global minimum is at 1,1
      2. Gradient descent method: x = 1,1 at 75032 iterations with b = 5.2931e-27
      3. Newton method: x = 1,1 with 312 iterations & b = 1.7749e-30
      4. Quasi-Newton method: x = 1, 1 with 57 iteration & b is 0.
   6. Quasi-Newton method performs the best among three methods with less iteration and more accuracy. As we can see, the Rosenbrock function is extremely slow for gradient descent and converges at 75032 iterations as the step size is very small. Newton method performs well at 313 iterations because of the analytic hessian matrix.
4. Conjugate Descent Method
5. ***Summary***

The three methods- gradient descent, Newton and Quasi-Newton, each has its pros & cons. Gradient descent method is simple and low storage. Since the method doesn’t compute second-derivative (Hessian matrix), there is no need to store matrices. It runs computationally fast per iteration. However, it could be very slow if the scale of the problem is complicated. The direction of gradient descent is not well-scaled. When the scale of the problem gets complicated, the number of iteration becomes large. It runs extremely slow.

Newton's method is simple to apply and fast convergence. It converges much faster towards a local maximum or minimum than gradient descent if the function converges. We observed this fast convergence when we ran function 1. The disadvantage of Newton method is it computes and stores Hessian matrix for each iteration. It becomes expensive if the function is highly dimensional and requiring solving linear equation. Hessian matrix would be very high dimensions and expensive to store. Newton method works well with quadratic function. Function quadratically is the unique property of Newton. The method might fail because of Hessian matrix.

Quasi-Newton method like the Newton method is simple to apply. It avoids the calculation of second derivatives and uses inexact line searches to find the direction of global convergence. The method uses approximate Hessian matrix to save the computation. We observed this nice feature when we ran function 3. Quasi-Newton method is faster than Newtown method. The optimization is more accurate than Newton.

***Figures***

Figure 1. Gradient Descent Method for Function 1

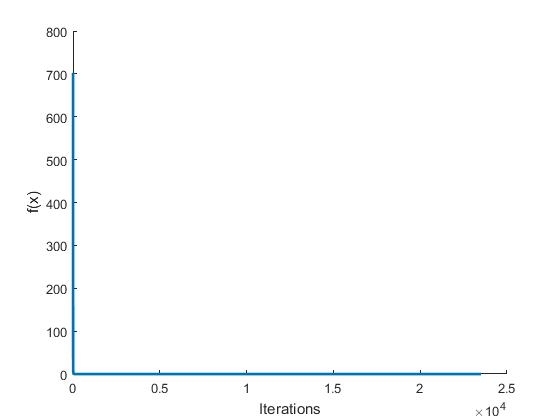


Figure 2. Newton Method for Function 1

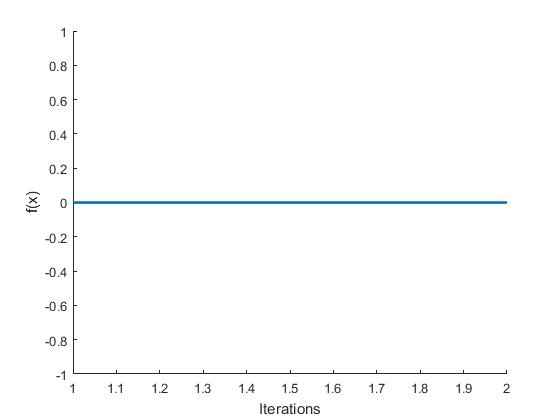


Figure 3. Quasi Newton Method for Function 1

