CS421 Assignment 6: 1997

Due Date: December 5, 1997

1. (Heath 6.9) Write a Matlab program to find a minimum of Rosenbrock's function
$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$
using each of the following methods:

   (a) steepest descent with stepsize α determined by
   $$\alpha = \textbf{fmin}(\text{fdname}, 0, \alpha_{\text{max}}, [\,], x, d)$$
   where $\alpha_{\text{max}} = \min(1, \frac{1}{\|d\|_2})$ and fdname names the function which computes $f(x + \alpha d)$, see help **fmin**. The computed stepsize α is an approximate minimizer of f along the descent direction d.

   (b) Newton's method;

   (c) Newton's method with the backtracking line search below

   Given $\gamma \in (0, \frac{1}{2})$ and $0 < \rho < 1$ (e.g., $\gamma = 10^{-4}$, $\rho = 0.9$)
   $\alpha_k = 1$
   while $f(x_k + \alpha_k d_k) > f(x_k) + \gamma \alpha_k \nabla f(x_k)^T d_k$
   $\quad \alpha_k \rightarrow \rho \alpha_k$
   end
   $x_{k+1} = x_k + \alpha_k d_k$;

   You can stop computation when
   $$\|\nabla f(x_k)\| < \text{tol} \mid \frac{\|x_k - x_{k-1}\|}{\|x_k\| + 1} < \text{tol} \mid \frac{|f_k - f_{k-1}|}{|f_k| + 1} < \text{tol} \mid k > \text{kmax}$$
   where tol $= 10^{-12}$ and kmax $= 25$. You should try each of the methods from each of the three starting points $x_0 = [-1 \;\; 1]^T$, $[0 \;\; 1]^T$ and $[2 \;\; 1]^T$. Plot the path taken in the plane by the approximate solutions for each method from each starting point.

2. In the trust region approach for $\min_{x \in \Re^n} f(x)$, the ratio

$$\rho_k = \frac{f(x_k + s_k) - f(x_k)}{\nabla f(x_k)^T s_k + \frac{1}{2} s_k^T \nabla^2 f(x_k) s_k}$$

measures the agreement between the change $f(x_k + s_k) - f(x_k)$ of the nonlinear function and the quadratic approximation $\nabla f(x_k)^T s_k + \frac{1}{2} s_k^T \nabla^2 f(x_k) s_k$. Let s_k be a solution to the trust region subproblem

$$\min_{||s|| \le \Delta_k} \nabla f(x_k)^T s + \frac{1}{2} s^T \nabla^2 f(x_k) s$$

Prove that $f(x_k + s_k) > f(x_k)$ if $\rho_k < 0$.

3. Let $A_k = J_k J_k^T$ be a symmetric positive definite approximation to the Hessian $\nabla^2 f(x_k)$ for some twice continuously differentiable $f(x)$. Assume that $y_k^T s_k > 0$. The BFGS update can be derived (as shown in class) by determining

$$A_{k+1} = J_{k+1} J_{k+1}^T$$

where

$$J_{k+1} = J_k + \frac{(y_k - J_k w) w^T}{w^T w}$$

with $w = \alpha J_k^T s_k$ and

$$\alpha^2 = \frac{y_k^T s_k}{s_k^T J_k J_k^T s_k}.$$

Prove that

$$A_{k+1} = A_k + \frac{y_k y_k^T}{s_k^T y_k} - \frac{A_k s_k s_k^T A_k}{s_k^T A_k s_k}$$

4. (Heath 6.14) The concentration of a drug in the bloodstream is expected to diminish exponentially with time. We will fit the model function

$$f(t, x) = x_1 e^{x_2 t}$$

to the following data:

t	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
y	7.80	4.00	2.50	1.75	1.48	1.25	1.20	1.15

(a) Perform the exponential fit using nonlinear least squares by implementing the Gauss-Newton method in Matlab. You can implement a similar backtrack line search algorithm and stopping criteria as in question 1.

(b) Taking logarithm of the model function gives $log(x_1) + x_2 t$, which is now linear in x_2. Thus, an exponential fit can also be done using linear least squares, assuming that we also take logarithms of the data points y_i. Using \ in Matlab, compute x_1 and x_2 in this manner. Do the values obtained agree with those determined in part (a)? Why?