Numerical Optimization HW 2

1. Write code for a general Steepest Descent Code. It should take as arguments a function f to minimize, a function df giving the gradient, and an intial point x0. It should work in $\mathbb{R}^n$ for any $n > 1$. It should use the Wolfe Conditions with the recomended parameter values from the book. The step size adjustment should be simple: multiply by 1.5 to enlarge α; multiply by 0.5 to shrink α. We want to compare the efficiency of different peoples code so your code needs to count the number and location of evaluations for f, count the number and location of of evaluations for df, and keep track of which points represent failed steps. The termination condition should be $|| \nabla f || <$ Tol where Tol = 10^{-7}. It should use the α that worked from one step on the next step and take an initial step of 1.
2. Write code for a general Newton Code. It should take as arguments a function f to minimize, a function df giving the gradient, a function ddf giving the second gradient, and an intial point x0. It should work in $\mathbb{R}^n$ for any $n > 1$. It should use the Wolfe Conditions with the recomended parameter values from the book. The step size adjustment should be simple: multiply by 1.5 to enlarge α; multiply by 0.5 to shrink α. We want to compare the efficiency of different peoples code so your code needs to count the number and location of evaluations for f, count the number and location of evaluations for df, count the number and location of evaluations for ddf, and keep track of which points represent failed steps. The termination condition should be $|| \nabla f || <$ Tol where Tol = 10^{-7}. At every step it should use an initial step of 1.
3. Test both codes on the SIAM function $f[x, y]$ from HW1 with the starting point {0.2, 0.1}. Describe how well each code worked. Specifically, compare the number and types of function evaluations, the convergence rate, and the number of steps.
4. Test both your codes on the scaled SIAM function $g[x, y] = f[100\,x, y]$ from HW1 with the starting point {0.002, 0.1}. Describe how well each code worked. Specifically, compare the number and types of function evaluations, the convergence rate, and the number of steps. Explain any differences between 3 and 4.
5. Finally test and discuss both your codes on the Rosenbrock function $h[x, y] = (1 - x)^2 + 100\left(y - x^2\right)^2$.