University of Delaware
Department of Mathematical Sciences

MATH-529 – Fundamentals of Optimization
Instructor: Dr. Marco A. Montes de Oca
Spring 2014

Homework 1

Due date: March 12, 2014

Problems

1. Suppose that $g(x)$ is a twice differentiable real-valued function that changes sign on the interval $[a, b]$, that is, $g(a)g(b) < 0$. Suppose further that there exist positive constants m and M such that

$$|g'(x)| \geq m, \qquad |g''(x)| \leq M$$

for all $x \in [a, b]$.

a) Prove that the equation $g(x) = 0$ in $[a, b]$ has one and only one solution r in $[a, b]$. (Hint: Apply the Intermediate Value Problem and Rolle's Theorem.)

b) If $x_k \in [a, b]$ and if x_{k+1} is defined in terms of x_k by Newton's Formula

$$x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)}$$

show that

$$|x_{k+1} - r| \leq \frac{M}{2m}|x_k - r|^2$$

where r is the unique solution to $g(x) = 0$ in $[a, b]$. (Hint: Apply Taylor's formula for $g(x)$ at the point $x = x_k$.)

2. In the steepest descent method, starting from $\boldsymbol{x}_k$, the next iterate is located along the line with direction vector $\boldsymbol{p}_k = -\nabla f(\boldsymbol{x}_k)$. Show that $\boldsymbol{p}_k$ is the direction of maximum decrease of the function f at $\boldsymbol{x}_k$.

3. In the multivariable Newton's method, the next estimate of the solution is calculated using the formula $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - (Hf(\boldsymbol{x}_k))^{-1}\nabla f(\boldsymbol{x}_k)$. Show that the vector $\boldsymbol{p}_k = -(Hf(\boldsymbol{x}_k))^{-1}\nabla f(\boldsymbol{x}_k)$ is a descent direction only if $Hf(\boldsymbol{x})$ is positive definite.

4. Show that if μ is a number greater than the absolute value of all eigenvalues of a symmetric matrix A, then

$$A + \mu I$$

is positive definite.

5. Use the results of the previous two exercises to modify the code of Newton's method (available in my homepage) to ensure that the step $\boldsymbol{p}_k$ is always a descent direction. (Tip: Set μ to be slighly larger than the absolute value of the most negative eigenvalue of the Hessian of the objective function.)

   Compare your implementation with the original one on the Rosenbrock function ($f(\boldsymbol{x}) = (x_1-1)^2+100(x_1^2-x_2)^2$). Use the following initial points for your comparisons: $(-1, 1)$, $(2, 2)$, and $(1, -1)$. Attach the output of each run.