Homework 3

Problem 1: Consider using a gradient algorithm to minimize the function

$$f(x) = \frac{1}{2} x^T \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} x$$

with the initial guess $x^{(0)} = [0.8, -0.25]^T$.

(a) To initialize the line search, apply the bracketing procedure along the line starting at $x^{(0)}$ in the direction of the negative gradient. Use $\epsilon = 0.075$.

(b) Apply the golden section method to reduce the width of the uncertainty region to 0.01. Organize the results of your computation in a table format or a plot.

Problem 2: Consider the minimization of

$$f(x_1, x_2) = x_1 + \frac{1}{2} x_2 + \frac{1}{2} x_1^2 + x_2^2 + 3.$$

Perform two iterations using the steepest descent method with the starting point $x^0 = (0, 0)$. Also determine an optimal solution analytically.

Problem 3: Consider the sequence $\{x^{(k)}\}$ given by $x^{(k)} = 2^{-2^{k^2}}$.

- Write down the value of the limit of $\{x^{(k)}\}$.
- Find the order of convergence of $\{x^{(k)}\}$.

Problem 4: Suppose that we use the golden section algorithm to find the minimizer of a function. Let u_k be the uncertainty range at the $k-$th iteration. Find the order of convergence of $\{u_k\}$.

Problem 5: Consider the function

$$f(x) = 3(x_1^2 + x_2^2) + 4x_1 x_2 + 5x_1 + 6x_2 + 7,$$

where $x = [x_1, x_2]^T \in \mathbb{R}^2$. Suppose that we use a fixed step size gradient algorithm to find the minimizer of f:

$$x^{(k+1)} = x^{(k)} - \alpha \nabla f(x^{(k)}).$$

Find the largest range of values of α for which the algorithm is globally convergent.

Problem 6: Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = (x - x_0)^4$, where $x_0 \in \mathbb{R}$ is a constant. Suppose that we apply Newton's method to the problem of minimizer f.

(a) Write down the update equation for Newton's method applied to the problem.

(b) Let $y^{(k)} = |x^{(k)} - x_0|$, where $x^{(k)}$ is the k-th iterate in Newton's method. Show that the sequence $\{y^{(k)}\}$ satisfies $y^{(k+1)} = \frac{2}{3}y^{(k)}$.

(c) Show that $x^{(k)} \to x_0$ for any initial guess $x^{(0)}$.

(d) Show that the order of convergence of the sequence $x^{(k)}$ in part (b) is 1.

(e) It is stated that under certain conditions, the order of convergence of Newton's method is at least 2. Why does that theorem not hold in this particular problem?

Problem 7: Consider Rosenbrock's function: $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$.

(a) Prove that $[1, 1]^T$ is the unique global minimizer of f over $\mathbb{R}^2$.

(b) With a starting point $[0, 0]^T$, apply two iterations of Newton's method.

(c) Repeat part (b) using a gradient descent algorithm using a fixed step size of $\alpha_k = 0.05$ at each step. Compare the distance to the true solution of the two algorithms.

Problem 8: Let Q be a real symmetric positive definite $n \times n$ matrix. Given an arbitrary set of linearly independent vectors $\{p^{(0)}, \cdots, p^{(n-1)}\}$ in $\mathbb{R}^n$, the Gram-Schimidt procedure generates a set of vectors $d^{(0)}, \cdots, d^{(n-1)}$ as follows :

$$d^{(0)} = p^{(0)}$$
$$d^{(k+1)} = p^{(k+1)} - \sum_{i=0}^{k} \frac{p^{(k+1)^T} Q d^{(i)}}{d^{(i)} Q d^{(i)}} d^{(i)}.$$

Show that the vectors $d^{(0)}, \cdots, d^{(n-1)}$ are $Q-$ conjugate.

Problem 9: Let Q be a real $n \times n$ symmetric matrix.

Show that there exists a Q-conjugate set $\{d^{(1)}, \cdots, d^{(n)}\}$ such that each $d^{(i)} (i = 1, \cdots, n)$ is an eigenvector of Q.

Problem 10: [Programming problem] Write a simple program for implementing the steepest descent algorithm using the secant method for the line search (using the program implemented in last homework). For the stopping criterion, use the condition $\|g^{(k)}\| \leq \epsilon$, where $\epsilon = 10^{-4}$. Test your program on the following functions and determine the number required to satisfy the stopping criterion and evaluate the objective function at the final point to see how close it is to 0.

- $f(x_1, x_2, x_3) = (x_1 - 4)^4 + (x_2 - 3)^2 + 4(x_3 + 5)^4$ using an initial point $x^{(0)} = [-4, 5, 1]^T$.
- Rosenbrock's function $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ (known to be a "nasty" function-often used as a benchmark for testing algorithms) using an initial point $x^{(0)} = [-2, 2]^T$.

Problem 11: [Programming problem] Let $f(x_1, x_2, x_3, x_4) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$ be the Powell function. Implement Newton's method to minimize it using an initial point $[3, -1, 0, 1]^T$ and stopping criteria $\frac{\|x^{(k+1)} - x^{(k)}\|}{\max\{1, \|x^{(k)}\|\}} \leq \epsilon$, where $\epsilon = 1e - 3$. Plot the figure of $\|x^{(k)} - x^*\|$ where $x^* = (0, 0, 0, 0)^T$ and $f(x^{(k)})$.

Problem 12: [Programming problem] Implement rank one, DFP, BFGS method for solving a problem in the form:

$$\min_{x \in \mathbb{R}^n} c^T x - \mu \sum_{i=1}^{m} \log(b_i - a_i^T x)$$

by generating a random set of inequalities constraint $a_i^T x \leq b_i$, and a vector c with $m = 500, n = 100$. Compare the performance of the three methods on your example.