ECE 506 Optimization Theory

Problem 1 ___ /10

Problem 2 ___ /25

Problem 3 ___ /10

Problem 4 ___ /20

Problem 5 ___ /10

Problem 6 ___ /25

Good Luck!

Problem 1 (10 points)
1(a)(**8 points**) For two-dimensional problems, we can generate all possible sign combinations of the Hessian eigenvalues by considering:

$$f_1(x_1,\, x_2) = x_1^2 + x_2^2 \tag{1}$$
$$f_2(x_1,\, x_2) = x_1^2 - x_2^2 \tag{2}$$
$$f_3(x_1,\, x_2) = -x_1^2 + x_2^2 \tag{3}$$
$$f_4(x_1,\, x_2) = -x_1^2 - x_2^2 \tag{4}$$

For each function:

- Classify the stationary point for which $\nabla f_i = 0$ as a minimum point, a maximum point, or a saddle point.
- Sketch the contours around each stationary point.

1(b)(**2 points**) Based on your answer in 1(a), can you identify which two functions exhibit the same type of stationary point?

Problem 2 (25 points)
We begin by restating theorem 3.2 from your text.

Consider any iteration of the form:

$$x_{k+1} = x_k + \alpha_k p_k, \tag{5}$$

where a_k satisfies the Wolfe conditions:

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T p_k, \tag{6}$$

$$\nabla f(x_k + \alpha_k p_k)^T p_k \geq c_2 \nabla f_k^T p_k, \qquad 0 < c_1 < c_2 < 1. \tag{7}$$

Suppose that f is bounded below in $\mathbb{R}^n$ and that f is continuously differentiable in an open set N containing:

$$L = \{x : \ f(x) \leq f(x_0)\}, \tag{8}$$

where x_0 is the starting point of the iteration. Assume also that the gradient ∇f is Lipschitz continuous on N. Then

$$\sum_{k \geq 0} \cos^2 \theta_k \|\nabla f_k\|^2 < \infty \tag{9}$$

2(a)(**5 points**) Use theorem 3.2 to show that the steepest descent algorithm will always converge.
2(b)(**8 points**) Provide a numerical example where the Wolfe conditions are satisfied. For full credit, you must provide the formula for the function and evaluate the Wolfe conditions at an admissible value of α_k, and you must also specify c_1, c_2.
2(c)(**6 points**) Suppose that you have developed a new algorithm for computing candidate directions p_k. Indicate how you would provide a new line-search algorithm that converges **at-least as many iterations** as steepest-descent or the Newton's method.

2(d)(**6 points**) Indicate how you would establish that your new algorithm proposed in 2(c) is **at-least as fast** as the steepest-descent or the Newton's method. In other words you must specifically show how you would assess your algorithm's

- accuracy,
- robustness, and
- efficiency.

Problem 3 (10 points)
Given a Hessian matrix H, outline an algorithm that can be used to:

- establish when H is positive definite, and
- modify H so that it becomes positive definite.

Problem 4 (20 points).
Suppose that f is twice differentiable and that the Hesian $\nabla^2 f(x)$ is Lipschitz continuous in a neighborhood of the solution x^*, at which the second order sufficient conditions for a strict minimum are satisfied. Show that the Newton algorithm:

$$x_{k+1} = x_k + p_k \tag{10}$$
$$p_k^N = -\nabla^2 f_k^{-1} \nabla f_k \tag{11}$$

converges quadratically, assuming that the starting point is sufficiently close to x^*.

Problem 5 (10 points)

for $k = 0, 1, \ldots,$ *MaxIterations* **do**
    Compute p_k
    Compute α_k so that the Wolfe, strong Wolfe, or Goldstein conditions are satisfied.
    $x_{k+1} = x_k + \alpha_k p_k$
end

Algorithm 1: General Framework for Line-Search Algorithms

5(a)(6 points) Give the formulas for computing p_k using

- the steepest descent method, and
- the exact Newton's method.

5(b)(4 points) Explain why we can always expect to find α_k that satisfies the Wolfe conditions.

Problem 6 (25 points)

Given $\bar{\Delta} > 0$, $\Delta_0 \in (0, \tilde{\Delta})$ and $\eta \in [0, 1/4]$.
for $k = 0, 1, 2, \ldots,$ *MaxIterations* **do**
  Obtain p_k by approximately solving:
    $\min_{p \in \mathbb{R}^n} m_k(p) = f_k + \nabla f_k^T p + \frac{1}{2} p^T B_k p$ such that: $\|p\| \le \Delta_k$
  Evaluate the reduction ratio:
    $\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)}$
  if $\rho_k < \frac{1}{4}$ **then**
    $\Delta_{k+1} = \frac{1}{4}\Delta_k$;
  else
    if $\rho_k > \frac{3}{4}$ *and* $\|p_k\| = \Delta_k$ **then**
      $\Delta_{k+1} = \min(2\Delta_k, \bar{\Delta})$
    else
      $\Delta_{k+1} = \Delta_k$
    end
  end
  if $\rho_k > \eta$ **then**
    $x_{k+1} = x_k + p_k$
  else
    $x_{k+1} = x_k$
  end
end

Algorithm 2: General Framework for Trust-Region Algorithms

6(a)(6 points) Based on the trust-region algorithm, please explain:

- How is the trust region size increased?
- How is the trust region size decreased?
- When is a new p_k direction taken?

6(b)(13 points) Suppose that an efficient algorithm is available to you that can estimate any one eigenvalue or eigenvector of a given matrix. Indicate how you to use this algorithm to improve upon the Cauchy point given as:

$$p_k^c = \tau_k p_k^s \tag{12}$$

where:

$$p_k^s = \frac{-\Delta_k}{\|\nabla f_k\|} \nabla f_k \tag{13}$$

$$\tau_k = \begin{cases} 1, & \text{if} \quad \nabla f_k^T B_k \nabla f_k \leq 0 \\ \min\left(\frac{\|\nabla f_k\|^3}{\Delta_k \nabla f_k^T B_k \nabla f_k}, 1\right), & \text{otherwise.} \end{cases} \tag{14}$$

6(c)(6 points) Explain how your algorithm will work on

$$g(x_1, x_2) = x_1^2 + a x_2^2, \tag{15}$$

for the cases when (i) $a \gg 1$, (ii) $a \ll 1$, and $a = 1$.