MA/OR/ST 706: Nonlinear Programming
Midterm Exam
Instructor: *Dr. Kartik Sivaramakrishnan*

INSTRUCTIONS

1. Please write your name and student number clearly on the front page of the exam.
2. The exam is due in my office (HA 235) by 10 am on Friday, the 29th of February. No late exams will be accepted without prior instructor approval. There is no time limit on the exam.
3. This has to be your own work. You are not allowed to consult anyone else on the exam.
4. I will be holding office hours on Thursday, the 28th of February between 11.45-1 pm to answer any questions on the exam. I will also answer any questions via email.
5. There are 5 pages and 4 questions on the exam. Each question appears on a different page. Read each question carefully.
6. The exam is worth 100 points. The distribution of points is clearly indicated on the exam.
7. Solve each problem in sufficient detail in the space provided. Please use both sides of each page as needed.
8. Write clearly, including all the steps to the final solution. If I can't read it, you won't get credit.
9. Sources: Course textbook (Nocedal & Wright), class notes, homeworks, and solutions to the homework assignments. You are NOT allowed to consult any other textbooks, journals, or online material during the exam.
10. You may use an electronic calculator and MATLAB on the exam.
11. When you turn in the exam, you are acknowledging that you have followed the above rules and regulations. Any violation of these rules is regarded as academic dishonesty, and will fetch you a zero on the exam.

1. `[30 points]` Consider the application of the steepest descent method with exact line search to minimize $f(x)$ where $x \in \mathbb{R}^n$ versus an application of the method to minimize $F(x) = \frac{1}{2}||\nabla f(x)||^2$. We will assume that $f(x) = \dfrac{1}{2}x^T Q x - b^T x$ where Q is a symmetric positive definite matrix.

   (a) `[10 points]` The steepest descent with exact line search iteration formula for $f(x)$ is given by (3.26) on page 42 in Nocedal and Wright. What is the corresponding iteration formula for $F(x)$?

   (b) `[10 points]` Let $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ be the eigenvalues of Q and x^* be an optimal solution. The error $f(x^{k+1}) - f(x^*)$ in the $(k+1)$th iteration of the steepest descent method with exact line search (see formulas (3.27) and (3.29) in Nocedal and Wright) satisfies

$$f(x^{k+1}) - f(x^*) \quad \leq \quad \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right)^2 \left(f(x^k) - f(x^*)\right). \tag{1}$$

   What is the corresponding formula for $F(x^{k+1}) - F(x^*)$? Use this result to justify why the equivalent minimization of $F(x)$ is not an attractive strategy.

   (c) `[10 points]` Consider the problem

$$\min \quad -12x_2 + 4x_1^2 + 4x_2^2 + 4x_1x_2.$$

   The optimal solution is $x^* = (-1, 2)$. Suppose we start with the point $x^0 = (0, 2)$. Using (1), find an upper bound on the number of iterations k such that $f(x^k) - f(x^*) \leq 10^{-5}$. Also, use the formula derived in part (b) to find an upper bound on the number of iterations k such that $F(x^k) - F(x^*) \leq 10^{-5}$.

2. [25 points] Consider Newton's method with constant step size $\alpha_k = 1$ applied to minimize the function $f(x) = ||x||^3$, where $x \in I\!R^n$ and $||x|| = (x_1^2 + \ldots + x_n^2)^{\frac{1}{2}}$.

   (a) [10 points] Compute the gradient $\nabla f(x)$ and the Hessian $\nabla^2 f(x)$.

   (b) [10 points] Use the Sherman-Morrison formula

   $$(I + uu^T)^{-1} \;=\; I - \tfrac{1}{2}uu^T,$$

   where I is the identity matrix of size n and $u \in I\!R^n$ is a unit vector, to compute the inverse of the Hessian $(\nabla^2 f(x))^{-1}$. Write the iteration formula for Newton's method.

   (c) [5 points] Show that Newton's method is linearly convergent to the optimal solution $x^* = 0$. Why does quadratic convergence not occur?

3. `[20 points]` Consider the problem of minimizing the function

$$f(x_1, x_2) = x_1^2 + 2x_2^2 - 2x_1x_2 - 2x_2 + 2x_1.$$

(a) `[5 points]` Find the optimal solution x^* to this problem. Is the optimal solution unique? Give reasons for your answer.

(b) `[15 points]` Starting from the point $x^0 = (2, 2)$, minimize the function using the conjugate gradient algorithm (Algorithm 5.2 on page 112). Show that the algorithm converges to the optimal solution x^* in 2 iterations.

4. [25 points] Consider the following variant of the BFGS algorithm (Algorithm 6.1 on page 140) for minimizing a function $f(x)$ with the following changes:

   (a) The initial search direction $p_0 = -\nabla f(x_0)$.

   (b) The BFGS update formula (6.17) is replaced by

   $$H_{k+1} = (I - \rho_k s_k y_k^T)(I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T.$$

   In other words, we have $H_k = I$ in the original formula (6.17).

   (c) An exact line search is carried out in every iteration.

   Let $\nabla f_k = \nabla f(x_k)$.

   (a) [5 points] Show that $\nabla f_{k+1}^T s_k = 0$. (**Hint:** The line search is exact).

   (b) [20 points] Use induction to show that the search direction in the $(k+1)$th iteration satisfies

   $$p_{k+1} = -\nabla f_{k+1} + \frac{(\nabla f_{k+1} - \nabla f_k)^T \nabla f_{k+1}}{||\nabla f_k||^2} p_k. \tag{2}$$

   This shows that the new BFGS algorithm coincides with the nonlinear conjugate gradient method (Algorithm 5.4 on page 121) with exact line search and the Polak-Ribiere formula (5.44) for β_{k+1} in (5.41a).
   (**Hint:** Show the result for $k = 1$. Assume (2) for k, and show that it continues to hold for $k + 1$).