This assignment is due at the **start** of your lecture on Friday, 28 February 2020.

1. [20 marks]

   Do Questions 3.1 and 3.9 on pages 63 and 64, respectively, of the Nocedal and Wright textbook.

   For Question 3.9, you are asked to use the Wolfe conditions. You can base your implementation of the Wolfe conditions on Algorithms 3.5 and 3.6 discussed on pages 60–62 of the Nocedal and Wright textbook.

   For each of the three line-search methods that you are asked to program, you need a *stopping criterion* for the iteration. One very simple one is to stop the iteration when

   $$\|\nabla f(x_k)\|_2 \leq \text{tol}$$

   where tol is a small value such as $\text{tol} = 10^{-3}$. This stopping criterion is a little too simple for use in a *production code*, but it should be fine for this question.

   There are a total of six problem–method combinations required for this question. For each problem–method combination, print two tables of numerical results. For the first table, print a title, an appropriate column-header line and then, on each following line, print

   $$k,\ x_k^T,\ f(x_k),\ \|\nabla f(x_k)\|_2$$

   for $k = 0, 1, 2, \ldots$. For the second table, print a title, an appropriate column-header line and then, on each following line, print

   $$k,\ \alpha_{k-1},\ \|x_k - x^*\|_2,\ \frac{\|x_k - x^*\|_2}{\|x_{k-1} - x^*\|_2},\ \frac{\|x_k - x^*\|_2}{\|x_{k-1} - x^*\|_2^2}$$

   for $k = 1, 2, \ldots$, where $x^* = (1, 1)^T$ is the minimizer of Rosenbrock's function. (I asked you to print two tables for each problem–method combination because I think it would be hard to fit all of these values into one table.)

   Try to make your table well-formatted and easy to read.

   You will likely find that some of the tables requested above are very long. For those tables that are very long, feel free to print a subset of the results. For example, you could print some results from the beginning of the iteration, some from the middle of the iteration and some from the end of the iteration. Alternatively, you could print every n^{th} line for some reasonable choice of $n > 1$. Anything sensible is fine. However, if you print a subset of the results, add a comment to your output to let your TAs know what you have chosen to print.

   Also, for each problem–method combination, print a graph of the iterates $\{x_k : k = 0, 1, 2, \ldots\}$ along with a few level sets of Rosenbrock's function. See Figure 3.7 on page 42 of the Nocedal and Wright textbook for an example of such a graph.

If your graph has too many points, feel free to print a subset of the points. However, in this case, add a comment to your output to let your TAs know what you have chosen to print.

Include sufficient documentation in your programs that your TAs will be able to understand them easily. In particular, when you are using a result from the book, reference it by equation number, theorem number, algorithm number or page number.

Write a short discussion explaining what the ratios

$$\frac{\|x_k - x^*\|_2}{\|x_{k-1} - x^*\|_2} \quad \frac{\|x_k - x^*\|_2}{\|x_{k-1} - x^*\|_2^2} \tag{1}$$

say about the rate of convergence of each problem–method combination. (In some cases, the numerical results might not be very conclusive. Try to give a fair discussion of the numerical results, but don't overstate your conclusions.)

Hand in your programs, the output requested above and the written answer to the question about the ratios (1).

You can use any programming language you like for this question, but I think MatLab (or one of the free variants of MatLab) is probably the best choice. If you use MatLab, you might find it helpful to read "help fprintf", "help open" and "help close". These functions might help you to produce well-formatted tables. In addition, if you use MatLab, you might find it helpful to read "help contour", "help plot" and "help hold". These functions might help you to produce the graphs for this question.

2. [10 marks]

   Do Question 3.7 on page 64 of the Nocedal and Wright textbook.

   Then use the Kantorovitch inequality shown in Question 3.8 on page 64 of the Nocedal and Wright textbook to show that inequality (3.29) follows from equation (3.28) on page 43 of the Nocedal and Wright textbook.

   Note that I am not asking you to prove the Kantorovitch inequality; for this question, you can just use the Kantorovitch inequality shown in Question 3.8 on page 64 of the Nocedal and Wright textbook without proof. However, for your own interest, you might try to prove it. If you can't prove it yourself, you can find several proofs on the internet.

3. [15 marks: 5 marks for each part]

   Typical convergence results for minimization problems show that either

   (a) $\nabla f(x_k) \to 0$ as $k \to \infty$, or

   (b) $f(x_k) \to f^*$ as $k \to \infty$ and $f^* \leq f(x)$ for x in some appropriate set.

They often don't claim that $x_k \to x^*$ for some $x^* \in \mathbb{R}$. In this problem, we'll see why (a) and/or (b) might hold, even though $x_k \not\to x^*$ for any $x^* \in \mathbb{R}$.

We'll also use this problem to get some experience with the Wolfe conditions on page 34 of the Nocedal and Wright textbook.

Consider the problem of minimizing $f(x) = \mathrm{e}^{-x}$ for $x \in \mathbb{R}$. This function is

- strictly convex,
- strictly decreasing,
- $f(x) > 0$ for all $x \in \mathbb{R}$,
- $f(x) \to 0$ as $x \to \infty$,
- $\nabla f(x) = f'(x) < 0$ for all $x \in \mathbb{R}$,
- $\nabla f(x) = f'(x) \to 0$ as $x \to \infty$,
- $\nabla^2 f(x) = f''(x) > 0$ for all $x \in \mathbb{R}$,
- $\nabla^2 f(x) = f''(x) \to 0$ as $x \to \infty$.

We'll see that if we apply the line-search versions of either steepest descent or Newton's method with an appropriate choice of the step-length parameter α_k to this problem, then (a) and (b) will hold and $x_k \to \infty$. So, $x_k \not\to x^*$ for any $x^* \in \mathbb{R}$.

(a) Consider the line-search steepest-descent method with $x_0 = 0$ and $\alpha_k = 1$ for all $k = 0, 1, 2, \ldots$

First show that, with this choice of α_k, $x_k \to \infty$ as $k \to \infty$.

Next show that, with this choice of α_k, the line-search steepest-descent method does not satisfy the Wolfe conditions (3.6a)–(3.6b) on page 34 of the Nocedal and Wright textbook for any choice of the constants c_1 and c_2 satisfying $0 < c_1 < c_2 < 1$.

(b) Consider the line-search steepest-descent method again with $x_0 = 0$, but this time with $\alpha_k = e^{x_k}$ for all $k = 0, 1, 2, \ldots$

First show that, with this choice of α_k, $x_k \to \infty$ as $k \to \infty$.

Next show that, with this choice of α_k, the line-search steepest-descent method does satisfy the strong Wolfe conditions (3.7a)–(3.7b) on page 34 of the Nocedal and Wright textbook for some constants c_1 and c_2 satisfying $0 < c_1 < c_2 < 1$.

Are there any additional constraints on the constants c_1 and c_2 other than $0 < c_1 < c_2 < 1$?

(c) Consider the line-search Newton's method with $x_0 = 0$ and $\alpha_k = 1$ for all $k = 0, 1, 2, \ldots$

First show that, with this choice of α_k, $x_k \to \infty$ as $k \to \infty$.

Next show that, with this choice of α_k, the line-search Newton's method does satisfy the strong Wolfe conditions (3.7a)–(3.7b) on page 34 of the Nocedal and Wright textbook for some constants c_1 and c_2 satisfying $0 < c_1 < c_2 < 1$.

Are there any additional constraints on the constants c_1 and c_2 other than $0 < c_1 < c_2 < 1$?

(To avoid repeating yourself, you can reference some of your results from (b) in your answer to (c)).