This assignment is due at the **start** of your lecture/tutorial on Monday, 3 March 2014. (We will likely have the midterm test the same day.)

1. [20 marks]

   Do questions 3.1 and 3.9 on pages 63 and 64, respectively, of your textbook.

   For both questions 3.1 and 3.9, print out the iterates, x_k, the step lengths, $\|x_{k+1} - x_k\|$, and the ratios

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

   where x^* is the solution to the problem. (Try to make your output easy to read.)

   What do the ratios (1) say about the rate of convergence of the methods?

   Also, for each method, print a graph of the iterates $\{x_k\}$ along with a few level sets of the Rosenbrock function.

   Comment your programs. In particular, when you are using a result from the book, reference it by equation number, theorem number, algorithm number or page number.

   Hand in your programs, the output requested above and the answer to the question above.

   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.

2. [10 marks]

   Do question 3.7 on page 64 of your textbook.

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

   Note that I am not asking you to prove the Kantorovitch inequality; for this question, you can just use the result shown in question 3.8 on page 64 of your 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 your 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 your 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 your 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 your 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 $0 < c_1 < c_2 < 1$?

(You can simply quote some of the results from (b) in your answer to (c)).