1. (Heath) The concentration of a drug in the bloodstream is expected to diminish exponentially with time. We will fit the model function

$$y = f(t, x_1, x_2) = x_1 e^{x_2 t},$$

to the following data

<table>
<thead>
<tr>
<th>$t$</th>
<th>0.5</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>2.5</th>
<th>3.0</th>
<th>3.5</th>
<th>4.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y$</td>
<td>6.80</td>
<td>3.00</td>
<td>1.50</td>
<td>0.75</td>
<td>0.48</td>
<td>0.25</td>
<td>0.20</td>
<td>0.15</td>
</tr>
</tbody>
</table>

(a) Perform the exponential fit using non-linear least squares. You should use the Newton’s method.

(b) Taking the logarithm of $f$ gives $\log(x_1) + x_2 t$, which is now linear in $x_2$. Thus, an exponential fit can also be done using linear least squares, assuming that we also take logarithms of the data points $y_i$. Use linear least squares to compute $x_1$ and $x_2$ in this manner. Do the values obtained agree with those determined in the previous part? Why?

2. (Heath) Let $A$ be an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. We know that $\lambda_1 = \min_x \frac{x^T Ax}{x^T x}$ and $\lambda_n = \max_x \frac{x^T Ax}{x^T x}$, with the minimum and the maximum occurring at the corresponding eigenvectors.

(a) Use an unconstrained optimization routine to compute the extreme eigenvalues and the corresponding eigenvectors of the matrix $A = \begin{pmatrix} 6 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Is the solution unique in each case? Why?

(b) The foregoing characterization of $\lambda_1$ and $\lambda_n$ remain valid if we restrict the vector $x$ to be normalized by taking $x^T x = 1$. Repeat the above part, but use a constrained optimization routine to impose this normalization constraint. What is the significance of the Lagrange multiplier in this context?

3. (Boyd, Vandenberghe) Newton’s method with fixed step size ($\eta = 1$) can diverge if the initial point is not close to $x^*$. In this problem we consider two examples (plot the function values at each of the iterates):

(a) $f(x) = \log(e^x + e^{-x})$ has a unique minimizer $x^* = 0$. Run Newton’s method with fixed step size $\eta = 1$, starting at $x(0) = 1$ and at $x(0) = 1.1$. 

1
(b) \( f(x) = -\log x + x \) has a unique minimizer at \( x^* = 1 \). Run Newton’s method with fixed step size \( \eta = 1 \), starting at \( x(0) = 3 \).

4. (Boyd, Vandenberghe) Consider the optimization problem:

\[
\min f(x) = -\sum_{i=1}^{n} x_i \log x_i, \text{ subject to } Ax = b,
\]

where \( x \in \mathbb{R}^n \) with all coordinates being positive, and \( A \) is \( p \times n \) matrix, where \( p < n \). Generate a problem instance with \( n = 100 \) and \( p = 30 \) by choosing \( A \) randomly (checking that it has full rank), choosing \( \hat{x} \) as a random positive vector (e.g., with entries uniformly distributed on \([0, 1]\)) and then setting \( b = A\hat{x} \) (Thus, \( \hat{x} \) is feasible).

5. (Boyd, Vandenberghe) Let \( \gamma > 1 \) and consider the function

\[
f(x_1, x_2) = \begin{cases} 
\sqrt{x_1^2 + \gamma x_2^2} & \text{if } |x_2| < x_1 \\
\frac{x_1 + \gamma |x_2|}{\sqrt{1 + \gamma}} & \text{otherwise}
\end{cases}
\]

Prove that \( f \) is convex. Consider the gradient descent algorithm applied to \( f \), with starting point \( x(0) = (\gamma, 1) \) and exact line search. Show that the iterates are

\[
x_1^k = \gamma \left( \frac{\gamma - 1}{\gamma + 1} \right)^k, \quad x_2^k = \left( -\frac{\gamma - 1}{\gamma + 1} \right)^k.
\]

Therefore \( x^k \) converges to \((0, 0)\). However, this is not the optimum, since \( f \) is unbounded below.

6. (Boyd, Vandenberghe) Consider the unconstrained problem

\[
\min f(x) = -\sum_{i=1}^{m} \log(1 - a_i^T x) - \sum_{i=1}^{n} \log(1 - x_i^2),
\]

where \( x \in \mathbb{R}^n \) and \( a_1, \ldots, a_m \) are vectors in \( \mathbb{R}^n \). Note that we can choose \( x(0) = 0 \) as our initial point. You can generate instances of this problem by choosing \( a_i \) from some distribution on \( \mathbb{R}^n \).

- Use the gradient descent method to solve the problem, using reasonable choices for the back-tracking parameters, and a stopping criterion of the form \( \|\nabla f(x)\| \leq \eta \). Plot the objective function and step length versus iteration number. (Once you have determined \( f(x^*) \) to high accuracy, you can also plot \( f(x) - f(x^*) \) versus iteration.) Experiment with the backtracking parameters \( \alpha \) and \( \beta \) to see their effect on the total number of iterations required. Carry these experiments out for several instances of the problem, of different sizes.
- Repeat using Newton’s method, with stopping criterion based on the Newton decrement. Look for quadratic convergence. You do not have to use an efficient method to compute the Newton step; you can use a general purpose solver for system of equations.