Linear Least Squares

1. Solving systems of linear equations, $Ax = b$, only has a solution if the rank of the matrix $A$ is less than or equal to $n$ (the number of variables, $x$), and has a unique solution only if the rank is equal to $n$. Under what conditions does the linear least squares problem, $Ax \cong b$, have a solution and under which conditions is the solution unique?
2. Consider the following linear least squares approximation problem where a function, \( f(\alpha, x) \) that is linear in the parameters \( \alpha \) is to be fitted to a set of data points such that the square residual (difference between \( f(\alpha, x) \) and \( y \) of the data) is minimized.

\[
f(\alpha, x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2, \quad \text{Data} \ (x, y): \{(0, 1), (1, 3), (2, 4), (3, 5)\}
\]

a) Provide the augmented system formulation for the problem.

b) Discuss some of the differences (in particular in terms of sensitivity and complexity) between the normal equation solution and the solution using the augmented formulation.
Nonlinear Least Squares

3. Consider the following nonlinear least squares data fitting problem.

\[ f(\alpha, x) = \alpha_1 + (x + \alpha_2)^2, \quad Data \ (x, y) : \{(0, 1), (1, 3), (2, 4)\} \]

a) Derive the Gauss-Newton formulation for this problem.

b) Show the first 2 steps of Gauss-Newton on the resulting system. (You can use normal equations to solve the linear least squares step).
Unconstrained Optimization

4. Optimization is concerned with finding an optimum of a function.
   a) What is the difference between a local and a global optimum?
   b) Name conditions under which all local minima of a function are also global minima.
One-Dimensional Optimization

5. In the Bisection method for root finding, the concept of a bracket was based on the observation that if the two endpoints of an interval have opposite signs and the function is continuous, then the function has to take on a value of 0 somewhere in the interval. Using this, computing the value of the midpoint allowed to select which half of the interval formed a new bracket. Golden Section Search uses a similar concept but uses three points to characterize the active interval. Using this and picking one new point and computing its value here again allows the algorithm to deterministically remove one of the endpoints, effectively making the interval narrower.

a) Discuss how the data point is chosen and how the endpoint that is removed is selected.

b) Describe the rationale and assumptions behind this choice and how they ensure that there will always remain a minimum in the remaining interval.
6. Consider the following nonlinear optimization problem to find the value for $x$ that (locally) minimizes the function $f(x) = x^4 + 3x^2 + 4x - 6$.

   a) Formulate Newton’s method for one-dimensional optimization for this problem (i.e. derive all the terms you need to apply Newton’s method).

   b) Show the first three iterations of Newton’s method on this problem starting from $x_0 = 0$. 

Multi-Dimensional Optimization

7. Present the basic idea behind the Nelder-Mead method for direct search-based multi-dimensional optimization. In particular, how does it address the shortcomings of Golden Section search in multi-dimensional optimization problems and how do the different operations address finding an area containing a minimum.
8. Consider the following unconstrained optimization problem to find the values for $x$ and $y$ that minimize the value for the function $f(x, y) = x^2 + y^4 + x + y$.

   a) Derive the terms needed for an optimization using Newton’s method for multi-dimensional optimization.

   b) Show the first 2 steps of Newton’s method on this problem starting with $x_0 = y_0 = 0$. 
Constrained Optimization

9. Constrained optimization problems allow for two types of constraints: equality constraints and inequality constraints.

a) Discuss some of the differences between optimization with only equality constraints and with inequality constraints. In particular focus on the applicability of different methods to the solution.

b) Why is optimization with inequality constraints more difficult than with equality constraints?
10. Derive the extended form with slack variables for the following linear programming problem.

Objective function : \( f(\vec{x}) = x_1 + 5x_2 - 7x_3 \)

Constraints :
\[
\begin{align*}
    h_1(\vec{x}) &= 2x_1 - 3x_2 - 2 \leq 0 \\
    h_2(\vec{x}) &= x_1 + 2x_3 - 4 \leq 0 \\
    h_3(\vec{x}) &= 3x_2 - 6x_3 \leq 0 \\
    x_1 \geq 0, \ x_2 \geq 0, \ x_3 \geq 0
\end{align*}
\]
11. One way to address constrained optimization with equality constraints is to convert the original objective function into a merit function by adding a weighted penalty function for the constraints. The most common one of these is the square of the constraints, $g(x)^T g(x)$.

Consider the following optimization problem with equality constraints:

Objective function : $f(x) = x_1 + x_2^2$

Constraints : $g_1(x) = x_1 - x_2 = 0$

a) Show the merit function for this problem.

b) Provide a description of the basic algorithm to solve this problem using the merit function.
c) Show the first 2 iterations of the solution using the unconstrained optimization scheme of choice in each iteration. Note: an iteration here refers to a choice of mixing parameter between objective function and penalty term, not an iteration of the unconstrained optimization algorithm you choose (e.g. Steepest descent, Newton’s method, etc). You can be very lax about the termination condition for your optimization approach in each iteration (i.e. you can let it converge after a very small number of iterations - e.g. 3). You should start at $x_1 = x_2 = 0$ in the first step of the first iteration.