Homework 2, 550.362 Optimization, Spring 2016

Important Instructions: You may discuss this homework with students currently in the class, with the TAs, and with me, up until the time that you do your write-up. At no time may you consult any existing written solutions; this would be in violation of the homework rules and, in addition, would be plagiarism if sources are not cited.

Problem 1: Suppose $A \in \mathbb{R}^{m \times n}$ is full column rank, and $b \in \mathbb{R}^m$. Derive a closed-form solution (meaning a solution expressed involving the parameters $A$ and $b$) to the optimization problem

$$\min \| Ax - b \|_2 \text{ such that } x \in \mathbb{R}^n.$$ 

Be clear why your proposed solution is indeed the solution. [ Hint: $(A^T A)^{-1}$ exists since $A$ is full column rank. Another hint: it doesn’t matter if the objective function is instead $\| Ax - b \|_2^2$. ]

Problem 2: Suppose the height $Y$ that a particular species of plant grows to can be expressed as $Y = \alpha_0 + \alpha_1 t_1 + \alpha_2 t_2 + \alpha_3 t_3 + \epsilon$ where $t_1$ is the amount of nutrient 1 in the soil, $t_2$ is the amount of nutrient 2 in the soil, $t_3$ is the amount of nutrient 3 in the soil, $\alpha_0$, $\alpha_1$, $\alpha_2$, $\alpha_3$ are fixed constants, and $\epsilon$ is a random variable with mean 0. You don’t know the value of these constants $\alpha_0$, $\alpha_1$, $\alpha_2$, $\alpha_3$ and you wish to estimate them. To do this, you do six tests; for each of $i = 1, 2, 3, 4, 5, 6$ you choose levels of nutrients 1, 2, 3 to be $\tau_{i1}$, $\tau_{i2}$, $\tau_{i3}$, respectively, and with these respective levels of nutrients the plant height was, say, $y_i$ where the matrix $\tau$ and vector $y$ are as follows:

$$\tau = \begin{bmatrix} 2 & 3 & 4 \\ 6 & 7 & 2 \\ 1 & 9 & 6 \\ 4 & 8 & 3 \\ 2 & 6 & 1 \\ 7 & 1 & 3 \end{bmatrix}, \quad y = \begin{bmatrix} 5.14 \\ 8.41 \\ 7.98 \\ 7.91 \\ 5.01 \\ 6.91 \end{bmatrix}$$

(For example, in the fifth test you used 2 units of nutrient 1, 6 units of nutrient 2, and 1 unit of nutrient 3 in the soil, and the plant grew to the height 5.01.) In order to estimate $\alpha_0$, $\alpha_1$, $\alpha_2$, $\alpha_3$, you decide to select the values of $a_0, a_1, a_2, a_3$ which minimizes $\sum_{i=1}^{6} (a_0 + a_1 \tau_{i1} + a_2 \tau_{i2} + a_3 \tau_{i3} - y_i)^2$, and these optimal values of $a_0, a_1, a_2, a_3$ will be taken as estimates for $\alpha_0, \alpha_1, \alpha_2, \alpha_3$, respectively.

a) Express this problem as precisely the kind of problem in this homework’s Problem 1.

b) Solve this problem using your solution to Problem 1, ie give your estimates of $\alpha_0, \alpha_1, \alpha_2, \alpha_3$. 

**Problem 3:** Let $S \subset \mathbb{R}^n$ be a convex set, and suppose $f : S \to \mathbb{R}$ is a function. Prove that $f$ is a convex function if and only if the epigraph of $f$ is a convex set.

**Problem 4:** Suppose you are given $A \in \mathbb{R}^{n \times n}$ symmetric positive definite, $b \in \mathbb{R}^n$, and $c \in \mathbb{R}$, and you are considering the function $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(x) = \frac{1}{2}x^T Ax + b^T x + c$. Now, suppose you are given a particular point $\bar{x} \in \mathbb{R}^n$ and a nonzero direction $d \in \mathbb{R}^n$. Derive the optimal solution to the one-dimensional subproblem

$$\min f(\bar{x} + \alpha d) \text{ such that } \alpha \in \mathbb{R}.$$  

Hint: The answer will be that the optimal value $\alpha$ is $\alpha^* = -\frac{(A\bar{x} + b)^T d}{d^T Ad}$.

**Problem 5:** Suppose you want to minimize the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by, for all $x \in \mathbb{R}^2$, $f(x) = \frac{1}{2}x^T Ax + b^T x + c$ where $A = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$, and $c = 10$.

a) Solve this optimization problem exactly using first and second order conditions.

b) Show that for any initial guess $x^{(0)}$, steepest descent method generates $x^{(1)}$, $x^{(2)}$, $x^{(3)}$, $x^{(4)}$ . . . , where, for all $k$, $x^{(k+1)} = x^{(k)} + \frac{d^T d}{d^T Ad} d$ with $d = -(Ax^{(k)} + b)$. Hint: Use Problem 4 directly.

c) On MATLAB, run twenty iterations of steepest descent method starting from point $x^{(0)} = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$.

d) Plot the sequence of steepest descent iterates from part c in $\mathbb{R}^2$ (and zoom in as necessary to follow the iterates). What do you notice?