ESE 415: Optimization (FL2024) Instructor: Shen Zeng
https://wustl.instructure.com/courses/132084 **Due:** November 15, 2024

Homework 6

1. Find the quadratic polynomial $p(t) = x_0 + x_1 t + x_2 t^2$ that best fits the following data in the least squares sense:

t	-2	-1	0	1	2
y	2	-10	0	2	1

2. Show that the least squares optimization problem of

$$\text{minimize } \|Ax - b\|^2$$

can be obtained from solving the linear algebraic system

$$\begin{pmatrix} I & A \\ A^\top & 0 \end{pmatrix} \begin{pmatrix} r \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

in an exact sense. Here I denotes the identity matrix of appropriate size.

Hint: The auxiliary variable r introduced in the algebraic approach can be viewed as the residual / error $r = b - Ax = -(Ax - b)$. Moreover, the second equation in the linear algebraic system can be viewed through the lens of the normal equations.

3. Consider a quadratic cost functional of the form

$$f(x) = \frac{1}{2} x^\top Q x - c^\top x.$$

(a) Calculate the gradient and the Hessian matrix of f.

(b) Provide the Taylor series expansion of $f(x)$ at the point $a \in \mathbb{R}^n$.

(c) Consider a single step of a gradient descent algorithm $x^+ = x - \gamma(\nabla f)(x)$. Derive closed formula for the optimal step size $\gamma^\star$ as a function of $(\nabla f)(x)$ and Q.

4. Consider a $n \times n$ symmetric and positive definite matrix $Q \succ 0$ and nonzero vectors $d_1, \ldots, d_n \in \mathbb{R}^n$ that are Q-conjugate, i.e., $d_i^\top Q d_j = 0$ for all i, j with $i \neq j$. Show that the vectors $d_1, \ldots, d_n$ are then also linearly independent in the standard sense.

5. The *Conjugate Direction Theorem* reads as follows. Let $d_0, d_1, \ldots, d_{n-1}$ be a set of nonzero Q-conjugate vectors ($Q \succ 0$). For any $x^0 \in \mathbb{R}^n$, the sequence generated via

$$x^{k+1} = x^k + \alpha_k d_k \quad \text{with} \quad \alpha_k = -\frac{g_k^\top d_k}{d_k^\top Q d_k} \quad \text{where} \quad g_k = Qx^k - b,$$

converges to the unique solution x^* of $Qx = b$ after n steps, i.e., $x^n = x^*$.

Work out a proof of the Conjugate Directions Theorem using the following strategy:

(a) Expand $x^* - x^0 = \alpha_0 d_0 + \cdots + \alpha_{n-1} d_{n-1}$, apply $d_k^\top Q$ from the left, and use the fact that $d_0, \ldots, d_{n-1}$ are Q-conjugate to obtain an explicit representation of α_k.

(b) The iteration $x^{k+1} = x^k + \alpha_k d_k$ yields $x^k = x^0 + \alpha_0 d_0 + \cdots + \alpha_{k-1} d_{k-1}$. How does x^n relate to x^*? Further consider $d_k^\top Q(x^k - x^0)$ and show that $d_k^\top Q x^0 = d_k^\top Q x^k$.

(c) Now consider $d_k^\top Q(x^* - x^0)$, which is equal to $d_k^\top Q(x^* - x_k)$ due to $d_k^\top Q x^0 = d_k^\top Q x^k$, and use $Qx^* = b$, as well as the given formula for $g_k = Qx^k - b$ to show that $d_k^\top Q(x^* - x^0) = -g_k^\top d_k$, which when plugged into the formula for α_k obtained in part (a) yields $\alpha_k = -\frac{g_k^\top d_k}{d_k^\top Q d_k}$, as claimed.

6. Consider the matrix

$$Q = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(a) Verify that the following vectors are Q-conjugate:

$$d_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad d_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad d_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

(b) Now consider with the above matrix Q and $b = (1, 1, 0)^\top$ the cost functional $f(x) = \frac{1}{2} x^\top Q x - b^\top x$. Use the above vectors d_t and the initialization $x^0 = (0, 0, 0)^\top$ to implement the method described in the Conjugate Directions Theorem.

7. (**Programming Problem**) Consider the inconsistent linear system

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \\ 2 & 0 & 5 \\ -7 & 8 & 0 \\ 1 & 2 & -1 \end{pmatrix} x = \begin{pmatrix} 3 \\ 1 \\ 2 \\ 8 \\ 0 \\ 1 \end{pmatrix}.$$

(a) Find the least squares solution.

(b) Write code implementing the *Conjugate Gradient Method* to get the same answer.