Math 408 Homework Set 2

This homework set will focus on the linear least squares problem

$$\text{LLS} \qquad \min_{x\in\mathbb{R}^n} \frac{1}{2}\|Ax-b\|_2^2 ,$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.

(1) Listed below are two functions. In each case write the problem $\min_x f(x)$ as a linear least squares problem by specifying the matrix A and the vector b, and then solve the associated problem.

SOLUTION METHOD:
Observe that

$$\|Ax-b\|_2^2 := (A_{1\cdot}\bullet x - b_1)^2 + (A_{2\cdot}\bullet x - b_2)^2 + \dots (A_{m\cdot}\bullet x - b_1 m)^2 ,$$

where $A_{i\cdot}$ is the ith row of A. We apply this to the first function below. A similar procedure applies to the second.

(a) $f(x) = (2x_1 - x_2 + 1)^2 + (x_2 - x_3)^2 + (x_3 - 1)^2$

Solution: *Note that $x \in \mathbb{R}^3$ and three terms in the sum defining f so $A \in \mathbb{R}^{3\times 3}$ and $b \in \mathbb{R}^3$. Using the description of $\|Ax-b\|_2^2$ given above, we have*

$$\begin{array}{ccc} (A_{1\cdot}\bullet x - b_1)^2 + & (A_{2\cdot}\bullet x - b_2)^2 + & (A_{3\cdot}\bullet x - b_3)^2 \\ (2x_1 - x_2 + 1)^2 + & (x_2 - x_3)^2 + & (x_3 - 1)^2 \end{array}$$

Therefore,

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \quad \textit{and} \quad b = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} .$$

(b) $f(x) = (1 - x_1)^2 + \sum_{j=1}^{4}(x_j - x_{j+1})^2$

(2) Consider the data points $(\lambda_i, y_i) \in \mathbb{R}$, $(1,1)$, $(2,0)$, $(-1,2)$, and $(0,-1)$. We wish to determine a real polynomial of degree 2 that best fits this data. A general real polynomial of degree 2 has the form $p(\lambda) = x_0 + x_1\lambda + x_2\lambda^2$, where $x = (x_0, x_1, x_2)^T \in \mathbb{R}^3$. Note that there are more data points that there are unknown coefficients x_0, x_1, and x_2 and so it is unlikely that there exists a second degree polynomial that fits this data precisely.

SOLUTION METHOD:
The solution technique for this problem is described on the first page of Chapter 2 in the notes (i.e., page 7).

(a) Write the problem of determining the quadratic polynomial that "best" fits this data as a linear least squares problem by specifying the matrix A and the vector b.

Solution: *Here we are trying to satisfy $y_i = x_0 + x_1\lambda_i + x_2\lambda_i^2$ for the given data (λ_i, y_i). The associated linear least squares problem is to minimize the sum of the squares of the misfits:*

$$(x_0 + x_1 + x_2 - 1)^2 + (x_0 + 2x_1 + 4x_2)^2 + (x_0 - x_1 + x_2 - 2)^2 + (x_0 + 1)^2,$$

in which case

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \textit{and} \quad b = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix} .$$

(b) Solve this linear least squares problem.

Solution: *Solve the linear least squares problem by solving the associated normal equations, $A^TAx = A^Tb$ (see Theorem 2.1 on page 26 of the notes):*

$$A^TA = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \quad \textit{and} \quad A^Tb = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}.$$

Hence, $x = (1/5,\ -9/10,\ 1/2)^T$ which tells us that the polynomial of degree two that best fits this data in the least squares sense is $\frac{\lambda^2}{2} - \frac{9\lambda}{10} + \frac{1}{5}$.

(3) Find the quadratic polynomial $p(t) = x_0 + x_1t + x_2t^2$ that best fits the following data in the least-squares sense:

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

Solution: *Apply the same procedure as described in problem (2) above.*

(4) Consider the problem LLS with

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

(a) What are the normal equations for this A and b.

Solution: *The normal equations are $A^TAx = A^Tb$ (see Theorem 2.1 on page 13 of the notes), where*

$$A^TA = \begin{bmatrix} 4 & 0 & 4 \\ 0 & 4 & 4 \\ 4 & 4 & 8 \end{bmatrix} \quad \text{and} \quad A^Tb = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}.$$

(b) Solve the normal equations to obtain a solution to the problem LLS for this A and b.

Solution: *The set of all solutions to the normal equations are*

$$x = \frac{1}{4}\begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \quad t \in \mathbb{R}.$$

(c) What is the general reduced QR factorization for this matrix A?

Solution:

$$QR = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 2 \end{bmatrix}.$$

(d) Compute the orthogonal projection onto the range of A.

Solution: *Apply Lemma 3.1 on page 15 of the notes to answer this question. This lemma tells us that we need to obtain an orthonormal basis for the range of A, write these basis vectors as the columns of a matrix Q, and then QQ^T is the orthogonal projection onto* $\mathrm{Ran}(A)$.

Begin by applying the Gram-Schmidt orthogonalization process to the columns of A since the range of A is the linear span of its columns. This yields the two vectors

$$q^1 = \frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} \quad \text{and} \quad q^2 = \frac{1}{2}\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}, \quad \textit{so that} \ \ Q = \begin{bmatrix}1/2 & -1/2\\1/2 & 1/2\\1/2 & -1/2\\1/2 & 1/2\end{bmatrix}.$$

Notice that this matrix Q is the same matrix that appears in the QR factorization of A. The orthogonal projection can now be written as

$$QQ^T = \begin{bmatrix}1/2 & 0 & 1/2 & 0\\0 & 1/2 & 0 & 1/2\\1/2 & 0 & 1/2 & 0\\0 & 1/2 & 0 & 1/2\end{bmatrix}.$$

(e) Use the recipe

$$\begin{aligned} AP &= Q[R_1\ R_2] && \text{the general reduced QR factorization}\\ \hat{b} &= Q^T b && \text{a matrix-vector product}\\ \bar{w}_1 &= R_1^{-1}\hat{b} && \text{a back solve}\\ \bar{x} &= P\begin{bmatrix}R_1^{-1}\hat{b}\\0\end{bmatrix} && \text{a matrix-vector product.}\end{aligned}$$

to solve LLS for this A and b.

Solution:

$$\begin{aligned}\hat{b} &= Q^T b = \begin{pmatrix}3/2\\-1/2\end{pmatrix}\\ \bar{w}_1 &= R_1^{-1}\hat{b} = \begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}^{-1}\begin{pmatrix}3\\-1\end{pmatrix} = \begin{pmatrix}3/4\\-1/4\end{pmatrix}\\ \bar{x} &= P\begin{bmatrix}R_1^{-1}\hat{b}\\0\end{bmatrix} = \begin{pmatrix}3/4\\-1/4\\0\end{pmatrix}.\end{aligned}$$

Note that this gives only the particular solution given in Part (b) above and not the entire solution set. How might you recover the entire solution set from th QR factorization?

(f) If $\bar{x}$ solves LLS for this A and b, what is $A\bar{x} - b$?

Solution:

$$A\bar{x} - b = \frac{1}{2}\begin{pmatrix}0\\-1\\0\\1\end{pmatrix}.$$

(5) Consider the matrix

$$A = \begin{bmatrix}1 & 1 & 1\\1 & 1 & 0\\1 & 0 & 0\\1 & 0 & 1\end{bmatrix}.$$

SOLUTION METHOD*:*
In this problem we again apply Lemma 3.1 on page 15 of the notes, but we could also apply Corollary 5.1.2 on page 19 of the notes.

(a) Compute the orthogonal projection onto $\mathrm{Ran}(A)$.

Solution: *After applying Gram-Schmidt to the columns of A and them writing them as the columns of the matrix Q, we obtain*

$$Q = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \end{bmatrix}.$$

This gives the orthogonal projection

$$QQ^T = \frac{1}{4}\begin{bmatrix} 3 & 1 & -1 & 1 \\ 1 & 3 & 1 & -1 \\ -1 & 1 & 3 & 1 \\ 1 & -1 & 1 & 3 \end{bmatrix}.$$

(b) Compute the orthogonal projection onto $\mathrm{Null}(A^T)$.

Solution: *Since* $\mathrm{Null}(A^T) = \mathrm{Ran}(A)^\perp$*, the projection onto* $\mathrm{Null}(A^T)$ *is just $I - QQ^T$, where Q is given above (see the discussion on page 28 of the notes):*

$$I - QQ^T = \frac{1}{4}\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}.$$

(6) Let $A \in \mathbb{R}^{m\times n}$. Show that $\mathrm{Null}(A) = \mathrm{Null}(A^TA)$.

Solution: *See Lemma 2.1 on page 13 of the notes.*

(7) Let $A \in \mathbb{R}^{m\times n}$ be such that $\mathrm{Null}(A) = \{0\}$.

(a) Show that A^TA is invertible.

Solution: *In problem (6) above we see that* $\mathrm{Null}(A) = \mathrm{Null}(A^TA)$*, and so* $\mathrm{Null}(A^TA) = \{0\}$*. Since A^TA is square, this implies that A^TA is invertible.*

(b) Show that the orthogonal projection onto $\mathrm{Ran}(A)$ is the matrix $P := A(A^TA)^{-1}A^T$.

Solution: *By Lemma 3.1 on page 15 of the notes, we need to show that $P^2 = P$, $P^T = P$ and* $\mathrm{Ran}(P) = \mathrm{Ran}(A)$*. Direct computation shows that $P^2 = P$ and $P^T = P$. To see that* $\mathrm{Ran}(P) = \mathrm{Ran}(A)$ *first note that for any $y \in \mathbb{R}^m$ set $z = (A^TA)^{-1}A^Ty$, then $Py = A\left[(A^TA)^{-1}A^Ty\right] = Az \in \mathrm{Ran}(A)$. Consequently,* $\mathrm{Ran}(P) \subset \mathrm{Ran}(A)$*. On the other hand, if $w \in \mathrm{Ran}(A)$, there is an $x \in \mathbb{R}^n$ such that $w = Ax$, hence, $Pw = PAx = A(A^TA)^{-1}A^TAx = Ax = w$ so that* $\mathrm{Ran}(A) \subset \mathrm{Ran}(P)$.