DISCUSSION 6 SOLUTIONS

MATH 170A

Problem 0.1. Set up and solve the following least squares problem. (You can just use the normal equations, which are easier to do by hand than the QR decomposition method.) The data is \((x, y) = (3, 2), (2, 5), (5, 8)\) and your model is \(y = a_1 + a_2(x - 3)^2\). (I did not choose the numbers to come out nicely.)

Solution. For instance, this means I want to have \(2 \approx a_1 + a_2(3 - 3)^2\). My \(Ax = b\) is thus

\[
\begin{bmatrix}
1 & 1 \\
1 & 0 \\
1 & 4
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2
\end{bmatrix}
= 
\begin{bmatrix}
5 \\
2 \\
8
\end{bmatrix}.
\]

(You may have your rows in different orders, but the solutions will be the same.) Multiplying both sides by \(A^T\), we get

\[
\begin{bmatrix}
3 & 5 \\
5 & 17
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2
\end{bmatrix}
= 
\begin{bmatrix}
15 \\
37
\end{bmatrix}.
\]

Solving this via row reduction (or any other method) gives

\[
\begin{bmatrix}
a_1 \\
a_2
\end{bmatrix}
= 
\begin{bmatrix}
2.6923 \\
1.3846
\end{bmatrix}.
\]

My best fit model is thus \(y = 2.6923 + 1.3846(x - 3)^2\).

Problem 0.2. Explain why the solution of the normal equation \(A^T Ax = A^T b\) is the solution to the least squares problem \(Ax = b\). You may assume that theorem 3.5.15 (and its proof) are already understood. That theorem says that \(\|b - y\|_2\) is minimized (over any choice of \(y \in S\)) when \(y\) is the projection of \(b\) onto \(S\). In other words,

\[\|b - y\|_2 = \min_{s \in S} \|b - s\|_2.\]

You can also assume that \(R(A)^\perp = N(A^T)\).

Solution. When solving the least squares problem \(Ax = b\), we are trying to find a vector \(x\) which minimizes the quantity \(\|b - Ax\|_2\). Since \(Ax\) must be in \(R(A)\), and, in fact, can be anything in \(R(A)\), Theorem 3.5.15 implies that \(x\) minimizes that norm exactly when \(Ax\) is the projection of \(b\) into \(R(A)\). But if \(Ax\) is the projection of \(b\) into \(R(A)\), then \(b - Ax\) must be the remaining part of \(b\), i.e., the projection of \(b\) into \(R(A)^\perp\).

Since \(R(A)^\perp\) is equal to \(N(A^T)\), then \(b - Ax\) must be in \(N(A^T)\). Thus, if \(x\) is a solution, \(A^T(b - Ax) = 0\). If we rearrange this, we get the normal equations,
$A^T Ax = A^T b$. Thus the solution to the normal equation is the solution to the least squares problem. \qed