Notes from Section 10

Least Squares Solutions

Situation: We want to solve $Ax = b$ but there is no solution (since $b$ is not in $\text{Col}(A)$). What do we do? Solution: Instead of giving up, we try to find $x$ that makes $Ax$ as close as possible to $b$.

In previous sections, we learned (via the Best Approximation Theorem) that if $Ax$ is as close as possible to $b$ then $Ax = \hat{b}$ where $\hat{b}$ is the orthogonal projection of $b$ onto $\text{Col}(A)$.

Let $\hat{x}$ be a solution to “$Ax$ is as close as possible to $b$.” Then $A\hat{x} = \hat{b}$.

To find $\hat{x}$ we could first find $\hat{b}$, by finding an orthonormal basis for the column space of $A$ and projecting $b$ onto $\text{Col}(A)$. Finding orthonormal bases is no fun, so the book gives us an easier way to find $\hat{x}$.

**Theorem 0.1** $\hat{x}$ satisfies $A\hat{x} = \hat{b}$ if and only if $\hat{x}$ is a solution to $A^T Ax = A^T b$.

Proof: Assume $A\hat{x} = \hat{b}$. Recall we can write $b$ as $b = \hat{b} + y$ where $y$ is the part of $b$ orthogonal to $\text{Col}(A)$. Then $y = b - \hat{b}$. Since $A\hat{x} = \hat{b}$, $y = b - A\hat{x}$. Now we multiply both sides by $A^T$. Let $A = [a_1 a_2 \ldots a_n]$

$$A^T y = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} y = \begin{pmatrix} a_1 \cdot y \\ a_2 \cdot y \\ \vdots \\ a_n \cdot y \end{pmatrix}$$

Since $a_i$ are the column vectors of $A$, multiplying these two matrices row by column we get the rows of $A^T$, which are the columns of $A$ multiplied by $y$, giving us a dot product for each entry. Since $y$ is orthogonal to $\text{Col}(A)$, all the dot products are 0. Thus $A^T y = 0$. So

$$0 = A^T (y) = A^T (b - A\hat{x})$$

$$0 = A^T (b - A\hat{x}) = A^T b - A^T (A\hat{x}) = A^T b - A^T \hat{b}$$

Thus if $A\hat{x} = \hat{b}$, $\hat{x}$ satisfies $A^T Ax = \hat{b}$. Following the argument backwards, we get that if $\hat{x}$ satisfies $A^T Ax = \hat{b}$ then $A\hat{x} = \hat{b}$. 