Lecture 2i

Overdetermined Systems

(pages 345-346)

Another use of the Approximation Theorem is to find the "best fit" solution to an inconsistent system. Specifically, we are looking for a solution for an **overdetermined system**.

<u>Definition</u>: An **overdetermined system** of linear equations is a system that has more equations than variables.

These systems do sometimes have solutions, but that requires one of the equations to be a linear combination of the others. If the equations are independent, then overdetermined systems are always inconsistent. But real life often causes us to want a solution anyway! So, while we may not be able to find a perfect solution, we can use the approximation theorem to find a vector $\vec{x}$ that is as close as possible to a solution. As in the previous lecture, we will develop our techniques while we work through an example. To that end, consider the following system:

$$\begin{array}{rrl} x_1 & +3x_2 & = -2 \\ 3x_1 & -x_2 & = 4 \\ 2x_1 & +2x_2 & = 1 \end{array}$$

Let's see if there are any solutions to this system by row reducing its augmented matrix:

$$\left[\begin{array}{rr|r} 1 & 3 & -2 \\ 3 & -1 & 4 \\ 2 & 2 & 1 \end{array}\right] \sim \left[\begin{array}{rr|r} 1 & 3 & -2 \\ 0 & -10 & 10 \\ 0 & -4 & 5 \end{array}\right] \sim \left[\begin{array}{rr|r} 1 & 3 & -2 \\ 0 & 1 & -1 \\ 0 & -4 & 5 \end{array}\right]$$
$$\sim \left[\begin{array}{rr|r} 1 & 3 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right]$$

The last row indicates that the system is inconsistent. Since we can't find an actual solution to the system, we will now try to find an approximate solution to the system. To that end, consider that a solution to our system is a vector $\vec{x}$ such that $A\vec{x} = \vec{b}$, where $A = \begin{bmatrix} 1 & 3 \\ 3 & -1 \\ 2 & 2 \end{bmatrix}$ is the coefficient matrix for our system, and $\vec{b}$ is our solution vector $\begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix}$. We already know that no such $\vec{x}$ exists, and another way of phrasing this is to say that $\vec{b}$ is not in the columnspace of A. And this is key, because now we have our subspace that we are trying to be closest to–the columnspace of A. So, since we can't find $\vec{x}$ such that $A\vec{x} = \vec{b}$, we will instead look for the unique $\vec{s}$ in Col(A) that minimizes $||\vec{b} - \vec{s}||$, where $\vec{s} = A\vec{x}$ for some $\vec{x}$. If we simply square our distance, we find ourselves in the same situation we had in the previous lecture: looking for $\vec{x}$ that minimizes $||\vec{b} - A\vec{x}||^2$. So, using the exact same argument that we used in the last lecture, we know that

$$\vec{x} = (A^TA)^{-1}A^T\vec{b}$$

So we calculate $A^TA = \begin{bmatrix} 1 & 3 & 2 \\ 3 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 3 & -1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 4 \\ 4 & 14 \end{bmatrix}$.

Next, we fine $(A^TA)^{-1}$ using the matrix inverse algorithm:

$$\left[\begin{array}{cc|cc} 14 & 4 & 1 & 0 \\ 4 & 14 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 2/7 & 1/14 & 0 \\ 1 & 7/2 & 0 & 1/4 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 2/7 & 1/14 & 0 \\ 0 & 45/14 & -1/14 & 1/4 \end{array}\right]$$
$$\sim \left[\begin{array}{cc|cc} 1 & 2/7 & 1/14 & 0 \\ 0 & 1 & -1/45 & 7/90 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 7/90 & -1/45 \\ 0 & 1 & -1/45 & 7/90 \end{array}\right]$$

So $(A^TA)^{-1} = \frac{1}{90}\begin{bmatrix} 7 & -2 \\ -2 & 7 \end{bmatrix}$. This means that

$(A^TA)^{-1}A^T = \frac{1}{90}\begin{bmatrix} 7 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 3 & -1 & 2 \end{bmatrix} = \frac{1}{90}\begin{bmatrix} 2 & 23 & 10 \\ 19 & -13 & 10 \end{bmatrix}$.

And this means that $\vec{x} = (A^TA)^{-1}A^T\vec{b} = \frac{1}{90}\begin{bmatrix} 2 & 23 & 10 \\ 19 & -13 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix} =$ $\frac{1}{90}\begin{bmatrix} 98 \\ -80 \end{bmatrix} = \begin{bmatrix} 49/45 \\ -8/9 \end{bmatrix}$.

So $\vec{x} = \begin{bmatrix} 49/45 \\ -8/9 \end{bmatrix}$ is the vector that minimizes $||\vec{b} - A\vec{x}||$.

So, we see from our example that the question of finding an "approximate" solution to the system $A\vec{x} = \vec{b}$ is the question of finding the vector $\vec{x}$ that minimizes $||\vec{b} - A\vec{x}||^2$. And using the same argument that we did when looking to find the least squares approximation, we know that this $\vec{x}$ is $(A^TA)^{-1}A^T\vec{b}$.