22M:33
Discussion of "Normal Equations: Approximate solution of an overdetermined system of linear equations

J. Simon *

May 30, 2006

Suppose you are trying to solve a system of linear equations such as

$$\begin{aligned} 2x + 3y &= 4 \\ x - y &= 5 \\ x + 2y &= 1 \\ -3x + 4y &= 7 \end{aligned}$$

Even though there are more equations than unknowns, we might still hope there exist solutions. We can form the augmented coefficient matrix, and row-reduce it. In this case,

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & -1 & 5 \\ 1 & 2 & 1 \\ -3 & 4 & 7 \end{bmatrix} \Longrightarrow rref(A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} .$$

The third row of $rref(A)$ says "zero equals 1", which tells us that the system has no solution.

Let's look at the problem stated in terms of vectors. We seek coefficients $\{x, y\}$ such that

$$x \begin{bmatrix} 2 \\ 1 \\ 1 \\ -3 \end{bmatrix} + y \begin{bmatrix} 3 \\ -1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 1 \\ 7 \end{bmatrix} .$$

The desired output vector $[4, 5, 1, 7]^T$ does not lie in the span of the two vectors on the left side; that is the same thing as saying that there are no values for v and y to solve the system.

> IF we cannot find a linear combination of $[2, 1, 1, -3]^T$ and $[3, -1, 2, 4]^T$ that exactly equals the one on the right, we can try to *find a linear combination that is as close as possible to the one on the right.*
>
> We do this by finding the projection of $[4, 5, 1, 7]^T$ into the subspace $\mathcal{U}$ of $\mathbb{R}^4$ spanned by $[2, 1, 1, -3]^T$ and $[3, -1, 2, 4]^T$.

Here is how to find that projection (without going through the complicated process of making the basis orthonormal). We are going to pretend we have the projection (as a linear combination of the two column vectors that span $\mathcal{U}$) and find equations that the coefficients have to satisfy, then use those equations to find the coefficients.

The projection vector $\mathbf{p} = Proj_{\mathcal{U}}(\mathbf{v})$ is a some linear combination of the basis vectors for $\mathcal{U}$, that is

$$\mathbf{p} = x \begin{bmatrix} 2 \\ 1 \\ 1 \\ -3 \end{bmatrix} + y \begin{bmatrix} 3 \\ -1 \\ 2 \\ 4 \end{bmatrix} ,$$

where x and y are coefficients we want to find. Note that

$$\mathbf{p} = M \begin{bmatrix} x \\ y \end{bmatrix} . \tag{1}$$

Let $\mathbf{n} = \mathbf{v} - \mathbf{p}$. We have observed before that if we subtract from $\mathbf{v}$ its projection to some subspace, then what is left is orthogonal to that subspace. So $\mathbf{n}$ is orthogonal to both columns of the matrix

$$M = \begin{bmatrix} 2 & 3 \\ 1 & -1 \\ 1 & 2 \\ -3 & 4 \end{bmatrix} .$$

We can say the same thing in terms of the transpose of M, and then do some manipulations of the matrix equation:

$$\begin{aligned} \begin{bmatrix} 2 & 1 & 1 & -3 \\ 3 & -1 & 2 & 4 \end{bmatrix} (\mathbf{v} - \mathbf{p}) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \implies M^T\mathbf{v} &= M^T\mathbf{p} \\ \implies \text{(using equation \textbf{??})} M^T\mathbf{v} &= M^T M \begin{bmatrix} x \\ y \end{bmatrix} . \end{aligned}$$

So there we have our approach: We cannot solve the system

$$M \begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{v} ,$$

instead we solve the so-called "normal system" associated with the given system

$$M^T M \begin{bmatrix} x \\ y \end{bmatrix} = M^T\mathbf{v} .$$

[End of handout]