Math 24
Winter 2010
Friday, March 5

We saw in class that if $Ax = b$ has no solution, we can find x such that Ax is as close as possible to b by choosing x so that

$$A^*Ax = A^*b.$$

If A^*A is invertible, we can rewrite this as

$$x = (A^*A)^{-1}A^*b.$$

This connects to the least squares method of fitting a line to data in the following way: Suppose we have data points (t_1, y_1), (t_2, y_2), $\ldots$, (t_n, y_n) we want to fit a line to. If the line is given by $y = ct + d$, the y-coordinates of the points on the line with t-coordinates t_1, t_2, $\ldots$, t_n are given by $ct_1 + d$, $ct_2 + d$, $\ldots$, $ct_n + d$. We would like the n-tuple of these y-coordinates, $(ct_1 + d, ct_2 + d, \ldots, ct_n + d)$, to be as close as possible to the n-tuple of actual y-coordinates of our data points, $b = (y_1, y_2, \ldots, y_n)$.

The function $T(c, d) = (ct_1+d, ct_2+d, \ldots, ct_n+d)$ is linear, and its matrix in the standard bases for $\mathbb{R}^2$ and $\mathbb{R}^n$ is

$$A = \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{pmatrix}.$$

Hence we are looking for an approximate solution $x = \begin{pmatrix} c \\ d \end{pmatrix}$ to $Ax = b$, which we find by setting

$$\begin{pmatrix} c \\ d \end{pmatrix} = x = (A^*A)^{-1}A^*b = \left(\begin{pmatrix} t_1 & t_2 & \cdots & t_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} t_1 & t_2 & \cdots & t_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

This is why the method in the example on page 363 works.