Objectives

- ☐ I understand how the best-fit line/curve is related to the equation
$$\vec{x} = (A^TA)^{-1}(A^T\vec{b})$$

In the last class, we stated that the method of least squares in an optimization problem. In particular, we showed that the sum of squares

$$S = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

is minimized when

$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i$$

and

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_iy_i.$$

For weeks now, we have been fitting data to functions, but we never used the equations stated above. Instead, we used

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

Today, we show that these two are the *same* equation.

First, observe that $\vec{x} = (A^TA)^{-1}A^T\vec{b}$ can be rewritten as

$$A^TA\vec{x} = A^T\vec{b}$$

if we multiply both sides of the equation by (A^TA).

Example 25.1: Show how $A^TA\vec{x} = A^T\vec{b}$ is equivalent to solving

$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i$$

and

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_iy_i.$$

For ease of calculation, assume $n = 4$.

Solution 25.2: Let's write out each equation with the assumption that $n = 4$.

$$a\sum_{i=1}^{4} x_i + 4b = \sum_{i=1}^{4} y_i$$

means

$$a(x_1 + x_2 + x_3 + x_4) + 4b = y_1 + y_2 + y_3 + y_4$$

and

$$a\sum_{i=1}^{4}x_i^2 + b\sum_{i=1}^{4}x_i = \sum_{i=1}^{4}x_iy_i$$

means

$$a(x_1^2 + x_2^2 + x_3^2 + x_4^2) + b(x_1 + x_2 + x_3 + x_4) = x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4.$$

Now, let's write out our usual matrices. Remember that the function is $y = ax + b$.

$$A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{pmatrix} \qquad \vec{x} = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \vec{b} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

Now, let's calculate

$$A^TA = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 + x_4^2 & x_1 + x_2 + x_3 + x_4 \\ x_1 + x_2 + x_3 + x_4 & 4 \end{pmatrix}$$

Therefore,

$$\begin{aligned} A^TA\vec{x} &= \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 + x_4^2 & x_1 + x_2 + x_3 + x_4 \\ x_1 + x_2 + x_3 + x_4 & 4 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} a(x_1^2 + x_2^2 + x_3^2 + x_4^2) + b(x_1 + x_2 + x_3 + x_4) \\ a(x_1 + x_2 + x_3 + x_4) + 4b \end{pmatrix} \end{aligned}$$

Finally, we calculate $A^T\vec{b}$.

$$\begin{aligned} A^T\vec{b} &= \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \\ &= \begin{pmatrix} x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4 \\ y_1 + y_2 + y_3 + y_4 \end{pmatrix} \end{aligned}$$

Putting this all together, $A^TA\vec{x} = A^T\vec{b}$ is

$$\begin{pmatrix} a(x_1^2 + x_2^2 + x_3^2 + x_4^2) + b(x_1 + x_2 + x_3 + x_4) \\ a(x_1 + x_2 + x_3 + x_4) + 4b \end{pmatrix} = \begin{pmatrix} x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4 \\ y_1 + y_2 + y_3 + y_4 \end{pmatrix}$$

which is the same as

$$a(x_1 + x_2 + x_3 + x_4) + 4b = y_1 + y_2 + y_3 + y_4$$

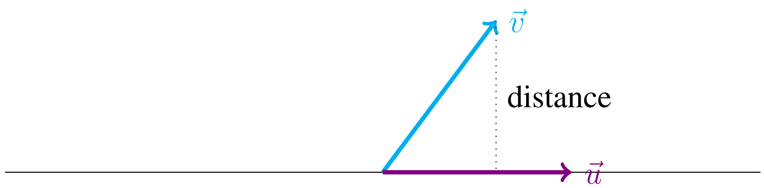
and

$$a(x_1^2 + x_2^2 + x_3^2 + x_4^2) + b(x_1 + x_2 + x_3 + x_4) = x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4.$$

If we let n be an arbitrary number (not just 4), we would see that $A^TA\vec{x} = A^T\vec{b}$ is

$$\begin{pmatrix} a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i \\ a\sum_{i=1}^{n} x_i + nb \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{pmatrix}$$

Finally, I have mentioned in class numerous times that $A^T A\vec{x} = A^T\vec{b}$ is a "projection." This seems contradictory because we have just shown that fitting data to a line (or curve) is actually a minimization problem. It turns out, all projections are minimization problems. For example, when we project one vector $\vec{v}$ onto another vector $\vec{u}$, we are *minimizing* the distance from the tip of $\vec{v}$ to the line that contains $\vec{u}$:



Summary of Ideas

- The equation
$$\vec{x} = (A^T A)^{-1}(A^T\vec{b})$$
is equivalent to the equations that minimize
$$S(a, b):$$
$$a\sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i$$
and
$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i.$$