Least-Squares
Lay 6.5

1 The problem

We are given a matrix A and a vector $\mathbf{b}$. We want to solve $A\mathbf{x} = \mathbf{b}$. However, $\mathbf{b}$ is not in the column space of A, so there is no solution. What can we do?

This kind of problem arises often–for instance, linear algebra problems come up in engineering where you want to build a system that has certain properties, but you are unable to get them exactly right. The approach we will take here is to instead ask for an approximate solution. What is the $\hat{\mathbf{x}}$ such that $A\mathbf{x}$ is as close to $\mathbf{b}$ as possible, in the sense that $\|A\hat{\mathbf{x}} - \mathbf{b}\|$ is as small as possible?

2 Solving the problem

Such an $\hat{\mathbf{x}}$ always exists; indeed, by the same line of reasoning as in our orthogonal projections lecture, what solving this problem amounts to is figuring out the value of $\hat{\mathbf{x}}$ such that $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, the orthogonal projection of $\mathbf{b}$ onto $\operatorname{col} A$.

The problem is that computing orthogonal projections without an orthogonal basis is somewhat tedious. How can we make this simpler? Note that our solution $\hat{\mathbf{x}}$ satisfies

$$\begin{aligned} A\hat{\mathbf{x}} &= \hat{\mathbf{b}} \\ \implies (\mathbf{b} - \hat{\mathbf{b}}) &= (\mathbf{b} - A\hat{\mathbf{x}}) \in (\operatorname{col} A)^{\perp} \\ \implies A^T(\mathbf{b} - A\hat{\mathbf{x}}) &= 0 \\ \implies A^T A\hat{\mathbf{x}} &= A^T\mathbf{b}. \end{aligned}$$

Thus, the vector $\mathbf{x}$ that gets us closest to solving $A\mathbf{x} = \mathbf{b}$ is the solution of the so-called "normal equations"

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

The "least-squares error" is how far we are from the real value of $\mathbf{b}$:

$$\|A\hat{\mathbf{x}} - \mathbf{b}\|$$

3 Example

I will work an easy example for the sake of time. The textbook has more complicated ones.

Example 3.1. Say that

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Clearly $A\mathbf{x} = \mathbf{b}$ has no solution. The column space of A is in fact the set of all vectors with 0 in their third coordinate (pictorially, it is the x-y plane in 3d space). So the least-squares solution had better satisfy

$$A\hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

We need the following calculations for the normal equations.

$$A^T A = \begin{bmatrix} 5 & 1 \\ 1 & 1 \end{bmatrix}; \quad A^T\mathbf{b} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Solving the normal equations gives

$$\hat{\mathbf{x}} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}.$$

This gives us the answer we expect for $A\hat{\mathbf{x}}$, and the least-squares error is 1.

4 Uniqueness

We want to know also whether the least-squares solution is unique, or whether there is more than one way to solve the normal equations. The following theorem tells us exactly when this is true:

Theorem 4.1. *If A is $m \times n$, the following are equivalent:*

1. *$A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for every $\mathbf{b}$;*
2. *The columns of A are linearly independent;*
3. *$A^T A$ is invertible.*

5 Least squares: the easy case

If the columns of A are orthogonal, then rather than writing down the normal equations, we can just explicitly read off the least-squares solution by taking orthogonal projections.

Example 5.1. Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

The columns of A are orthogonal. So writing $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix}$,

$$\hat{\mathbf{b}} = \frac{\mathbf{b} \cdot \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_1}\mathbf{a}_1 + \frac{\mathbf{b} \cdot \mathbf{a}_2}{\mathbf{a}_2 \cdot \mathbf{a}_2}\mathbf{a}_2 = 1\mathbf{a}_1 + 0\mathbf{a}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Solving $A\mathbf{x} = \hat{\mathbf{b}}$ is a matter of finding out the weights to put on the columns of A to produce $\hat{\mathbf{b}}$. But these are given as the coefficients in the above equation! so the solution is

$$\hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$