

Example 41. Find the least squares solution to $Ax = b$, where

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

Solution. First, $A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$ and $A^T b = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$.

Hence, the normal equations $A^T A \hat{x} = A^T b$ take the form $\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \hat{x} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$.

Solving, we immediately find $\hat{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Check. The error $A\hat{x} - b = \begin{bmatrix} 2 \\ 4 \\ -8 \end{bmatrix}$ is indeed orthogonal to $\text{col}(A)$. Because $\begin{bmatrix} 2 \\ 4 \\ -8 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} = 0$ and $\begin{bmatrix} 2 \\ 4 \\ -8 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = 0$.

Definition 42. The **(orthogonal) projection** \hat{b} of a vector b onto a subspace Y is the vector in Y closest to b .

The (orthogonal) projection \hat{b} of b onto $\text{col}(A)$ is $\hat{b} = A\hat{x}$.
Here, \hat{x} is a least squares solution to $Ax = b$ (i.e. \hat{x} solves $A^T A \hat{x} = A^T b$).

Why? Why is $A\hat{x}$ the projection of b onto $\text{col}(A)$?

Because, for a least squares solution \hat{x} , $A\hat{x} - b$ is as small as possible.

Note. This is a recipe for computing any orthogonal projection! That's because every subspace Y can be written as $\text{col}(A)$ for some choice of the matrix A (take, for instance, A so that its columns are a basis for Y).

Example 43. What is the orthogonal projection of $\begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$ onto $\text{span}\left\{\begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right\}$?

Solution. This is the same question as: what is the projection of b onto $\text{col}(A)$, with A and b as in the previous example.

The projection of b onto $\text{Col}(A)$ is $A\hat{x} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$.

Example 44. (homework)

(a) What is the orthogonal projection of $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ onto $\text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right\}$?

(b) What is the orthogonal projection of $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ onto $\text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}$?

Solution. (final answer only) The projections are $\left(\frac{11}{6}, \frac{1}{3}, \frac{7}{6}\right)^T$ and $\left(\frac{3}{2}, 0, \frac{3}{2}\right)^T$.

Example 45. (bonus challenge!) Can you show that, if the columns of a matrix A are independent, then $A^T A$ is always invertible?

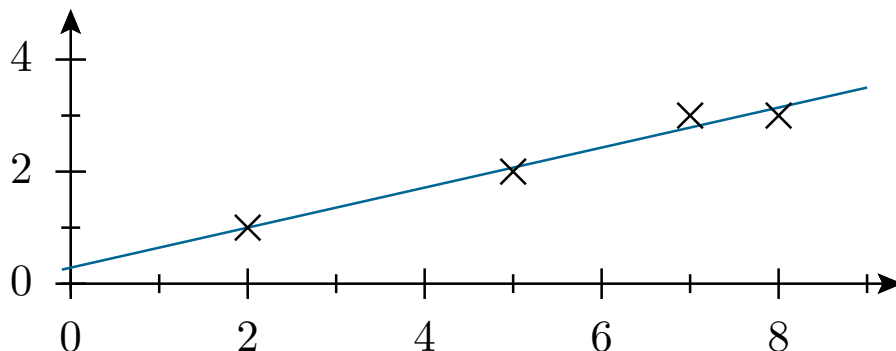
For instance. In our example, the columns of $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$ are independent. The bonus claim is that it follows that $A^T A = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$ is invertible. That's certainly true here. Why is it always true?

Hint: Assume $A^T A$ is not invertible, so that $A^T A x = 0$ for some x . Multiply both sides with x^T and ...

4.2 Application: least squares lines

Experimental data: (x_i, y_i)

Wanted: parameters β_1, β_2 such that $y_i \approx \beta_1 + \beta_2 x_i$ for all i



This approximation should be so that $SS_{\text{res}} = \underbrace{\sum_i [y_i - (\beta_1 + \beta_2 x_i)]^2}_{\text{residual sum of squares}}$ is as small as possible.

Example 46. Determine the line that best fits the data points $(2, 1), (5, 2), (7, 3), (8, 3)$.

Solution. We need to determine the values a, b for the best-fitting line $y = a + bx$.

If there was a line that fit the data perfectly, then:

$$a + 2b = 1 \quad (2, 1)$$

$$a + 5b = 2 \quad (5, 2)$$

$$a + 7b = 3 \quad (7, 3)$$

$$a + 8b = 3 \quad (8, 3)$$

In matrix form, this is: $\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} a \\ b \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{observation vector } \mathbf{y}}$ (writing the points as (x_i, y_i))

Using our points, these equations become $\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$. [This system is inconsistent (as expected).]

We compute a least squares solution.

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}, \quad X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}.$$

Solving the normal equations $\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$, we find $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$.

Hence, the least squares line is $y = \frac{2}{7} + \frac{5}{14}x$.

The plot above shows our points together with this line. It does look a very good fit!

Example 47. (homework) Find the line that best fits the points $(-1, 5), (0, 0), (1, 5), (2, 10)$.

Solution. (final answer only) The least squares line is $4 + 2x$. (Compare with earlier homework!)