

Recall that, by the FTLA, $\text{col}(A)$ and $\text{null}(A^T)$ are orthogonal complements.

The following is a useful consequence:

Theorem 34. $Ax = b$ is consistent $\iff b$ is orthogonal to $\text{null}(A^T)$

Proof. $Ax = b$ is consistent $\iff b$ is in $\text{col}(A) \xleftrightarrow{\text{FTLA}} b$ is orthogonal to $\text{null}(A^T)$

Note. b is orthogonal to $\text{null}(A^T)$ means that $y^T b = 0$ whenever $y^T A = 0$. Why?!

Example 35. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 5 \end{bmatrix}$. For which b does $Ax = b$ have a solution?

Solution. (old)

$$\left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 3 & 1 & b_2 \\ 0 & 5 & b_3 \end{array} \right] \xrightarrow{R_2 - 3R_1 \Rightarrow R_2} \left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & -5 & -3b_1 + b_2 \\ 0 & 5 & b_3 \end{array} \right] \xrightarrow{R_3 + R_2 \Rightarrow R_3} \left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 0 & -5 & -3b_1 + b_2 \\ 0 & 0 & -3b_1 + b_2 + b_3 \end{array} \right]$$

So, $Ax = b$ is consistent if and only if $-3b_1 + b_2 + b_3 = 0$.

Solution. (new) We determine a basis for $\text{null}(A^T)$:

$$\left[\begin{array}{ccc} 1 & 3 & 0 \\ 2 & 1 & 5 \end{array} \right] \xrightarrow{R_2 - 2R_1 \Rightarrow R_2} \left[\begin{array}{ccc} 1 & 3 & 0 \\ 0 & -5 & 5 \end{array} \right] \xrightarrow{-\frac{1}{5}R_2 \Rightarrow R_2} \left[\begin{array}{ccc} 1 & 3 & 0 \\ 0 & 1 & -1 \end{array} \right] \xrightarrow{R_1 - 3R_2 \Rightarrow R_1} \left[\begin{array}{ccc} 1 & 0 & 3 \\ 0 & 1 & -1 \end{array} \right]$$

We read off from the RREF that $\text{null}(A^T)$ has basis $\begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$.

b has to be orthogonal to $\text{null}(A^T)$. That means $b \cdot \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} = 0$. As above!

4 Least squares

Example 36. Not all linear systems have solutions.

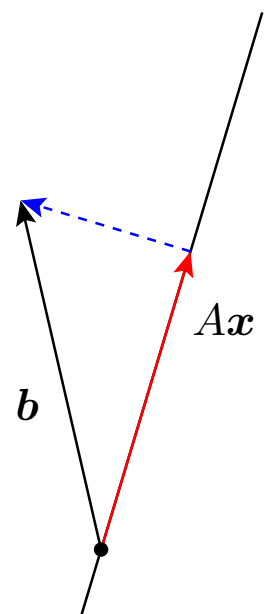
In fact, for many applications, data needs to be fitted and there is no hope for a perfect match.

For instance, $Ax = b$ with

$$\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 5 \end{bmatrix} x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

has no solution:

- $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is not in $\text{col}(A)$ since $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} \neq 0$ (see previous example).
- Instead of giving up, we want the x which makes Ax and b as close as possible.
- Such x is characterized by the error $Ax - b$ being **orthogonal** to $\text{col}(A)$ (i.e. all possible Ax).



Definition 37. $\hat{\mathbf{x}}$ is a **least squares solution** of the system $A\mathbf{x} = \mathbf{b}$ if $\hat{\mathbf{x}}$ is such that $A\hat{\mathbf{x}} - \mathbf{b}$ is as small as possible (i.e. minimal norm).

- If $A\mathbf{x} = \mathbf{b}$ is consistent, then $\hat{\mathbf{x}}$ is just an ordinary solution (in that case, $A\hat{\mathbf{x}} - \mathbf{b} = \mathbf{0}$)
- Interesting case: $A\mathbf{x} = \mathbf{b}$ is inconsistent. (in particular, if the system is overdetermined)

4.1 The normal equations

The following result provides a straightforward recipe (thanks to the FTLA) to find least squares solutions for any system $A\mathbf{x} = \mathbf{b}$.

Theorem 38. $\hat{\mathbf{x}}$ is a least squares solution of $A\mathbf{x} = \mathbf{b}$
 $\iff A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$ (the **normal equations**)

Proof.

$\hat{\mathbf{x}}$ is a least squares solution of $A\mathbf{x} = \mathbf{b}$

$\iff A\hat{\mathbf{x}} - \mathbf{b}$ is as small as possible

$\iff A\hat{\mathbf{x}} - \mathbf{b}$ is orthogonal to $\text{col}(A)$

$\stackrel{\text{FTLA}}{\iff} A\hat{\mathbf{x}} - \mathbf{b}$ is in $\text{null}(A^T)$

$\iff A^T(A\hat{\mathbf{x}} - \mathbf{b}) = \mathbf{0}$

$\iff A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$ □

Example 39. Find the least squares solution to $A\mathbf{x} = \mathbf{b}$, where

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

Solution. First, $A^T A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ and $A^T \mathbf{b} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Hence, the normal equations $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$ take the form $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Solving, we immediately find $\hat{\mathbf{x}} = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}$.

Check. Since $A\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$, the error is $A\hat{\mathbf{x}} - \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$. Recall that the error must be orthogonal to $\text{col}(A)$!

This error is indeed orthogonal to $\text{col}(A)$ because $\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = 0$ and $\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 0$.

Example 40. (homework) Find the least squares solution to $\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 5 \\ 0 \\ 5 \\ 10 \end{bmatrix}$.

Verify that the error is indeed orthogonal to $\text{col}(A)$.

Solution. (final answer only) The least squares solution is $\hat{\mathbf{x}} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.

The error is $\begin{bmatrix} 2 \\ 4 \\ 6 \\ 8 \end{bmatrix} - \begin{bmatrix} 5 \\ 0 \\ 5 \\ 10 \end{bmatrix} = \begin{bmatrix} -3 \\ 4 \\ 1 \\ -2 \end{bmatrix}$, which is indeed orthogonal to $\text{col}(A)$: $\begin{bmatrix} -3 \\ 4 \\ 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 0$, $\begin{bmatrix} -3 \\ 4 \\ 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} = 0$.