

**Example 38. (warmup)**  $\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}$

Note that this means that the system of equations  $\begin{matrix} x_1 + 2x_2 = 1 \\ 3x_1 + x_2 = 1 \\ 5x_2 = 1 \end{matrix}$  can also be written as  $\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ .

[This was the motivation for introducing matrix-vector multiplication.]

In the same way, any system can be written as  $Ax = b$ , where  $A$  is a matrix and  $b$  a vector. In particular, this makes it obvious that:

$$Ax = b \text{ is consistent} \iff b \text{ is in } \text{col}(A)$$

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

**Theorem 39.**  $Ax = b$  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 40.** 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 is,  $b \cdot \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} = 0$ . As above!

**Comment.** Below is how we can use Sage to (try and) solve  $Ax = b$  for  $b = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$  and  $b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ .

```
Sage] A = matrix([[1,2],[3,1],[0,5]])
```

```
Sage] A.solve_right(vector([1,1,2]))
```

$$\left( \frac{1}{5}, \frac{2}{5} \right)$$

```
Sage] A.solve_right(vector([1,1,1]))
```

Traceback (most recent call last):  
ValueError: matrix equation has no solutions

## Least squares

**Example 41.** Not all linear systems have solutions.

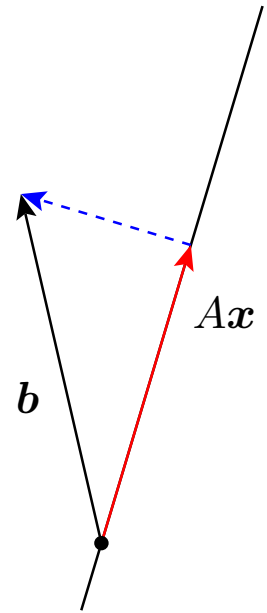
In fact, for many applications, data needs to be fitted and there is no hope for a perfect match.

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

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

## The normal equations

The following result provides a straightforward recipe (thanks to the FTLA) to find least squares solutions for all systems  $Ax = b$ .

**Theorem 43.**  $\hat{x}$  is a least squares solution of  $Ax = b$

$$\iff A^T A \hat{x} = A^T b \quad (\text{the normal equations})$$

**Proof.**

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

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

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

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

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

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

□