Sketch of Lecture 6

Mon, 1/23/2017

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

The following is a useful consequence:

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

Proof. $A\boldsymbol{x}=\boldsymbol{b}$ is consistent $\iff$ $\boldsymbol{b}$ is in $\mathrm{col}(A)$ $\overset{\text{FTLA}}{\iff}$ $\boldsymbol{b}$ is orthogonal to $\mathrm{null}(A^T)$

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

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

Solution. (old)

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

So, $A\boldsymbol{x}=\boldsymbol{b}$ is consistent if and only if $-3b_1+b_2+b_3=0$.

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

$$\begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & 5 \end{bmatrix} \overset{R_2-2R_1\Rightarrow R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 3 & 0 \\ 0 & -5 & 5 \end{bmatrix} \overset{-\frac{1}{5}R_2\Rightarrow R_2}{\rightsquigarrow} \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & -1 \end{bmatrix} \overset{R_1-3R_2\Rightarrow R_1}{\rightsquigarrow} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \end{bmatrix}$$

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

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

4 Least squares

Example 36. Not all linear systems have solutions.

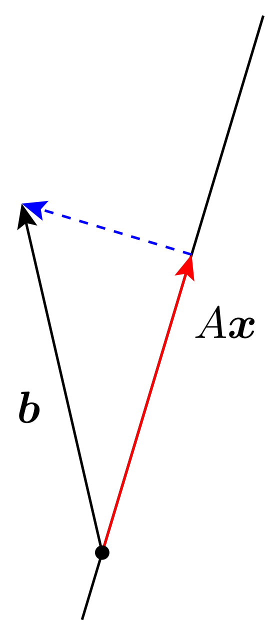
In fact, for many applications, data needs to be fitted and there is no hope for a perfect match.

For instance, $A\boldsymbol{x}=\boldsymbol{b}$ with

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

has no solution:

- $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is not in $\mathrm{col}(A)$ 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 $\boldsymbol{x}$ which makes $A\boldsymbol{x}$ and $\boldsymbol{b}$ as close as possible.
- Such $\boldsymbol{x}$ is characterized by the error $A\boldsymbol{x}-\boldsymbol{b}$ being **orthogonal** to $\mathrm{col}(A)$ (i.e. all possible $A\boldsymbol{x}$).



Armin Straub
straub@southalabama.edu

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

- If $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, then $\hat{\boldsymbol{x}}$ is just an ordinary solution (in that case, $A\hat{\boldsymbol{x}} - \boldsymbol{b} = \boldsymbol{0}$)
- Interesting case: $A\boldsymbol{x} = \boldsymbol{b}$ 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 $A\boldsymbol{x} = \boldsymbol{b}$.

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

Proof.
$\hat{\boldsymbol{x}}$ is a least squares solution of $A\boldsymbol{x} = \boldsymbol{b}$
$\iff A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is as small as possible
$\iff A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is orthogonal to $\operatorname{col}(A)$
$\overset{\text{FTLA}}{\iff} A\hat{\boldsymbol{x}} - \boldsymbol{b}$ is in $\operatorname{null}(A^T)$
$\iff A^T(A\hat{\boldsymbol{x}} - \boldsymbol{b}) = \boldsymbol{0}$
$\iff A^T A\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$ □

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

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

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 $A^T\boldsymbol{b} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Hence, the normal equations $A^T A\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$ take the form $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\hat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.
Solving, we immediately find $\hat{\boldsymbol{x}} = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}$.

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

This error is indeed orthogonal to $\operatorname{col}(A)$ 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 $\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\boldsymbol{x} = \begin{bmatrix} 5 \\ 0 \\ 5 \\ 10 \end{bmatrix}$.

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

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

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 $\operatorname{col}(A)$: $\begin{bmatrix} -3 \\ 4 \\ 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 0$, $\begin{bmatrix} -3 \\ 4 \\ 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} = 0$.