Last time

- $\text{proj}_V = QQ^T$, $Q = [\vec{u}_1 \cdots \vec{u}_r]$
- Orthogonal projection minimizes distance
- Least-squares solutions.

We know that $\vec{x}^*$ is a least-squares solution to $A\vec{x}=\vec{b}$ when $\vec{b}-A\vec{x}^*$ is orthogonal to $\text{im}(A)$.

Def The <u>orthogonal complement</u> of the subspace V of $\mathbb{R}^n$ is the collection $V^\perp$ of all vectors $\vec{x}$ in $\mathbb{R}^n$ orthogonal to every vector in V.

Facts about $V^\perp$

(1) $V^\perp$ is a subspace

(2) The only vector in V and in $V^\perp$ is $\vec{0}$

(3) $\dim V + \dim V^\perp = n$

(4) $(V^\perp)^\perp = V$

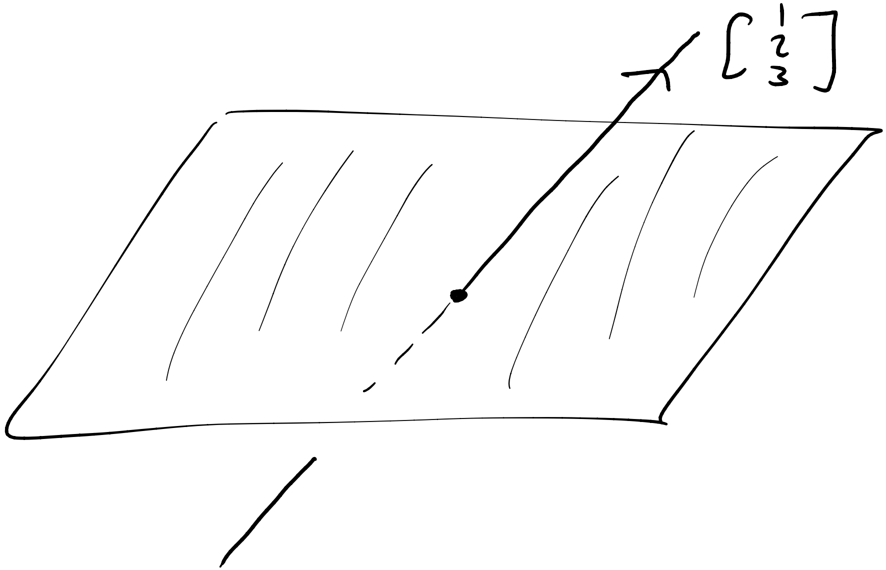
(5) Every vector $\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$ with $\vec{x}^{\parallel}$ in V and $\vec{x}^{\perp}$ in $V^\perp$, uniquely.

So $\vec{x}^*$ is a LSS to $A\vec{x}=\vec{b}$ precisely when $\vec{b}-A\vec{x}^*$ is in $\operatorname{im}(A)^\perp$. Since the columns of A span $\operatorname{im}(A)$, this condition is equivalent to being orthogonal to every column of A, which is to say dotting to 0 with every row of A^T.

$$\boxed{\operatorname{im}(A)^\perp = \ker(A^T)}$$

<u>Ex</u> If V is the line in $\mathbb{R}^3$ parallel to $\begin{bmatrix}1\\2\\3\end{bmatrix}$, then $V=\operatorname{im}(A)$, where $A=\begin{bmatrix}1\\2\\3\end{bmatrix}$,

while $\ker(A^T) = \ker([1\ 2\ 3])$ is the plane described by the equation $x + 2y + 3z = 0$



In summary, $\vec{x}^*$ is a LSS if and only if $A^T(\vec{b} - A\vec{x}^*) = \vec{0}$.

Normal equation The least-squares solutions of the system $A\vec{x} = \vec{b}$ are the solutions of the system

$$\boxed{A^T A \vec{x} = A^T \vec{b}}$$

Last time

- proj_V minimizes distance
- Least squares solutions
- Normal equation $\quad A^TA\vec{x} = A^T\vec{b}$

Ex Let's find the line that best fits the data points $(1,1)$, $(2,3)$, $(3,4)$

Writing $y = ax+b$, we have the system

$$\left.\begin{array}{l} 1 = a+b \\ 3 = 2a+b \\ 4 = 3a+b \end{array}\right\} \quad \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 2 & 0 & 3 \end{array}\right] \quad \begin{array}{l} a = 2 \\ a = 3/2 \end{array}$$

So the system is inconsistent. To find the LSS, we use the normal equations:

$$A^TA = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}$$

$$A^T\vec{b} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 19 \\ 8 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 14 & 6 & 19 \\ 6 & 3 & 8 \end{array}\right]$$

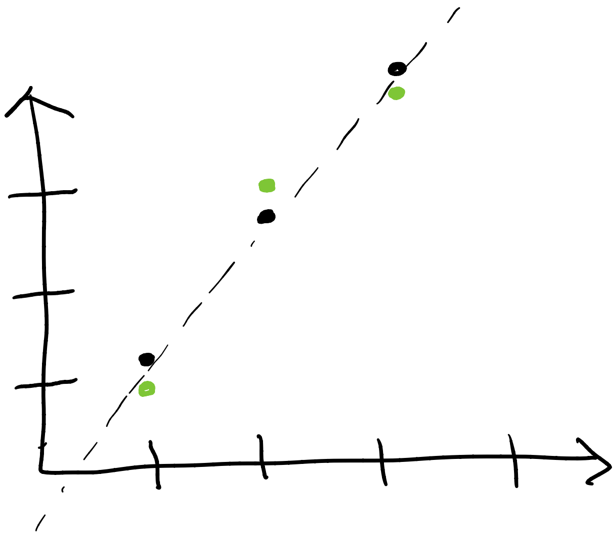
$$\downarrow$$

$$\left[\begin{array}{cc|c} 2 & 0 & 3 \\ 0 & 3 & -1 \end{array}\right]$$

$a = 3/2$

$b = -1/3$

The line of best fit is $y = \frac{3}{2}x - \frac{1}{3}$



Data	Line
$(1, 1)$	$(1, 1\frac{1}{6})$
$(2, 3)$	$(2, 2\frac{2}{3})$
$(3, 4)$	$(3, 4\frac{1}{6})$

In this example, there was a unique least squares solution because $\ker(A^TA) = \{\vec{0}\}$. Since A^TA is square, this happens precisely when A^TA is invertible.

<u>Unique least squares</u> A^TA is invertible if and only if $\ker(A) = \{\vec{0}\}$, in which case $A\vec{x} = \vec{b}$ has the unique LSS

$$\vec{x}^* = (A^TA)^{-1}A^T\vec{b}$$

<u>Ex</u> In our previous example,

$$(A^TA)^{-1} = \frac{1}{6}\begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix}$$

$$(A^TA)^{-1}A^T\vec{b} = \frac{1}{6}\begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$$

$$= \frac{1}{6}\begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix}\begin{bmatrix} 19 \\ 8 \end{bmatrix}$$

$$= \frac{1}{6}\begin{bmatrix} 9 \\ -2 \end{bmatrix}$$

$$= \begin{bmatrix} 3/2 \\ -1/3 \end{bmatrix}.$$

To justify the claim about A^TA, it suffices to show that $\ker(A) = \ker(A^TA)$, but

$$A^TA\vec{x} = \vec{0} \iff A\vec{x} \text{ is in } \ker(A^T) = \operatorname{im}(A)^\perp$$
$$\iff A\vec{x} \text{ is in } \operatorname{im}(A) \text{ and } \operatorname{im}(A)^\perp$$
$$\iff A\vec{x} = \vec{0}.$$

<u>Goal</u> Decide whether a matrix is invertible without having to row reduce.

For a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have already seen that the determinant $ad-bc$ tells the whole story.

<u>Recall</u> A is invertible if and only if its columns are linearly independent.

For two vectors, this just means $\begin{bmatrix} a \\ c \end{bmatrix}$ is a scalar multiple of $\begin{bmatrix} b \\ d \end{bmatrix}$, and this is what the determinant detects.

On the other hand, three vectors are linearly dependent precisely when all three lie in a plane, which is to say that there is some vector orthogonal to all three.

<u>Def</u> The <u>determinant</u> of the 3×3 matrix $A = [\vec{u}\ \vec{v}\ \vec{w}]$ is the number

$$\det(A) = \vec{u}\cdot(\vec{v}\times\vec{w}).$$

This number does the job since

$\det(A) = 0 \iff \vec{u}$ is orthogonal to $\vec{v} \times \vec{w}$

$\iff$ either $\vec{v}$ and $\vec{w}$ are parallel or $\vec{u}$ lives in the plane spanned by $\vec{v}$ and $\vec{w}$

$\iff \{\vec{u}, \vec{v}, \vec{w}\}$ is not linearly independent

$\iff$ A is not invertible.