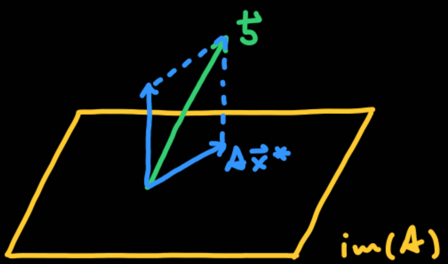
Least squares:

1. What it is and how to solve for least squares.
2. Examples.
3. Computing matrices of orthogonal projections.
4. Examples.

1. What are these and how to solve for them.



$A\vec{x} = \vec{b}$

$A\vec{x}^*$ is in $\text{im}(A)$

$\|\vec{b} - A\vec{x}^*\| \leq \|\vec{b} - A\vec{x}\|$

for all $\vec{x}$.

Def: The least squares solution to a system of equations $A\vec{x} = \vec{b}$ is the vector $\vec{x}^*$ satisfying that $A\vec{x}^*$ is the closest vector to $\vec{b}$ inside $\text{im}(A)$.

This amounts to finding $\vec{x}^*$ such that $\|\vec{b} - A\vec{x}^*\| \leq \|\vec{b} - A\vec{x}\|$ for all $\vec{x}$.

Namely, $A\vec{x}^* = \underbrace{\text{proj}_{\text{im}(A)}(\vec{b})}_{\text{vector in im}(A)\text{ closest to }\vec{b}}$.

Theorem: The least squares solution of $A\vec{x} = \vec{b}$ are the solutions of:

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

This is called the normal equation of the system.

2. Example.

① Find the least squares solution of:

$$\underbrace{\begin{bmatrix}1&4\\2&5\\3&6\end{bmatrix}}_{A}\begin{bmatrix}x\\y\end{bmatrix}=\underbrace{\begin{bmatrix}5\\7\\8\end{bmatrix}}_{\vec{b}}.$$

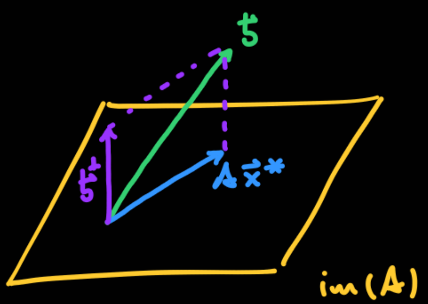
We solve: $A^TA\vec{x}=A^T\vec{b}$.

$$\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}\begin{bmatrix}1&4\\2&5\\3&6\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}\begin{bmatrix}5\\7\\8\end{bmatrix}$$

$$\begin{bmatrix}1+4+9&4+10+18\\4+10+18&16+25+36\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}5+14+24\\20+35+48\end{bmatrix}$$

$$\begin{bmatrix}14&32\\32&77\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}43\\103\end{bmatrix}$$

$$\left[\begin{array}{cc|c}14&32&43\\32&77&103\end{array}\right]\xrightarrow{\text{rref}}\left[\begin{array}{cc|c}1&0&5/18\\0&1&11/9\end{array}\right],\ \text{so}\ \underbrace{\vec{x}^*=\begin{bmatrix}5/18\\11/9\end{bmatrix}}_{\text{least squares solution.}}.$$



$$A\vec{x}^*=\begin{bmatrix}1&4\\2&5\\3&6\end{bmatrix}\begin{bmatrix}5/18\\11/9\end{bmatrix}=\begin{bmatrix}\frac{5}{18}+\frac{44}{9}\\\frac{10}{18}+\frac{55}{9}\\\frac{15}{18}+\frac{66}{9}\end{bmatrix}=\underbrace{\begin{bmatrix}\frac{31}{6}\\\frac{20}{3}\\\frac{49}{6}\end{bmatrix}}_{\text{not }\vec{b}}$$

should be $\text{proj}_{\text{im}(A)}(\vec{b})$.

$$\vec{b}^\perp=\vec{b}-A\vec{x}^*=\begin{bmatrix}5\\7\\8\end{bmatrix}-\begin{bmatrix}31/6\\20/3\\49/6\end{bmatrix}=\begin{bmatrix}(30-31)/6\\(21-20)/3\\(48-49)/6\end{bmatrix}=\begin{bmatrix}-1/6\\1/3\\-1/6\end{bmatrix}$$

Now: $\begin{bmatrix}-1/6\\1/3\\-1/6\end{bmatrix}\cdot\begin{bmatrix}1\\2\\3\end{bmatrix}=\frac{-1}{6}+\frac{2}{3}-\frac{3}{6}=0$; $\begin{bmatrix}-1/6\\1/3\\-1/6\end{bmatrix}\cdot\begin{bmatrix}4\\5\\6\end{bmatrix}=\frac{-4}{6}+\frac{5}{3}-\frac{6}{6}=0.$

② Find the least squares solution of:

$$\underbrace{\begin{bmatrix}1&4\\2&5\\3&6\end{bmatrix}}_{A}\begin{bmatrix}x\\y\end{bmatrix}=\underbrace{\begin{bmatrix}2\\1\\0\end{bmatrix}}_{\vec{b}}.$$

We have to solve $A^TA\vec{x}=A^T\vec{b}$.

$$\underbrace{\begin{bmatrix}14&32\\32&77\end{bmatrix}}_{A^TA}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}\begin{bmatrix}2\\1\\0\end{bmatrix}.$$

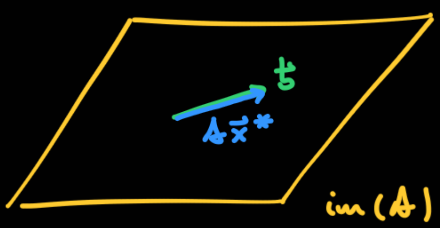
$$\begin{bmatrix}14&32\\32&77\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}4\\13\end{bmatrix}$$

$$\left[\begin{array}{cc|c}14&32&4\\32&77&13\end{array}\right]\xrightarrow{\text{rref}}\left[\begin{array}{cc|c}1&0&-2\\0&1&1\end{array}\right].\quad \text{So } \vec{x}^*=\begin{bmatrix}-2\\1\end{bmatrix}.$$

Now:

$$A\vec{x}^*=\begin{bmatrix}1&4\\2&5\\3&6\end{bmatrix}\begin{bmatrix}-2\\1\end{bmatrix}=\begin{bmatrix}-2+4\\-4+5\\-6+6\end{bmatrix}=\begin{bmatrix}2\\1\\0\end{bmatrix}=\vec{b}.$$

Note that when $\vec{b}$ is in $\operatorname{im}(A)$, computing $\vec{x}^*$ is the same thing as solving $A\vec{x}=\vec{b}$.



3. Matrices of orthogonal projections.

Theorem: Given $V=\operatorname{span}(\underbrace{\vec{v}_1,\vec{v}_2,\vec{v}_3}_{\text{basis, not necessarily orthonormal}})$, the matrix of the projection onto

V is given by:

$$A = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix}$$

$$P = A \cdot (A^T A)^{-1} \cdot A^T$$

Recall that when $V = \text{span}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ with $\vec{u}_1, \vec{u}_2, \vec{u}_3$ orthonormal basis of V then $P = A \cdot A^T$.

Question: What is $A^T A$?

4. Example.

① Find the matrix giving orthogonal projection onto $V = \text{span}\left(\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}4\\5\\6\end{bmatrix}\right)$.

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \frac{1}{54} \cdot \begin{bmatrix} 77 & -32 \\ -32 & 14 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} =$$

$$= \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

Sanity check:

$$P \cdot \begin{bmatrix} 5 \\ 7 \\ 8 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \\ 8 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 25 + 14 - 8 \\ 10 + 14 + 16 \\ -5 + 14 + 40 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 31 \\ 40 \\ 49 \end{bmatrix} = \begin{bmatrix} 31/6 \\ 20/3 \\ 49/6 \end{bmatrix}$$

$$P\cdot\begin{bmatrix}2\\1\\0\end{bmatrix}=\frac{1}{6}\begin{bmatrix}5&2&-1\\2&2&2\\-1&2&5\end{bmatrix}\begin{bmatrix}2\\1\\0\end{bmatrix}=\frac{1}{6}\begin{bmatrix}10+2\\4+2\\0\end{bmatrix}=\begin{bmatrix}2\\1\\0\end{bmatrix}$$

② Projection in $\mathbb{R}^2$ onto the line $y=x$.

$$L=\operatorname{span}\left(\begin{bmatrix}1\\1\end{bmatrix}\right) \qquad A=\begin{bmatrix}1\\1\end{bmatrix}$$

$$P=A(A^TA)^{-1}A^T=\begin{bmatrix}1\\1\end{bmatrix}\left(\begin{bmatrix}1&1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\right)^{-1}\begin{bmatrix}1&1\end{bmatrix}=$$

$$=\begin{bmatrix}1\\1\end{bmatrix}\left(\begin{bmatrix}2\end{bmatrix}\right)^{-1}\begin{bmatrix}1&1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}\frac{1}{2}\end{bmatrix}\begin{bmatrix}1&1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}\frac{1}{2}&\frac{1}{2}\end{bmatrix}=$$

$$=\begin{bmatrix}\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}.$$