MAT 022A Nov 6, 2019

Thursday, October 31, 2019 5:12 PM

__Recall__: To project vectors from a high dimensional space to a lower dimensional subspace, use projection matrix P

Projection matrix formula:
Create a matrix A whose columns give a basis for the subspace
Then

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

An equivalent way to understand a projection $P\vec{b}$ of a vector $\vec{b}$ in the big space to the smaller space:
We know Pb is in the column space of A: $C(A)$
So

$$A\vec{y} = P\vec{b} \text{ has a solution } \vec{y}.$$

$$A\vec{y} = \underbrace{A(A^TA)^{-1}A^T}_{P}\vec{b}$$

So try to solve $\vec{y} = (A^TA)^{-1}A^T\vec{b}$

equivalently $\boxed{A^TA\vec{y} = A^T\vec{b}}$

General idea: Sometimes $A\vec{y} = \vec{b}$ has no solution (if $\vec{b}$ is not in column space of A)

Next best thing: Solve $\boxed{A^T A\vec{y} = A^T b}$

This is always possible by solving $A\vec{y} = P\vec{b}$ (which is possible b/c Pb is in $C(A)$)

Easy example:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Trying to solve $Ax=b$:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \rightarrow \begin{cases} x=0 \\ y=0 \\ 0=1 \end{cases}$$

no solution!

Next best thing: solve $A^T A x = A^T b$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

This solution is ok but has an error term

$$A\vec{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

The difference between $A\vec{x}$ and $\vec{b}$ is the error

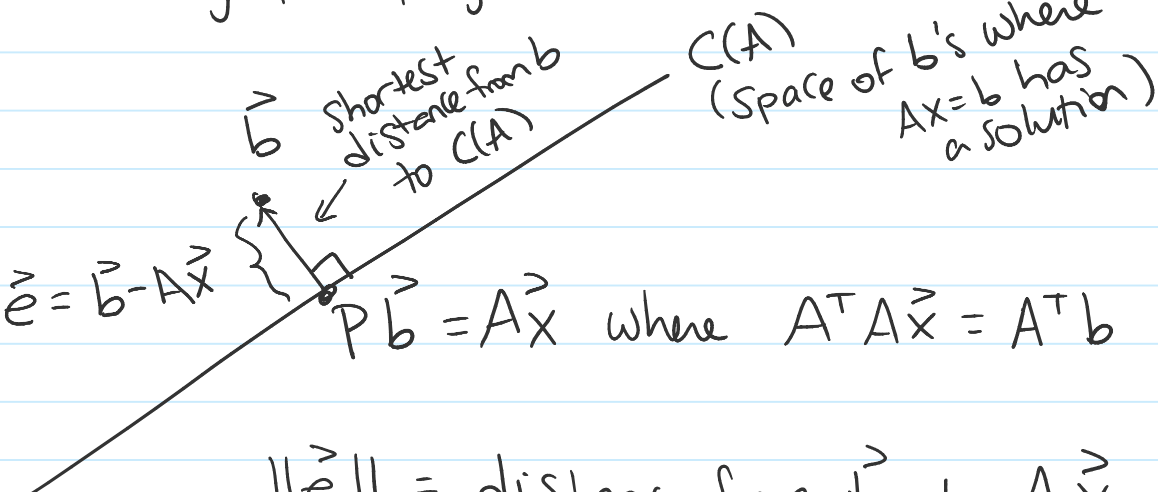
$$\underline{\underline{\vec{e} = \vec{b} - A\vec{x}}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

To measure error as a number (not a vector)
take length squared

$$\|\vec{e}\|^2 = \vec{e} \cdot \vec{e} = \left\| \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\|^2 = 1$$

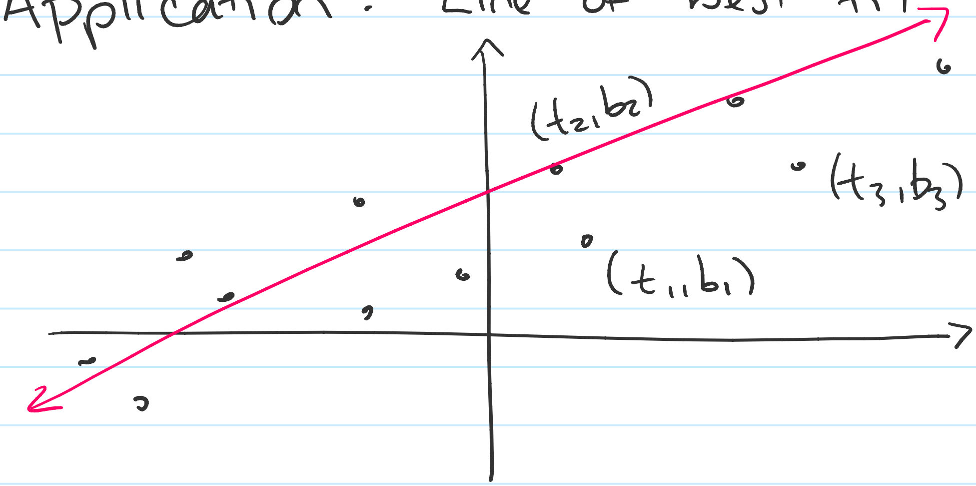
General goal: minimize error $\|\vec{e}\|^2$

Solving $A^TA\vec{x} = A^Tb$ gives a solution minimizing error because it uses the orthogonal projection:



$\|\vec{e}\|$ = distance from $\vec{b}$ to $A\vec{x}$

Application: Line of best fit



Have a bunch of points (t_i, b_i) in $\mathbb{R}^2$

Want to find a line $y = cx + d$ closest fit

If all the points were on the line $y = cx + d$

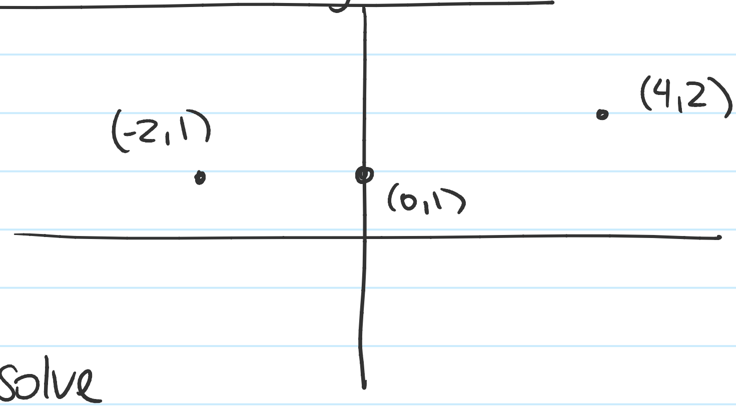
$$\begin{cases} b_1 = c t_1 + d \\ b_2 = c t_2 + d \\ \vdots \\ b_m = c t_m + d \end{cases}$$

$$\rightarrow \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{\vec{b}} = \underbrace{\begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} c \\ d \end{bmatrix}}_{\vec{x}}$$

We know they probably don't really all lie exactly on a line so probably there is no solution $A\vec{x} = \vec{b}$

So do the next best thing: Solve $A^TA\vec{x} = A^Tb$.

Example we can do by hand:



Try to solve

$$-2c + d = 1$$
$$0c + d = 1$$
$$4c + d = 2$$

$$\begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

(you can check there is no solution) so instead:

$$\begin{bmatrix} -2 & 0 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} -2 & 0 & 4 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 20 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 20 & 2 \\ 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \dots = \begin{bmatrix} 5/28 \\ 17/14 \end{bmatrix}$$

So the line of best fit is: $\underline{y = \frac{5}{28}x + \frac{17}{14}}$