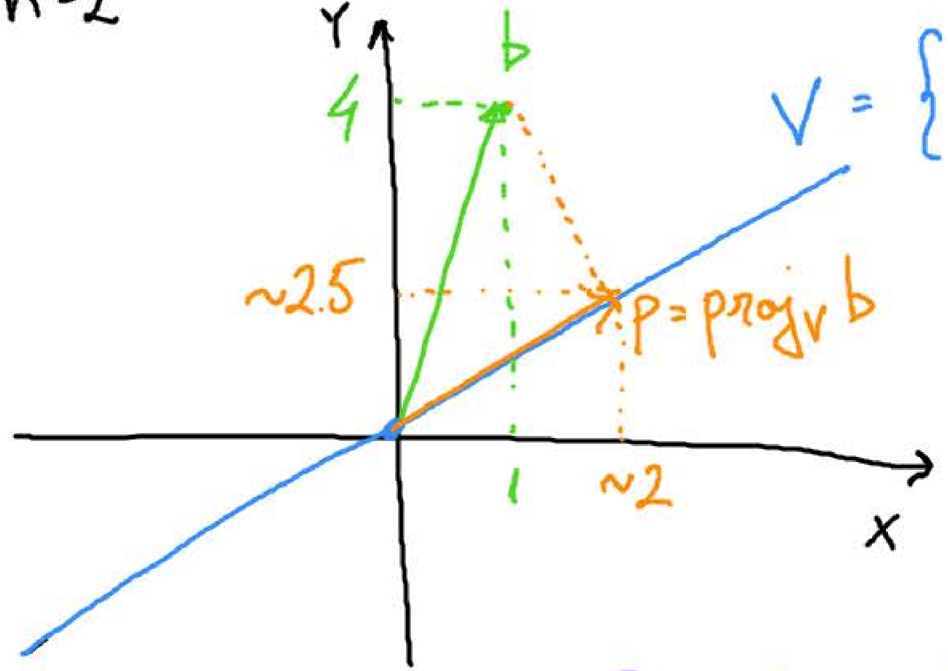
Projection matrices $P_V$

- are symmetric
- $P_V \cdot P_V = P_V$
- projection formula $\boxed{\operatorname{proj}_V b = P_V \cdot b}$

n×n matrices associated to a subspace $V \subset R^n$

n=2




$V = \{ Y = X \cdot \frac{3}{4} \}$

algorithm • write $V = C(A)$ for some matrix $A$

n×k

• $P_V = \underbrace{A}_{n\times k} (\underbrace{A^T}_{k\times n} \underbrace{A}_{n\times k})^{-1} \underbrace{A^T}_{k\times n}$

$A = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ $\quad \left( \underset{\shortparallel}{[25]} \right)^{-1} = [1/25]$

$P_V = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \left( \begin{bmatrix} 4 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} \right)^{-1} \begin{bmatrix} 4 & 3 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 16 & 12 \\ 12 & 9 \end{bmatrix}$

$p = P_V b = \frac{1}{25} \begin{bmatrix} 16 & 12 \\ 12 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 64 \\ 48 \end{bmatrix} \sim \begin{bmatrix} 2.5 \\ 2 \end{bmatrix}$

# "Least Squares" approximation

**Setup** fix $A, b$

approximate $\boxed{Av = b}$ if $b \notin C(A)$

find $v$ such that the error $b - Av$ is as small as possible

i.e. $\| b - Av \|$ should be as small as possible

**How to do this:** pick $p \in C(A)$ s.t.

$\| b - p \|$ is as small as possible

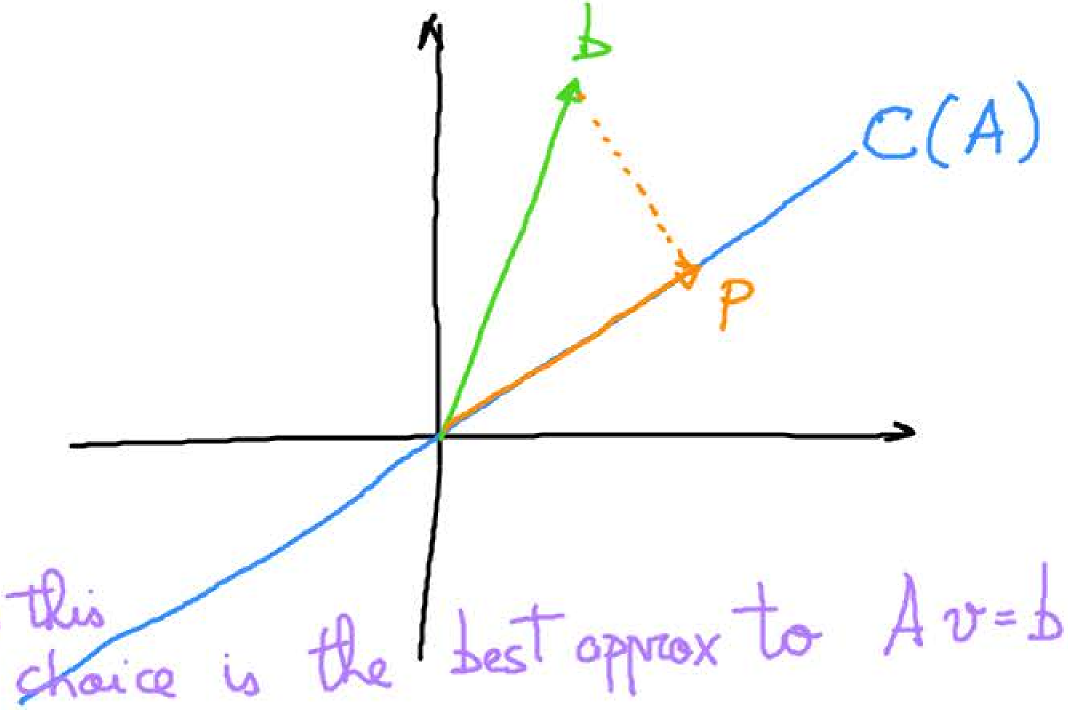
$\Updownarrow$

$p$ should be equal to $\mathrm{proj}_{C(A)} b$

$$p = P_{C(A)} b = A(A^T A)^{-1} A^T \cdot b$$

**What should $v$ be?** $p = Av \longrightarrow v = (A^T A)^{-1} A^T \cdot b$

and this choice is the best approx to $Av = b$



Problem: find $x, y, z$ s.t. $x\begin{bmatrix}1\\1\\0\end{bmatrix} + y\begin{bmatrix}0\\1\\1\end{bmatrix} + z\begin{bmatrix}1\\2\\1\end{bmatrix}$ is as close to $b = \begin{bmatrix}1\\2\\3\end{bmatrix}$ as possible

---

$A\begin{bmatrix}x\\y\\z\end{bmatrix}$ where $A = \begin{bmatrix}1 & 0 & 1\\1 & 1 & 2\\0 & 1 & 1\end{bmatrix}$

goal: find $\begin{bmatrix}x\\y\\z\end{bmatrix} = v$ s.t. $Av$ is as close to $b$ as possible

least squares formula:

choose $v = (A^T A)^{-1} A^T \cdot b$

let $S = A^T A = \begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\\1 & 2 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\1 & 1 & 2\\0 & 1 & 1\end{bmatrix} = \begin{bmatrix}2 & 1 & 3\\1 & 2 & 3\\3 & 3 & 6\end{bmatrix}$

the columns of $S$ are not independent $\Longrightarrow$ $S$ is not invertible
$\Updownarrow$
the columns of $A$ are not independent

but $\begin{bmatrix}1\\2\\1\end{bmatrix} = \begin{bmatrix}1\\1\\0\end{bmatrix} + \begin{bmatrix}0\\1\\1\end{bmatrix}$ implies that

$$x\begin{bmatrix}1\\1\\0\end{bmatrix} + y\begin{bmatrix}0\\1\\1\end{bmatrix} + z\begin{bmatrix}1\\2\\1\end{bmatrix} = \underbrace{(x+z)}_{\alpha}\begin{bmatrix}1\\1\\0\end{bmatrix} + \underbrace{(y+z)}_{\beta}\begin{bmatrix}0\\1\\1\end{bmatrix}$$

equivalently, it suffices to find $\alpha, \beta$ s.t. $\alpha\begin{bmatrix}1\\1\\0\end{bmatrix} + \beta\begin{bmatrix}0\\1\\1\end{bmatrix}$ is as close to $b = \begin{bmatrix}1\\2\\3\end{bmatrix}$ as possible

$\tilde{A} = \begin{bmatrix}1 & 0\\1 & 1\\0 & 1\end{bmatrix}$ we want to approximate $\tilde{A}\begin{bmatrix}\alpha\\\beta\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix}$

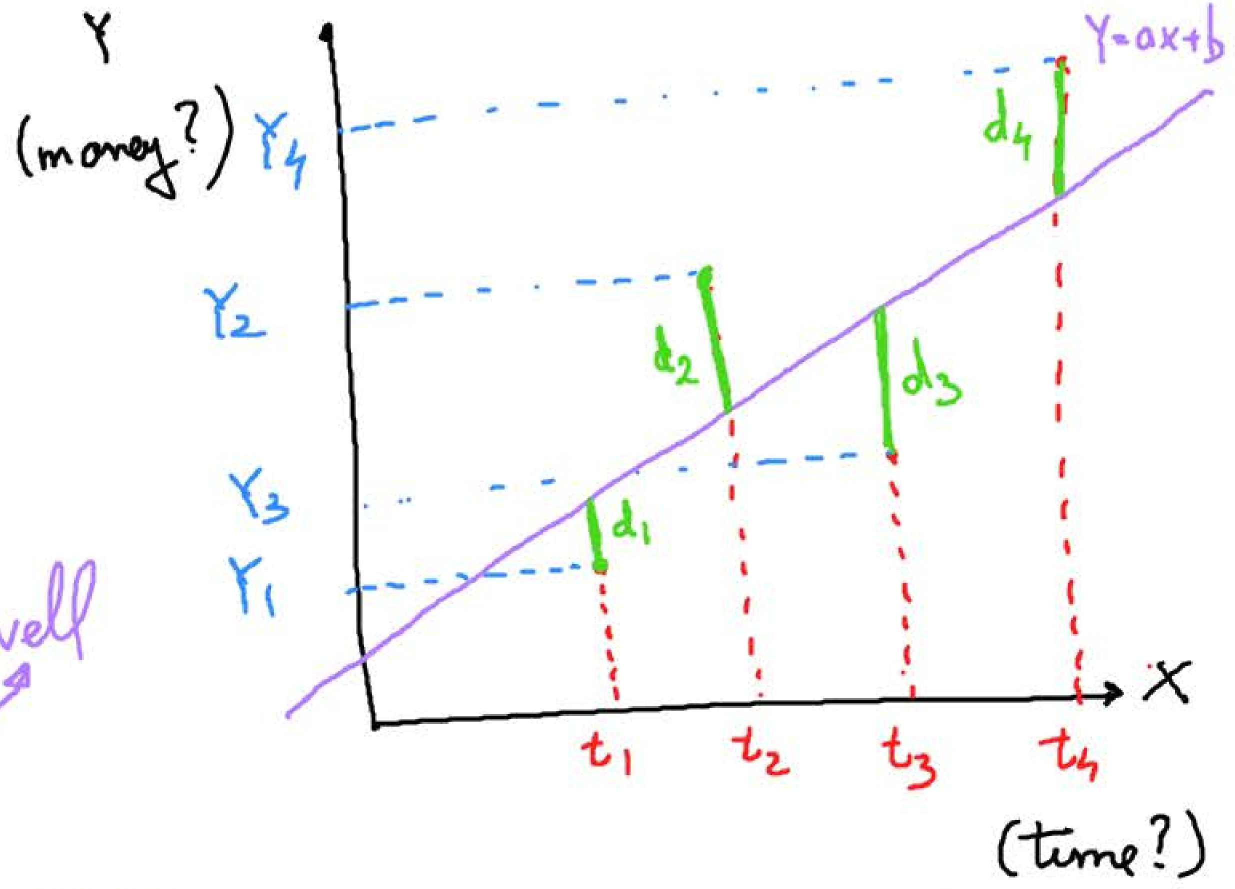
choose $\begin{bmatrix}\alpha\\\beta\end{bmatrix} = (\tilde{A}^T\tilde{A})^{-1}\tilde{A}^T \cdot b$ $\quad \tilde{A}^T\tilde{A} = \begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$

# Application: data points on an $(X,Y)$ plane

if time $\sim$ money you would expect a proportionality between the $t_i$'s and the $Y_i$'s

$\Updownarrow$

all data points on the same line




Goal: fit a line between the data points as well as possible

$Y = ax+b$

choosing a line of best fit comes down to choosing two real numbers $a$ and $b$

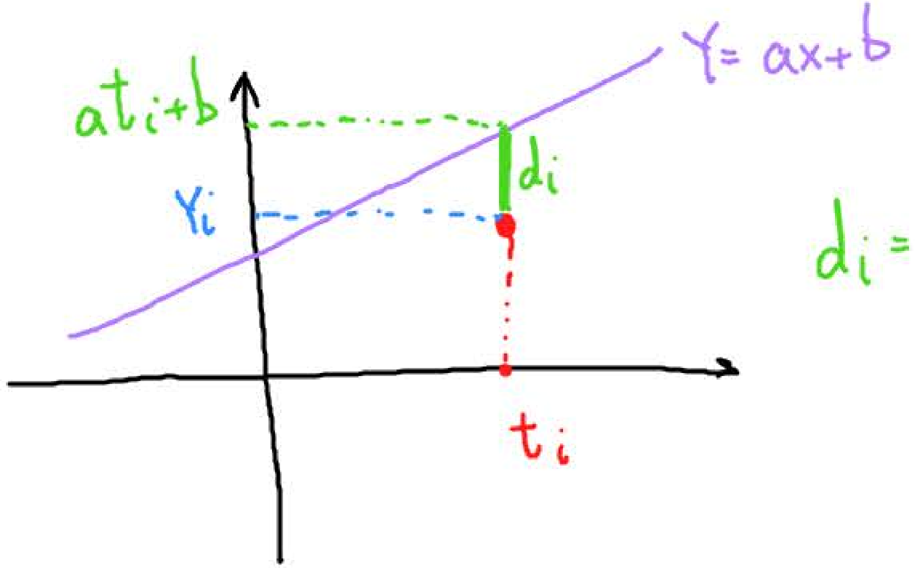
the sum of the squares of the vertical distances from the data points to the line should be as small as possible

$\longrightarrow$ $d_1^2+d_2^2+d_3^2+d_4^2$ should be as small as possible you are given $t_1, t_2, t_3, t_4$ and you want to find $a, b$

$Y_1, Y_2, Y_3, Y_4$

s.t. $d_1^2 + d_2^2 + d_3^2 + d_4^2$ is as small as possible $(Y_1 - at_1 - b)^2 + \dots + (Y_4 - at_4 - b)^2$

what is $d_i$ in terms of $Y_i$ and $t_i$?




$d_i = |Y_i - at_i - b|$

the system

$$\begin{cases} at_1 + b = Y_1 \\ \vdots \\ at_4 + b = Y_4 \end{cases} \iff \underbrace{\begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_4 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{v} = \underbrace{\begin{bmatrix} Y_1 \\ \vdots \\ Y_4 \end{bmatrix}}_{b}$$

least squares solution gives you

$v = \begin{bmatrix} a \\ b \end{bmatrix} = (A^T A)^{-1} A^T \cdot b$ = some formula in terms of $Y_i$'s and $t_i$'s