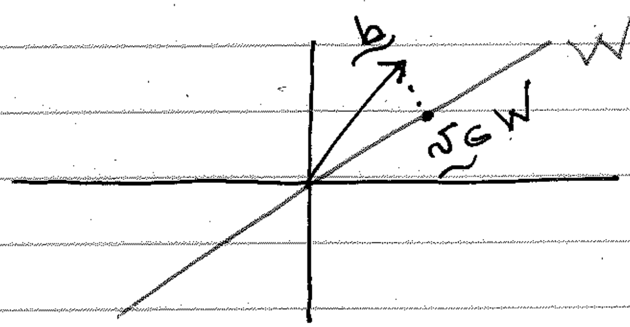
LECTURE 31 | 11/17/14. (1)

Projection Theorem. $V$ - vector space
$\langle \cdot , \cdot \rangle$ - inner product
$b \in V$
$W$ Subspace of $V$, $\dim W < \dim V$




$\underset{\sim}{v} \in W$ is the best approximation of $\underset{\sim}{b}$ in the Subspace $W$: $\min_{w \in W} \| \underset{\sim}{w} - \underset{\sim}{b} \| = \| \underset{\sim}{v} - \underset{\sim}{b} \|$.

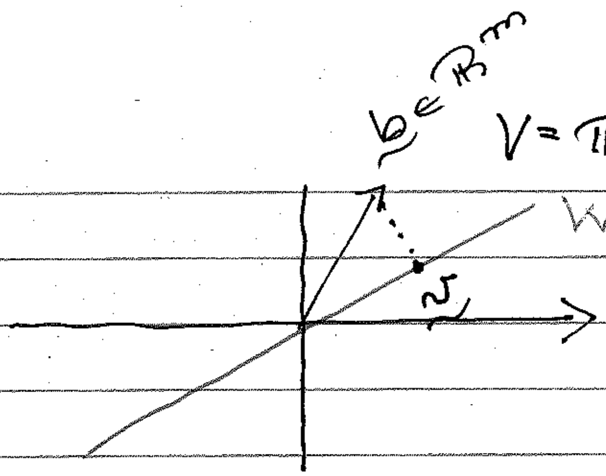
Two conditions are equivalent:

(1) $\underset{\sim}{v}$ is the closest element to $\underset{\sim}{b} \in V$
$$\| \underset{\sim}{b} - \underset{\sim}{v} \| \leq \| \underset{\sim}{b} - \underset{\sim}{w} \| \quad \forall w \in W$$

(2) $\underset{\sim}{v}$ is the orthogonal projection of $\underset{\sim}{b}$ onto $W$
$$\langle \underset{\sim}{b} - \underset{\sim}{v}, \underset{\sim}{w} \rangle = 0 \quad \forall w \in W.$$

Normal equations:
Consider $V = \mathbb{R}^m$ and $W = \operatorname{span}\{\underset{\sim}{a_i}\}_{i=1}^{n}$ where $\underset{\sim}{a_i} \in V$ and $n < m$.

Consider $\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle = u^T C \underset{\sim}{v}$, where $C$ is positive definite




$W = \text{Span}\{\underset{\sim}{a_i}\}_{i=1}^{n}$
$= \left\{ \sum_{i=1}^{n} x_i \underset{\sim}{a_i} ,\ x_i \in \mathbb{R} \right\}$

Question: What is the best approximation $\underset{\sim}{v} \in W$ to the vector $\underset{\sim}{b} \in \mathbb{R}^m$?

$$\underset{\sim}{v} \in W \Rightarrow \underset{\sim}{v} = \sum_{i=1}^{n} x_i \underset{\sim}{a_i} = \underbrace{[\underset{\sim}{a_1}\ \underset{\sim}{a_2} \cdots \underset{\sim}{a_n}]}_{A \in \mathbb{R}^{m \times n}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{x \in \mathbb{R}^n}$$

$x_1, x_2, \dots, x_n$ are the coordinates of $\underset{\sim}{v}$ with respect to the basis $\{a_i\}_{i=1}^{n}$

- Since $\underset{\sim}{b} \notin \text{rgn}(A) \Rightarrow A\underset{\sim}{x} = \underset{\sim}{b}$ is not solvable
- Recall if $b \in \text{rgn}(A) \Rightarrow A\underset{\sim}{x} = \underset{\sim}{b}$ is solvable and then $\underset{\sim}{v} = \underset{\sim}{b}$.

How do we find $v$ when $\underset{\sim}{b} \notin \text{rgn}(A)$? Using the projection theorem

$$\langle \underset{\sim}{b} - \underset{\sim}{v}, \underset{\sim}{w} \rangle = 0 \quad \forall \underset{\sim}{w} \in W$$
$$\Longleftrightarrow \langle \underset{\sim}{b} - \underset{\sim}{v}, \underset{\sim}{a_i} \rangle = 0 \quad \forall i = 1, \dots, n$$
$$\Longleftrightarrow \langle \underset{\sim}{b}, \underset{\sim}{a_i} \rangle = \langle \underset{\sim}{v}, \underset{\sim}{a_i} \rangle = \langle Ax, a_i \rangle, \ i=1,\dots,n$$

$$\Rightarrow \langle a_i, Ax \rangle = \langle a_i, b \rangle, \ i = 1, \dots, n$$
$$\underset{\sim}{a_i}^T C A \underset{\sim}{x} = \underset{\sim}{a_i}^T C \underset{\sim}{b}.$$

$$y = \alpha + \beta t$$

How do we choose $\alpha$ and $\beta$?

The error is $e_i =: y_i - (\alpha + \beta t_i)$, $i = 1, \ldots, m$.
We can write it as a system of equations.

$$\underset{\sim}{e} = \underset{\sim}{y} - A\underset{\sim}{x},$$

where $\underset{\sim}{e} = \begin{pmatrix} e_1 \\ \vdots \\ e_m \end{pmatrix}$, $\underset{\sim}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$; $A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ & \vdots \\ 1 & t_m \end{pmatrix}$

$$\underset{\sim}{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Notation: $\underset{\sim}{e} \rightarrow$ error vector
$\underset{\sim}{y} \rightarrow$ data vector

If $y \in \text{rgn}(A) \Rightarrow$ we can solve exactly and $\underset{\sim}{e} = 0$
(all data points lie on the straight line)
In applications, $\underset{\sim}{y} \notin \text{rgn}(A)$. Then, we need to minimize

$$\text{ERROR} = \| e \| = (e_1^2 + \cdots + e_m^2)^{1/2}$$

Least square approximation!

Take $C = I \Rightarrow$ Identity matrix

$$a_i^T A \underset{\sim}{x} = a_i^T \underset{\sim}{b}, \quad i = 1, \dots, m$$

$$\iff A^T A \underset{\sim}{x} = A^T \underset{\sim}{b} \rightarrow \text{Normal equations. } (\#)$$

Remarks

1) (#) always has a solution. Exercise.
2) Suppose $\{a_i\}_{i=1}^m$ is l.i $\Rightarrow A^T A$ is positive def.

$$\underset{\sim}{x}^T A^T A \underset{\sim}{x} = (A\underset{\sim}{x})^T A\underset{\sim}{x} = \|A\underset{\sim}{x}\|^2 \geq 0.$$

$\underset{\sim}{x} = 0 \Rightarrow A\underset{\sim}{x} = 0$, and $A\underset{\sim}{x} = 0 \Rightarrow \underset{\sim}{x} = 0$

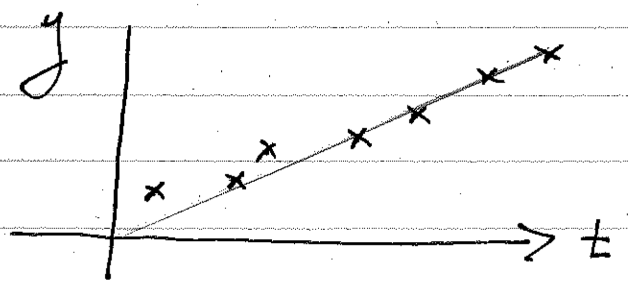
Since $\{a_i\}_{i=1}^m$ is l.i.

$\Rightarrow$ (#) has a unique solution.

§4.4 Data fitting and interpolation.

Suppose we are running an experiment, and we get at time $t_i$ the measurement $y_i$:

$$(t_1, y_1), (t_2, y_2), \dots, (t_m, y_m)$$



In mathematical terms, $\bar{x} = (\alpha \ \beta)^T$ minimizes

$$\| e \| = \| A x - y \|$$

if and only if $\underline{x}$ solves the normal equations

$$(A^T A) \underline{x} = A^T y$$

If columns of $A$ are l.i. $\Rightarrow$ $A^T A$ is invertible and then

$$\boxed{\underset{\sim}{x} = (A^T A)^{-1} A^T \underline{y}.}$$

Now,

$$A^T A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_m \end{pmatrix} \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \\ 1 & t_m \end{pmatrix} = \begin{pmatrix} m & \sum_1 t_i \\ \sum_1 t_i & \sum_1 (t_i)^2 \end{pmatrix}$$

$$= m \begin{pmatrix} 1 & \bar{t} \\ \bar{t} & \overline{t^2} \end{pmatrix}$$

$$A^T y = \begin{pmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_m \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} \sum_1 y_i \\ \sum_1 t_i y_i \end{pmatrix} = m \begin{pmatrix} \bar{y} \\ \overline{ty} \end{pmatrix}$$

where, $\bar{t} = \frac{1}{m} \sum_{i=1}^{m} t_i$, $\bar{y} = \frac{1}{m} \sum_{i=1}^{m} y_i$

for instance.