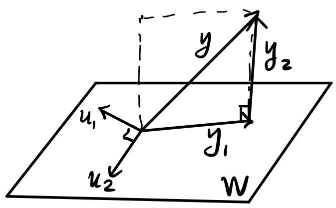
- $V \supseteq W$ inner product space.



$\{u_1, \dots, u_k\}$ o.n.b. of W

$y = y_1 + y_2 \quad y_1 \in W, \ \underline{y_2 \perp W}$

$\langle y_2, w \rangle = 0, \ \forall w \in W$

$y_1 = \mathrm{Proj}_W y = \sum_{i=1}^{k} \langle y, u_i \rangle u_i$ satisfies

$$\|y - y_1\| \le \|y - w\| \quad \forall w \in W$$

i.e. y_1 is the closest vector to y in W.

- A: $m \times n$ matrix $\quad A = (v_1 \ v_2 \cdots v_n) \quad v_i \in \mathbb{R}^m$.

$$A: \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto Ax \qquad \text{Set } V_1 = \mathbb{R}^n, \ V_2 = \mathbb{R}^m$$

Let $V = V_2 = \mathbb{R}^m$, $W = R(A) = \mathrm{Span}\{v_1, v_2, \dots, v_n\}$.

Let $y \in \mathbb{R}^m$. Find $y_1 = Ax_1$ that is closest to y in $R(A)$.

$Ax_1 = y_1$ is then characterized by the condition:

$$\langle y - Ax_1, Ax \rangle = 0 \quad \forall x \in \mathbb{R}^n.$$

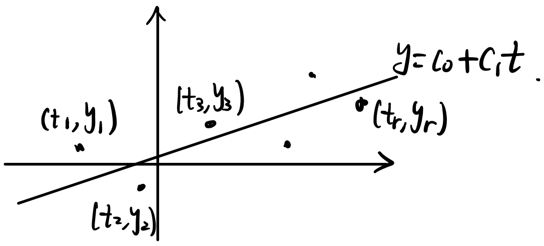
$$\langle y - Ax_1, Ax \rangle = \langle Ax, y - Ax_1 \rangle_{\mathbb{R}^m} = (Ax)^T (y - Ax_1) = (x^T A^T)(y - Ax_1) = x^T (A^T y - A^T A x_1)$$

$$x^T (A^T y - A^T A x_1) = \langle x, A^T y - A^T A x_1 \rangle_{\mathbb{R}^n} = 0$$

$$\Leftrightarrow A^T y - A^T A x_1 = 0 \Leftrightarrow x_1 = (A^T A)^{-1} A^T y.$$

(Assume $A^T A$ invertible).

Ex:



Find linear function $y = c_0 + c_1 t$ that best approximates the measurements $\{(t_i, y_i),\ 1 \le i \le r\}$.

W.r.T. the Error:
$$E = |c_0 + c_1 t_1 - y_1|^2 + |c_0 + c_1 \cdot t_2 - y_2|^2 + \cdots + |c_0 + c_1 t_r - y_r|^2$$
$$= \left\| \begin{pmatrix} c_0 + c_1 t_1 \\ c_0 + c_1 t_2 \\ \vdots \\ c_0 + c_1 t_r \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix} \right\|^2 = \left\| \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_r \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix} \right\|^2$$
$$= \|Ax - y\|^2.$$

$$A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_r \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix}. \quad \Rightarrow \quad \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = (A^T A)^{-1} A^T y$$

$y = c_0 + c_1 t$ gives the best linear approximation

Ex: $\{(t_i, y_i)\} = \{(-2, 4), (-1, 3), (0, 1), (1, -1), (2, -3)\}$.

$$A = \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad y = \begin{pmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{pmatrix}. \quad A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 10 \end{pmatrix}.$$

$$A^T y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ 1 \\ -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 4 \\ -18 \end{pmatrix}. \quad (A^T A)^{-1} A^T y = \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{10} \end{pmatrix} \begin{pmatrix} 4 \\ -18 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 4 \\ -9 \end{pmatrix} = \underbrace{\begin{pmatrix} c_0 \\ c_1 \end{pmatrix}}_{= x_1}$$

$$\Rightarrow y_1 = A x_1 = \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \frac{1}{5} \begin{pmatrix} 4 \\ -9 \end{pmatrix} = \frac{1}{5} \cdot \begin{pmatrix} 22 \\ 13 \\ 4 \\ -5 \\ -14 \end{pmatrix}.$$

$$\text{Error} = \|y-y_1\|^2 = \left\| \begin{pmatrix} 4\\3\\1\\-1\\-3 \end{pmatrix} - \frac{1}{5}\cdot\begin{pmatrix} 22\\13\\4\\-5\\-14 \end{pmatrix} \right\|^2 = \frac{1}{25}\cdot \left\| \begin{pmatrix} -2\\2\\1\\0\\-1 \end{pmatrix} \right\|^2 = \frac{4+4+1+1}{25} = \frac{2}{5}.$$

To find y_1, we can also project to $W = R(A) = \text{Span}\left\{ \underbrace{\begin{pmatrix} 1\\1\\1\\1\\1 \end{pmatrix}}_{w_1}, \underbrace{\begin{pmatrix} -2\\-1\\0\\1\\2 \end{pmatrix}}_{w_2} \right\}$

$v_1 = w_1$. $\|v_1\|^2 = 5$

$$v_2 = w_2 - \frac{\langle w_2, v_1\rangle}{\|v_1\|^2} v_1 = \begin{pmatrix} -2\\-1\\0\\1\\2 \end{pmatrix} - \frac{0}{5} * = \begin{pmatrix} -2\\-1\\0\\1\\2 \end{pmatrix}. \qquad y = \begin{pmatrix} 4\\3\\1\\-1\\-3 \end{pmatrix}$$

$$y_1 = \langle y, u_1\rangle u_1 + \langle y, u_2\rangle u_2 = \frac{\langle y, v_1\rangle}{\|v_1\|^2} v_1 + \frac{\langle y, v_2\rangle}{\|v_2\|^2} v_2$$

$$= \frac{4}{5}\begin{pmatrix} 1\\1\\1\\1\\1 \end{pmatrix} + \frac{-8-3-1-6}{4+1+1+4}\cdot\begin{pmatrix} -2\\-1\\0\\1\\2 \end{pmatrix} = \frac{4}{5}\cdot\begin{pmatrix} 1\\1\\1\\1\\1 \end{pmatrix} - \frac{9}{5}\begin{pmatrix} -2\\-1\\0\\1\\2 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 22\\13\\4\\-5\\-14 \end{pmatrix} \checkmark$$

$\underset{m\times n}{A}: \mathbb{R}^n \to \mathbb{R}^m$, $A^T: \mathbb{R}^m \to \mathbb{R}^n$ satisfy:

$$\langle Ax, y\rangle_{\mathbb{R}^m} = (Ax)^T\cdot y = (x^T A^T) y = x^T(A^T y) = \langle x, A^T y\rangle_{\mathbb{R}^n}$$

A^T = transpose of A represents the adjoint of the linear transformation $A: \mathbb{R}^n \to \mathbb{R}^m$.

Let $T: V_1 \to V_2$ be a linear transformation.

Assume V_i is equipped with an inner product structure $\langle \cdot, \cdot \rangle_{V_i}$ $i = 1, 2$.

· Def-Thm: There exists a linear transformation $T^*: V_2 \to V_1$ that satisfies the condition: $\langle Tx, y \rangle_{V_2} = \langle x, T^* y \rangle_{V_1}$, $\forall x \in V_1, y \in V_2$.

Such a linear transformation T^* is unique and is called the adjoint of T w.r.t. the inner products $\langle \cdot, \cdot \rangle_{V_1}$ and $\langle \cdot, \cdot \rangle_{V_2}$.

(T^* depends crucially on these two inner products).

· Choose o.n.b. $\alpha = \{u_1, u_2, \cdots, u_n\}$ for V_1 and o.n.b. $\beta = \{u_1', u_2', \cdots, u_m'\}$ for V_2

We prove: $[T^*]_\beta^\alpha = \left([T]_\alpha^\beta\right)^T$ (only true w.r.t. orthonormal bases)

Proof: i-th column of $[T^*]_\beta^\alpha = [T^* u_i']_\alpha$

$$\underset{\substack{\text{the Fourier coeff.}\\ \text{are given by inner prod.}}}{\overset{\substack{\text{because } \{u_1, \cdots, u_n\}\\ \text{are o.n.b.}}}{=\!=}} \begin{pmatrix} \langle T^* u_i', u_1 \rangle_{V_1} \\ \langle T^* u_i', u_2 \rangle_{V_1} \\ \cdots \\ \langle T^* u_i', u_m \rangle_{V_1} \end{pmatrix} \overset{\text{def. of } T^*}{=\!=} \begin{pmatrix} \langle u_i', T u_1 \rangle_{V_2} \\ \langle u_i', T u_2 \rangle_{V_2} \\ \cdots \\ \langle u_i', T u_n \rangle_{V_2} \end{pmatrix}$$

$$[T]_{\alpha}^{\beta} = \left[[T(u_1)]_{\beta}\ [T(u_2)]_{\beta} \cdots [T(u_n)]_{\beta} \right) = A$$

$$= \begin{pmatrix} \langle Tu_1, u_1' \rangle & \langle Tu_2, u_1' \rangle & & \langle Tu_n, u_1' \rangle \\ \langle Tu_1, u_2' \rangle & \langle Tu_2, u_2' \rangle & \cdots & \langle Tu_n, u_2' \rangle \\ \vdots & \vdots & & \vdots \\ \langle Tu_1, u_i' \rangle_{V_2} & \langle Tu_2, u_i' \rangle_{V_2} & & \langle Tu_n, u_i' \rangle_{V_2} \\ \vdots & \vdots & & \vdots \\ \langle Tu_1, u_m' \rangle & \langle Tu_2, u_m' \rangle & & \langle Tu_n, u_m' \rangle \end{pmatrix}$$

We compare to get: $[T^*]_{\beta}^{\alpha} = \left([T]_{\alpha}^{\beta}\right)^{T}$ ← transpose.