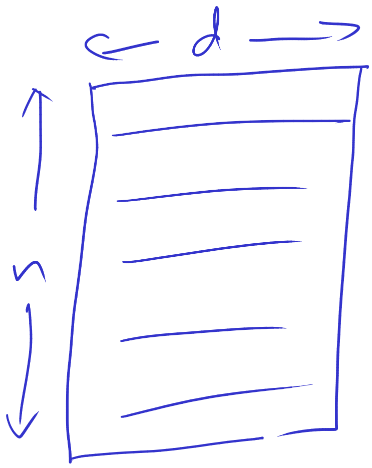
SVD and Best fit Subspace

$x_1 \; \dots \; x_n \in \mathbb{R}^d \qquad n >> d$

$M =$ 

$r = \text{rank}(M)$

$$M = \sum_{i=1}^{r} \sqrt{\lambda_i}\, u_i v_i^T$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \dots \geq \lambda_r > 0$

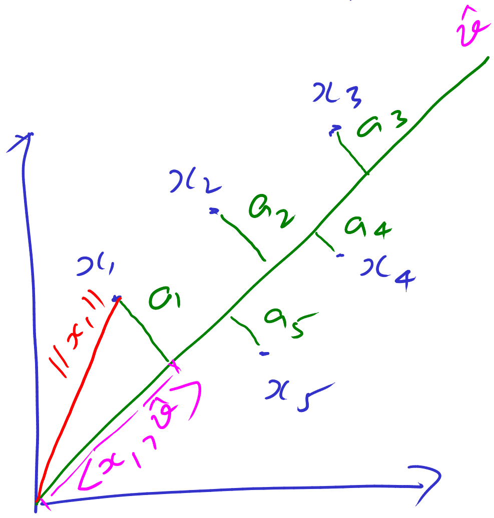
$\underset{\substack{\cap \\ \mathbb{R}^d}}{v_i}$'s are O.N.E.V of $M^T M \in \mathbb{R}^{d \times d}$ with eigenvalue λ_i

$$\underset{\substack{\cap \\ \mathbb{R}^n}}{u_i} = \frac{M v_i}{\sqrt{\lambda_i}}$$

O.N.E.V of $M M^T \in \mathbb{R}^{m \times m}$

$\text{Span}(v_1)$ = Best fit 1-dim (line through origin) subspace

$\text{Span}(v_1 \ldots v_k)$ = Best fit k-dim subspace



Best fit means $\sum_{i=1}^{n} a_i^2$ is minimized

$$= \sum_{i=1}^{n} \left(\|x_i\|^2 - \langle x_i, \hat{v} \rangle^2 \right)$$

$$\hat{v} = \underset{\|v\|=1}{\operatorname{argmin}} \sum_{i=1}^{n} \left(\|x_i\|^2 - \langle x_i, v \rangle^2 \right)$$

$$= \underset{\|v\|=1}{\operatorname{argmax}} \sum_{i=1}^{n} \langle x_i, v \rangle^2$$

$$= \underset{\|v\|=1}{\operatorname{argmax}} \|Mv\|^2$$

$$M \begin{bmatrix} v \end{bmatrix} = \begin{bmatrix} \langle x_1, v \rangle \\ \vdots \\ \langle x_n, v \rangle \end{bmatrix}$$

$$= \underset{\|v\|=1}{\arg\max} \, (Mv)^T (Mv) \to v^T \boxed{M^T M} v$$

s.D

$$= \underset{\|v\|=1}{\arg\max} \, v^T \left(\sum_{i=1}^{n} \lambda_i v_i v_i^T \right) v$$

$\langle v_i, v \rangle$

$= v_i \cdot v$

$$= \underset{\|v\|=1}{\arg\max} \sum_{i=1}^{n} \lambda_i \langle v_i, v \rangle^2$$

$v = (v_1 + v_2)/\sqrt{2}$

$\frac{\lambda_1}{2} + \frac{\lambda_2}{2}$

$\|v\| = 1$

$\Downarrow$

$\sum_i \langle v_i, v \rangle^2 = 1$

$= v_1$ with value λ_1

Hence Span(v_1)

= Best fit 1-dim subspace $\boxtimes$

$$v_1 = \underset{\|v\|=1}{\arg\max} \|Mv\|^2$$

Claim:

$$v_2 = \underset{\substack{v \in v_1^\perp \\ \|v\|=1}}{\arg\max} \|Mv\|^2$$

Proof:

$$\longrightarrow \underset{\substack{v \in v_1^\perp \\ \|v\|=1}}{\operatorname{argmax}} \sum_{i=1}^{n} \lambda_i \langle v_i, v \rangle^2$$

$$= \underset{\substack{v \in v_1^\perp \\ \|v\|=1}}{\operatorname{argmax}} \sum_{i=2}^{n} \lambda_i \langle v_i, v \rangle^2$$

$$= v_2 \text{ with value } \lambda_2 \qquad \boxtimes$$

Claim

$$v_k = \underset{\substack{v \in \operatorname{span}(v_1 \ldots v_{k-1})^\perp \\ \|v\|=1}}{\operatorname{argmax}} \|Mv\|^2$$

Claim:

Best fit 2-dim subspace $= \mathrm{Span}(v_1, v_2)$

Proof:

Let W be the best fit 2-dim subspace.

w_1, w_2 is O.N.B for W

We can chose $w_2 \in v_1^{\perp}$

$V = \mathrm{span}(v_1) \oplus v_1^{\perp}$

$\Downarrow$

$\exists\, w_2 \in W \cap v_1^{\perp}$ s.t $w_2 \neq 0$

$$w_1, w_2 = \underset{\substack{w_1^*, w_2^* \\ \|w_1^*\|, \|w_2^*\| = 1 \\ w_1^* \perp w_2^*}}{\arg\max} \|M w_1^*\|^2 + \|M w_2^*\|^2$$

sum of squares of proj onto W

$$\|M v_1\|^2 + \|M v_2\|^2 \leq \|M w_1\|^2 + \|M w_2\|^2$$

1.

$$\|Mv_1\|^2 \geq \|Mw_1\|^2$$

$$\|Mv_2\|^2 \geq \|Mw_2\|^2$$

$$\|Mv_1\|^2 + \|Mv_2\|^2 = \|Mw_1\|^2 + \|Mw_2\|^2$$

$\boxtimes$

General Case.

$w_k \in \mathrm{Span}(v_1 \ldots v_{k-1})^{\perp}$

(by induction)

Base case ✓

Best fit 1 dim = span(v_1)

Inductive

Assume that

Best Fit $k-1$ dim = span$(v_1 \ldots v_{k-1})$

$\Rightarrow$ Best Fit k dim = span$(v_1 \ldots v_k)$