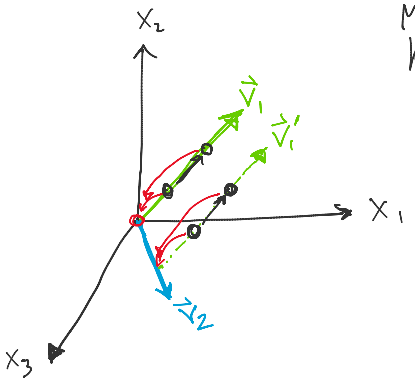
Section B

Wednesday, April 12, 2023 12:31 PM



$M = 3$
$K = 2$

$$\vec{U}^{(i)} = V \vec{X}^{(i)}$$

$$U_1^{(i)} = \vec{V}_1^T \vec{X}^{(i)}$$
$$U_2^{(i)} = \vec{V}_2^T \vec{X}^{(i)}$$
$$\vdots$$
$$U_k^{(i)} = \vec{V}_k^T \vec{X}^{(i)}$$

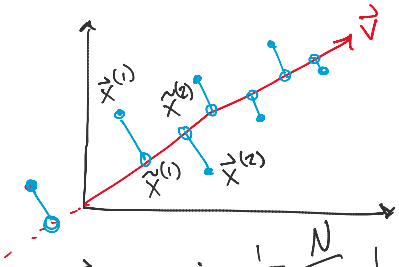
$$V = \begin{bmatrix} \text{—} \ \vec{V}_1^T \ \text{—} \\ \text{—} \ \vec{V}_2^T \ \text{—} \\ \vdots \\ \text{—} \ \vec{V}_k^T \ \text{—} \end{bmatrix}$$

Objectives for PCA

What is the first principal component for PCA?

$\vec{V}_1$

(1) Minimize the Reconstruction Error



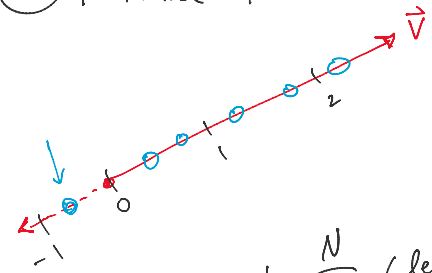
$$\vec{V}_1 = \underset{\vec{V}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \text{distance}\left(\vec{x}^{(i)}, \tilde{x}^{(i)}\right)^2$$

$\tilde{x}^{(i)}$ = (vector projection of $\vec{x}^{(i)}$ onto $\vec{V}$) $= \left(\vec{V}^T \vec{x}^{(i)}\right) \vec{V} / \|V\|_2^2$

$$= \underset{\vec{V}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \left\| \vec{x}^{(i)} - \left(\text{vector projection of } \vec{x}^{(i)} \text{ onto } \vec{V}\right) \right\|_2^2$$

$$= \underset{\vec{V} \text{ s.t. } \|\vec{V}\|_2 = 1}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \left\| \vec{x}^{(i)} - \left(\vec{V}^T \vec{x}^{(i)}\right) \vec{V} \right\|_2^2$$

(2) Maximize the Variance



$$\vec{V}_1 = \underset{\vec{V}}{\operatorname{argmax}} \frac{1}{N} \sum_{i=1}^{N} \left(\text{length of vector proj. of } x^{(i)} \text{ onto } \vec{V}\right)^2$$

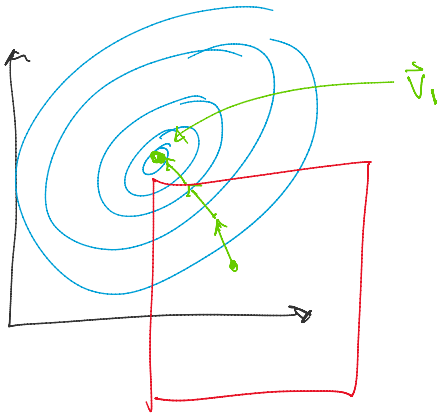
$$= \underset{\vec{V} \text{ s.t. } \|\vec{V}\| = 1}{\operatorname{argmax}} \frac{1}{N} \sum_{i=1}^{N} \left(\vec{V}^T \vec{x}^{(i)}\right)^2$$

$$= \underset{\vec{v} \text{ s.t. } \|\vec{v}\|_2 = 1}{\operatorname{argmax}} \frac{1}{N} \sum_{i=1} (\vec{v}^T \vec{x}^{(i)})$$

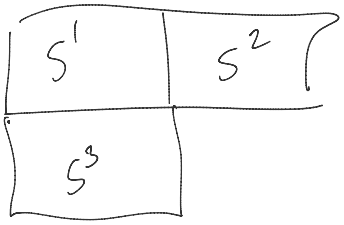
$$= \underset{\vec{v} \text{ s.t. } \|\vec{v}\|_2 = 1}{\operatorname{argmax}} \frac{1}{N} (\vec{v}^T X^T)(X \vec{v})$$

$$= \underset{\vec{v} \text{ s.t. } \|\vec{v}\|_2 = 1}{\operatorname{argmax}} \vec{v}^T \Sigma \vec{v}$$

b/c $\Sigma = \frac{1}{N} X^T X$



OH



$p(s^3 \mid s^1, \rightarrow) = 0.5$

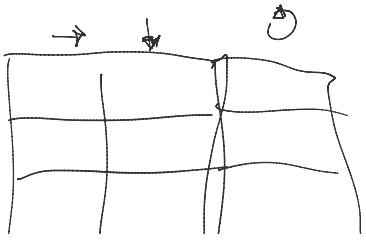
$p(s^2 \mid s^1, \rightarrow) = 0.5$

$(s^1, \rightarrow, 7, s^3)$

$(s^1, \rightarrow, 7, s^2)$

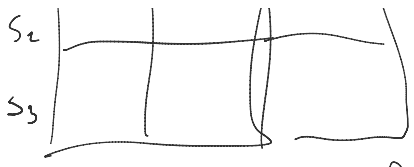
$Q(s^1, \rightarrow) = 7 + \gamma Q(s^3, \cdot)$, Q

$7 + \gamma Q(s^2, \cdot)$



2 of 3

$Q(s^1, \rightarrow) = 7 + \lambda Q(s^2, \circlearrowleft)$



$Q(s^1, \rightarrow) = (1-\alpha_t)\, Q(s^1, \rightarrow) + (\alpha_t)(7 + Q(s^{2 \text{ or } 3}, \circlearrowleft))$

$Q^*(s, a) = \max_a R(s,a) + \lambda \sum_{s'} p(s'|s,a) \left[\max_{a'} Q^*(s', a') \right]$

If $T = 1$

$ER = \vec{e} = [ER(1), ER(2), \ldots, ER(k)]$

$\vec{p} = \text{softmax}(\vec{e}/T)$

$P(a) = p_a$

$P(2) = p_2$