Introduction to Dimensionality Red.

$x \in \mathbb{R}^n$

$x = V\underset{\sim}{\alpha} = \alpha_1 \underline{u}^{(1)} + \dots + \alpha_n u^{(n)}.$

$[x]_\beta = \alpha \in \mathbb{R}^n.$

given $V^T V = I$

$\alpha = V^T x$

$x = V \underbrace{V^T x}$

Now let $V_d = [u^{(1)} | \dots | u^{(d)}]$

$\hat{V}_d = [u^{(d+1)} | \dots | u^{(n)}]$

$V V^T = V_d V_d^T + \hat{V}_d \hat{V}_d^T$

$$x = V_d V_d^T x + \hat{V}_d \hat{V}_d^T x$$

$$\mathbb{P} = V_d V_d^T$$

$$\mathbb{P}^2 = V_d \underbrace{V_d^T V_d}_{I} V_d^T = \mathbb{P}.$$

So $\mathbb{P} = V_d V_d^T$ is an orth. proj. onto $R(V_d)$.

$$V V^T = I = V_d V_d^T + \hat{V}_d \hat{V}_d^T$$

$\hat{V}_d \hat{V}_d^T = I - \mathbb{P}$ comp. proj.

$$x = \mathbb{P} x + (I - \mathbb{P}) x$$

$$V = R(\mathbb{P}) \oplus R(I - \mathbb{P})$$

$$\mathbb{P}^T = \mathbb{P}. \Rightarrow \mathbb{P}x \perp (\mathbb{I}-\mathbb{P})x$$

$$\begin{aligned}(\mathbb{P}x)^T(\mathbb{I}-\mathbb{P})x &= x^T\mathbb{P}^T(\mathbb{I}-\mathbb{P})x \\ &= x^T(\mathbb{P}-\mathbb{P}^2)x \\ &= x^T 0 x = 0.\end{aligned}$$

$\therefore \quad \|x\|^2 = \|\mathbb{P}x\|^2 + \|(\mathbb{I}-\mathbb{P})x\|^2.$

recall $\|x\|^2 = \langle x, x\rangle.$

<u>Novelty Detection</u> (Kohonen 1984)

Build $\mathbb{P}$ from "nominal data" and introduce

$x \leftarrow$ new observation

$$\underbrace{\mathbb{P}x}_{\text{"nominal"}} \quad \perp \quad \underbrace{(\mathbb{I}-\mathbb{P})x}_{\text{novelty}}$$

$\hat{x}$ $\quad$ $\hat{x}_\perp$

e.g.) Consider $V = [u^{(1)} \mid u^{(2)}]$

where $u^{(1)} = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$, $u^{(2)} = \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$

let $x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$$\alpha = V^T x = \begin{pmatrix} 2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4/\sqrt{6} \\ 1/\sqrt{3} \end{pmatrix}$$

$$\hat{x} = V\alpha = \left(\frac{4}{\sqrt{6}}\right) \cdot \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \left(\frac{1}{\sqrt{3}}\right)\left(\frac{1}{\sqrt{3}}\right)\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{4}{6}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Construction of $\mathbb{P}$ from data

Given $\underline{X} = [x^{(1)} \mid x^{(2)} \mid \cdots \mid x^{(k)}]$

with each $x^{(i)} \in \mathbb{R}^n$

① $\mathbb{P} = \underline{X}(\underline{X}^T \underline{X})^{-1} \underline{X}^T$

ex: show $\mathbb{P}^T = \mathbb{P}$, $\mathbb{P}^2 = \mathbb{P}$.

① Find an o.n. basis V for $R(\underline{X})$

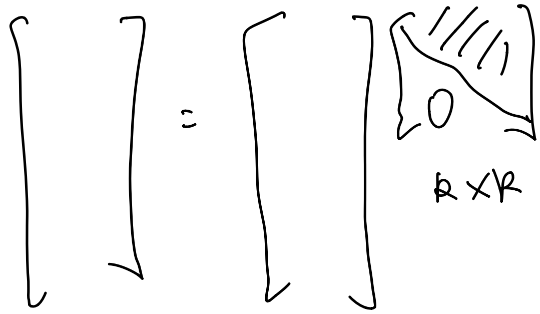
a) QR

b) SVD / PCA.

QR Decomposition

$$\underset{n\times k}{X} = \underset{n\times k}{Q}\,\underset{k\times k}{R}$$

where $Q^TQ = I$



R upper triangular.

$$x^{(1)} = r_{11} q^{(1)} \qquad r_{11} = x^T q^{(1)}$$

$$x^{(2)} = r_{21} q^{(1)} + r_{22} q^{(2)}$$

Construct the Q.

$$q^{(1)} = x^{(1)} / \|x^{(1)}\|$$

$$q^{(2)} = \left(I - q^{(1)} q^{(1)T}\right) x^{(2)} / \| \qquad \|$$

$$q^{(3)} = \left(I - q^{(2)} q^{(2)^T}\right)\left(I - q^{(1)} q^{(1)^T}\right) x^{(3)} / \| \cdot \|$$

$$= \left(I - q^{(2)} q^{(2)^T} - q^{(1)} q^{(1)^T}\right) \frac{\cancel{X}^{(3)}}{\| \cdot \|}$$

$$= \frac{x^{(3)}}{\| \quad \|} - \frac{q^{(2)^T} x^{(3)}}{\| \quad \|} q^{(2)} - \frac{q^{(1)^T} x^{(3)}}{\| \quad \|} q^{(1)}$$

$$x^{(3)} = \underbrace{\| \quad \|}_{r_{33}} q^{(3)} + \underbrace{q^{(2)^T} x^{(3)}}_{r_{23}.} q^{(2)} + \underbrace{q^{(1)^T} x^{(3)}}_{r_{13}} q^{(1)}$$

Principal Component Analysis.

$$x = \underbrace{\alpha_1 u^{(1)} + \alpha_2 u^{(2)} + \cdots + \alpha_d u^{(d)}}_{\text{keep}} + \alpha_{d+1} u^{(d+1)} + \cdots$$

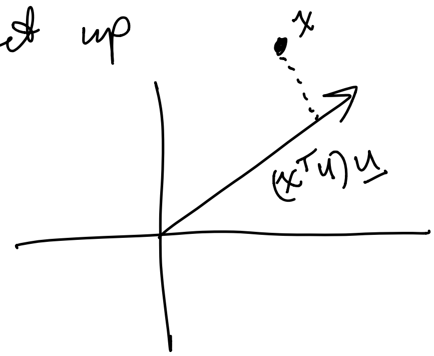
throw away.

$$\begin{aligned} \|x\|^2 &= \langle x, x \rangle \\ &= \left\langle \sum \alpha_i u^{(i)}, \sum \alpha_j u^{(j)} \right\rangle \\ &= \sum \alpha_i^2 \\ &= \sum_{i=1}^{d} \alpha_i^2 + \sum_{d+1}^{n} \alpha_i^2 \end{aligned}$$

$$\|x\|^2 = \|\hat{x}\|^2 + \|\hat{x}_\perp\|^2$$

$\Rightarrow \|\hat{x}\|^2 \leq \|x\|^2$ so vector shrinks

1D set up



$$\hat{x} = (x^T u)\, u$$

$$\max_u \sum_m \|\hat{x}^{(m)}\|^2$$

$$\|\hat{x}\|^2 = \hat{x}^T \hat{x}$$

$$= (x^T u)^2\, u^T u$$

$$= (x^T u)^2$$

$$= u^T x\, x^T u$$

$$\max_u \sum_m u^T x^{(m)} x^{(m)T} u$$