Relationship between PCA & Autoencoder..

I/p $X_{n\times m} = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ x_{21} & & \vdots \\ \vdots & & \vdots \\ x_{n1} & & x_{nm} \end{bmatrix}$

n - Sample
m - feature dimension.

PCA

$k < m$.

$$X_{n\times m} \xrightarrow{V_{m\times k}} X_{n\times k}$$

$$\boxed{X^{new}_{n\times k} = X_{n\times m} V_{m\times k}}$$

$$Cov(X) = \frac{1}{n}\left[(X-\mu_x)'_{m\times n}(X-\mu_x)_{n\times m}\right]$$

$\Sigma_{m\times m} \cdot V_{m\times m} = V_{m\times m} \cdot \lambda_{m\times m}$.

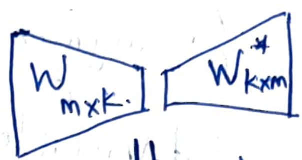
Eigen Vectors (Each Column is an eigen Vector)

Eigen Values.

→ Take eigen Vectors corresponds to top k eigen values.

$V_{m\times k}$ → Transformation matrix.

Auto Encoder



$X_{n\times m}$ $\hat{X}_{n\times m}$

$H_{n\times k}$

$\cancel{H_{n\times m} = f(X_{n\times m} W_{m\times k} + b_{1\times k})}$

$$H_{n\times k} = f(X_{n\times m} W_{m\times k} + 1_{n\times 1} \cdot b_{1\times k})$$

$$\hat{X}_{n\times m} = g(X_{n\times k} W^*_{k\times m} + 1_{m\times 1} b^*_{1\times m})$$

$$\boxed{\begin{aligned} H_{n\times k} &= f(X_{n\times m} W_{m\times k}) \\ \hat{X}_{n\times m} &= g(H_{n\times k} W^*_{k\times m}) \end{aligned}}$$

Under following Conditions encoder part of an auto encoder is equivalent to PCA.

(i) use a linear encoder
(ii) Use a linear decoder
(iii) Use a squared error loss function
(iv) normalise the inputs

$\cancel{x_{ij} = \frac{1}{\sqrt{m}}\left(x_{ij} - \frac{1}{m}\sum_{k=1}^{m} x_{kj}\right)}$

$\cancel{x_{ij} = x_{ij} - \frac{1}{m}\sum_{k=1}^{m} x_{kj}}$

$$\hat{x}_{ij} = \frac{1}{\sqrt{m}} \underbrace{\left(x_{ij} - \frac{1}{m}\sum_{k=1}^{m} x_{kj}\right)}_{x'_{ij}}$$

$X' \rightarrow$ be zero mean data

$$X = \frac{1}{\sqrt{m}} X'$$

$$X^T X = \frac{1}{m}(X')^T X' \rightarrow \text{Covariance matrix}$$

$\Rightarrow$ First we will show that if we use a linear encoder & Squared error loss function then.

The optimal solution to the following objective function

$$\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \hat{x}_{ij})^2 \Rightarrow \min_{\theta} \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \hat{x}_{ij})^2$$

is obtained when we use a linear encoder is equivalent to

$$\min_{W^*, W} \left(\|X - HW^*\|_F\right)^2 \qquad \boxed{H = XW}$$

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} \rightarrow \text{frobenius norm}$$

From SVD, the optimal solution to the above problem is given by

$$\underset{n\times m}{HW^*} = U_{n\times k}\,\Sigma_{k\times k}\,V^T_{k\times m} \quad [V_{m\times k}]$$

$$H = U_{n\times k}\,\Sigma_{k\times k}$$

$$W^* = V^T_{k\times m}$$

objective is to show $\boxed{H = XV}$

SVD

$$X_{n\times m} = U_{n\times m}\,\Sigma_{n\times m}\,V^T_{m\times m} \quad [V_{m\times m}]$$

$$X_{recons} = U_{n\times k}\,\Sigma_{k\times k}\,V^T_{k\times m} \quad [V_{m\times k}]$$

$$\boxed{X_{redu} = U_{n\times k}\,\Sigma_{k\times k}}$$

Each column is an eigen vector

$U \rightarrow$ Eigen Vectors of XX^T

$\Sigma \rightarrow$ Square root of non-zero eigenvalues of XX^T

$V \rightarrow$ Eigen vectors of X^TX

$$= (XX^T)(XX^T)^{-1} U_{n\times k} \Sigma_{k\times k} \quad \left[\text{Pre multiplying by } (XX^T)(XX^T)^{-1} = I\right]$$

Substitute $X = U\Sigma V^T$, keeping first X rest ~~[illegible]~~ substituted.

$$= (XV\Sigma^T U^T)(U\Sigma \underline{V^T V} \Sigma^T U^T)^{-1} U\Sigma \quad [V^T V = I]$$

Substitute $V^T V = I$

$$= XV\Sigma^T U^T (U\Sigma\Sigma^T U^T)^{-1} U\Sigma$$

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1} \quad (U\Sigma\Sigma^T U^T)^{-1} = (U^T)^{-1}(\Sigma\Sigma^T)^{-1}U^{-1}$$

$$= U(\Sigma\Sigma^T)^{-1}U^{-1} \quad U^{-1} = U^T, \quad (U^T)^{-1} = U$$

$$= XV\Sigma^T \underline{U^T U} (\Sigma\Sigma^T)^{-1} U^T U\Sigma$$

$$= XV\Sigma^T (\Sigma\Sigma^T)^{-1} U^T U \Sigma$$

$$= XV \underbrace{\Sigma^T (\Sigma^T)^{-1}}_{I} \Sigma^{-1} \underbrace{U^T_{n\times n} U \Sigma_{n\times k}}_{I_{n\times k}}$$

$$= XV\Sigma^{-1}_{m\times n} I_{n\times k} \Sigma_{k\times k}$$

$$= XV\Sigma^{-1}_{m\times k}\Sigma_{k\times k}$$

$$= XV I_{m\times k}$$

$$\boxed{H = XV_{m\times k}}$$

$U \to 2\times 3$

$U^{-1} = 3\times 2$

$U^{-1}_{3\times 3} U_{3\times 2}$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\left(\begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{pmatrix}\right)$$

remove last column