* Dimension Reduction :

high-dim $\longrightarrow$ low-dim

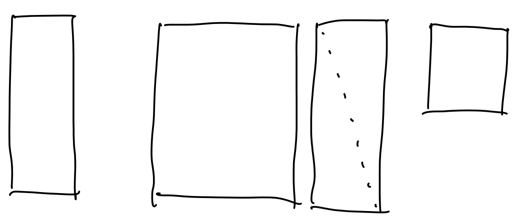
$R^d \longrightarrow R^k$, $d \gg k$

We want to preserve the structure in the data : e.g. distance pairwise

* SVD , or Singular Value Decomposition

$$\underset{m\times n}{A} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V}^{T}$$

orthogonal orthogonal



Eigenvalue and Eigenvector

$$\underset{n\times n}{A}\ \underset{n\times 1}{x} = \lambda \underset{n\times 1}{x}$$

"only for square matrix"

$A = V \operatorname{diag}(\lambda) V^{-1}$

eigenvectors eigen values

rank?

* SVD and rank-k approximation.

$$\tilde{A} = \underset{m\times k}{U}\ \underset{k\times k}{\Sigma}\ \underset{k\times n}{V}^{T}$$

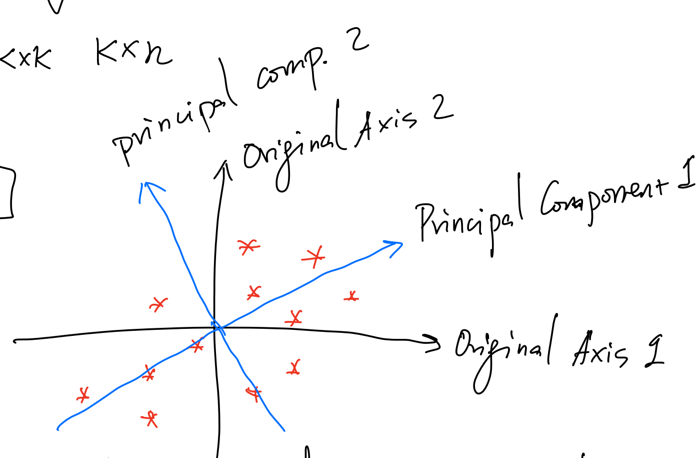
* PCA : Intuition

Example :

Save $\begin{bmatrix}1\\2\\3\end{bmatrix}$ $\begin{bmatrix}2\\4\\6\end{bmatrix}$ $\begin{bmatrix}4\\8\\12\end{bmatrix}$

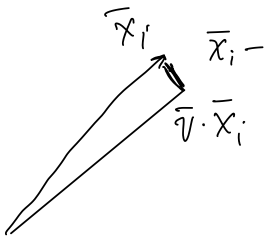
or $[1, 2, 4]$ w/ $\begin{bmatrix}1\\2\\3\end{bmatrix}$ as basis.

1/3 cost.



☑ Project high dim to low-dim $\underline{\text{LINEAR}}$ subspace

$\boxed{\checkmark}$ Correlated ft $\longrightarrow$ Uncorrelated ft



Let $\bar{v}_1, \bar{v}_2 \cdots \bar{v}_d$ be the d principal components

$$\bar{v}_i \cdot \bar{v}_j = 0 \ , \quad \bar{v}_i \cdot \bar{v}_i = 1$$

Assume data is centered,

Let $X = m\left\{\begin{bmatrix} | & | & | & | & | & | \\ | & | & | & | & | & | \end{bmatrix}_{m \times n}\right.$, $x_1 \cdots x_n \in R^m$

$x_1 \ x_2 \cdots x_n$

Objective : Find vector that maximizes sample variance of the projected data.

$$\sum_{i=1}^{n} (v^T x_i)^2 = v^T X X^T v$$

$$\Rightarrow \quad \max_{v} \ v^T X X^T v$$

$$\text{s.t.} \quad v^T v = 1$$

Lagrangian : $\displaystyle \max_{v} \ v^T X X^T v - \lambda v^T v$

$$\frac{d}{dv} = 0 \qquad \cancel{v^T} X X^T v - \lambda \cancel{v^T} v = 0$$

$$(XX^T - \lambda I) v = 0$$

$$(XX^T) v = \lambda v$$

$$\Updownarrow$$

$$A x = \lambda x$$

So v are eigenvectors of XX^T

Steps: ① Data matrix A : each row is a sample/data point

② center data by subtracting mean $A - \bar{A}$

③ Compute SVD of A : $\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V^T}$

④ Sort values in Σ,

top $\underset{\text{rows}}{k}$ in V^T are principal components

⑤ project data to those axis $\quad \underset{k\times n}{[V^T]_k} \cdot \underset{m\times n}{A}$

$$\underset{m\times n}{A} \cdot \underset{n\times k}{[V]_k} = \underset{m\times k}{\tilde{A}}$$

* Multidimensional Scaling or MDS

Setup : Instead of giving the raw data matrix,

we are given their similarity measure.

Can we find those data in k-dimensional space?

$\hookrightarrow$ get coordinates.

Let's say we have $n\times n$ matrix D. s.t. $d_{ij} = (x_i - x_j)^2$

$$D = \begin{array}{c|ccc} & x_1 & x_2 & x_3 \;\cdots \\ \hline x_1 & 0 & (x_1-x_2)^2 & (x_1-x_3)^2 \\ x_2 & (x_1-x_2)^2 & 0 & (x_2-x_3)^2 \\ x_3 & (x_1-x_3)^2 & (x_2-x_3)^2 & 0 \\ \vdots & & & \ddots \end{array}$$

Find $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix}$

$\because \; d_{ij}^2 = (x_i - x_j)^2 = x_i^2 - 2x_i x_j + x_j^2$

INPUT: n objects pairwise distances in p-dimension

OUTPUT: positions of those points up to $\left\{ \begin{array}{l} \text{rotation} \\ \text{translation} \end{array} \right.$

① Compute Gram matrix B.

② Compute positions given B.

$$B = X^T X \quad , \quad X = m\left\{ \overbrace{\begin{bmatrix} | & | & | & | & | & | \end{bmatrix}}^{n} \right._{m \times n}$$

distance: $d_{ij}^2 = b_{ii} + b_{jj} - 2b_{ij}$

$$b_{ij} = -\frac{1}{2}\left(d_{ij}^2 - \sum_j d_{ij}^2 - \sum_i d_{ij}^2 + \sum_{i,j} d_{ij}^2\right)$$, so we got B.

We need to find X.

$$B = X^T X = V \Sigma V^T$$, Because B is inner product symmetric, positive definite

$$X = V_{n \times n} \Sigma_{n \times n}^{-\frac{1}{2}}$$, similar to PCA, you choose how many eigenvectors to keep.

$$\approx V^*_{2 \times n} \Sigma_{n \times n}^{-\frac{1}{2}}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

If objects are Euclidean vectors, and Euclidean distances, $\|x_i - x_j\|_2 = d_{ij}$

MDS = PCA

* Valid Distance & Metric Space

$A > B > C > A \Rightarrow$ non-metric

$A > B, \; B > C, \; A > C \Rightarrow$ metric

MDS doesn't require distances to be metric.