$\boxed{\text{PCA vs MDS}}$

Let $\underline{P} \in \mathbb{R}^{n\times d}$ and assume $\underline{P}$ is already centered. (each column sums up to 0)

$\boxed{\text{PCA}}$

$$\underset{E\in Gr(d,k)}{\operatorname{argmin}} \ \frac{1}{n}\sum_{i=1}^{n} \|p_i - \underbrace{\Pi_E(p_i)}_{\text{orthogonal projection}}\|^2$$

Covariance: $C = \frac{1}{n}\underline{P}^T\underline{P} \in \mathbb{R}^{d\times d}$

Diagonalization: $C = Q^T D Q$

$\lambda_1 \geq \dots \geq \lambda_r > 0$ $\quad D = \left[\begin{array}{ccc|c} \lambda_1 & & & \\ & \ddots & & 0 \\ & & \lambda_r & \\ \hline & 0 & & 0 \end{array}\right]$

$Q = \begin{bmatrix} e_1^T \\ \vdots \\ e_d^T \end{bmatrix} \to$ take $\boxed{X = PQ^T}$

(for target $k<d$, take $P\cdot[\,e_1 \dots e_k \,|\, 0\,]$)

SVD: $\underline{P} = U^T D' V$ $\quad D' = \left[\begin{array}{ccc|c} \mu_1 & & & \\ & \ddots & & 0 \\ & & \mu_s & \\ \hline & 0 & & 0 \end{array}\right]$

$\Rightarrow \frac{1}{n}\underline{P}^T\underline{P} = \frac{1}{n} V^T D'^T D' V$

uniqueness of eigenvalues $\Rightarrow \begin{cases} D = D'^T D'/n \\ r = s \\ \forall i,\ \mu_i^2 = n\lambda_i \end{cases}$

↳ take $\boxed{X' = \underline{P}V^T}$

Then: $X'\underbrace{(VQ^T)}_{\text{orthogonal transform}} = \underline{P}V^TVQ^T = \underline{P}Q^T = X$

$\boxed{\text{MDS}}$

$$\underset{Y\in\mathbb{R}^{n\times k}}{\operatorname{argmin}} \ \|YY^T - \underbrace{\underline{P}\,\underline{P}^T}_{\text{(inner product/ Gram matrix)}}\|_F^2$$

(Frobenius norm)

Gram: $G = \underline{P}\,\underline{P}^T \in \mathbb{R}^{n\times n}$

Diagonalization: $G = R^T F R$ $\quad F = \left[\begin{array}{ccc|c} \delta_1 & & & \\ & \ddots & & 0 \\ & & \delta_r & \\ \hline & 0 & & 0 \end{array}\right]$

↳ take $Y = R^T\sqrt{F}$

$(\Rightarrow YY^T = R^TFR = G)$.

SVD: $\underline{P} = U^T D' V$ $\quad D' = \left[\begin{array}{ccc|c} \mu_1 & & & \\ & \ddots & & 0 \\ & & \mu_s & \\ \hline & 0 & & 0 \end{array}\right]$

↳ take $\boxed{Y' = U^T D'}$

$(\Rightarrow Y'Y'^T = U^TD'D'^TU = \underline{P}\,\underline{P}^T)$

uniqueness of eigenvalues

$\Rightarrow \begin{cases} F = D'D'^T \\ r = s \\ \mu_i^2 = \delta_i \ \forall i \end{cases}$

$\Rightarrow \underbrace{(R^TU)}_{\text{orthogonal transform}} Y' = R^T\sqrt{F}\,\underbrace{I_{n,d}}_{\left[\begin{smallmatrix}1 & & \\ & \ddots & \\ & & 1\end{smallmatrix}\right] \in \mathbb{R}^{n\times d}} = Y I_{n,d}$

(note: $r = s \leq \min\{n,d\}$)

$\underline{P} = U^T D' V \Rightarrow \underline{P}V^T = U^T D'$

$\Rightarrow \boxed{X' = Y'}$ (PCA and MDS are dual to each other)

metric MDS:

Input: a squared distance matrix $\Delta \in \mathbb{R}^{n \times n}$

$\forall i, j,\ \Delta_{ij} := d(p_i, p_j)^2$.

Case 1 $d(p_i, p_j) = \|p_i - p_j\|_2$ for some point cloud P in Euclidean space.

↳ let $G := -\frac{1}{2} H . \Delta . H$, where $H = I_n - \frac{1}{n}[1]$

(matrix with just 1's)

Prop: G is then the inner product matrix:

$\forall i, j,\ G_{ij} = \langle p_i | p_j \rangle$.

↳ apply MDS on G and get embedding X with $XX^T = G$.

Case 2 $d(p_i, p_j)$ is not a Euclidean distance,

↳ $G := -\frac{1}{2} H . \Delta . H$, where $H = I_n - \frac{1}{n}[1]$.

Then: G is <u>not</u> an inner product matrix

But: G remains symmetric $\Rightarrow$ diagonalizable (with some negative eigenvalues)

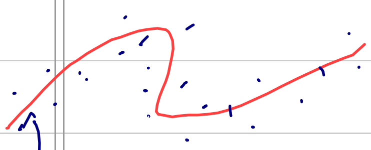
↳ apply MDS on G and get embedding X that minimizes $\|XX^T - G\|_F^2$

(ie. best preserves the Gram matrix).

Advantage: metric MDS can be applied in any metric space.

$\boxed{\text{Isomap}}$

Principle: apply metric MDS to the matrix of distances along the (curvy) object



take distances along this object.

$\boxed{\text{Thm}}$ [de Silva, Langford, Tenenbaum 2000]

If the object S is defined as $S = \varphi(\Omega)$, where:

- Ω is a convex set in $\mathbb{R}^k$
- $\varphi : \Omega \to \mathbb{R}^d$ preserves distances (isometry)

Then: metric MDS applied to the matrix of pairwise distances along S gives an embedding $X \in \mathbb{R}^{n \times k}$ that preserves these distances exactly.

In practice:

1. approximate the pairwise distances along S by:
   - computing some neighborhood graph (e.g. connect every data point to every other point within Euclidean distance ε, for some fixed $\varepsilon > 0$).
   - computing distances in the graph
2. apply metric MDS to the resulting distance matrix.