Unified framework

- Kernel PCA $\quad X \in R^{d \times n} \longrightarrow Y \in R^{p \times n}$

$K_{n \times n}$ : similarity between points

$Y = \Sigma V^T$

$\boxed{\underset{\sim}{y} = \Sigma^{-1} V^T K(X, \underset{\sim}{x})}$

$\underset{\sim}{x}$: out of sample (test) data
$\underset{\sim}{y}$: projected test data

- Dual PCA

$K = X^T X$ which is a linear kernel

- MDS

$K = -\frac{1}{2} H D_{(E)}^2 H$ : $D_{(E)}$ is a Euclidean distance

- Isomap

$K = -\frac{1}{2} H D_{(G)}^2 H$ : $D_{(G)}$ is a geodesic distance

- LLE

$M = (I - W)(I - W)^T$

Select 2 to p+1 smallest eigenvectors

Can you show me a matrix whose largest eigenvalues are small as the smallest eigenvalues of M. In othe words, we want to look for a matrix that has same eigenvectors as M, but the order of eigenvalues is fliped (reversed)

$K = M^{-1}$

$M = U \Sigma U^T$ $\quad V^T = U^T$ because $M$ is P.S.D

$M^{-1} = U \Sigma^{-1} U^T$ $\quad M M^{-1} = \underbrace{U \Sigma \underbrace{U^T U}_{I} \Sigma^{-1}}_{I} U^T = I$

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_p \end{pmatrix} \qquad \Sigma^{-1} = \begin{pmatrix} 1/\sigma_1 & & & \\ & 1/\sigma_2 & & \\ & & \ddots & \\ & & & 1/\sigma_p \end{pmatrix}$$

$K = \sigma_{max} I - M$ also satisfy the condition.

So in LLE, $K = M^{-1}$ or $\sigma_{max} I - M$

$Y$ consists of eigenvectors corresponding to largest eigenvalues of $K$

Note 1: LLE doesn't capture the variation in each direction so LLE is kernel PCA up to a scale.

Note 2: In kernel PCA, we can map out of sample data. Can we do it in Isomap or LLE?
No, we cannot. In kernel PCA, we use a closed-form kernel, but in Isomap and LLE, we don't have the closed-form kernel. For each new point, we need to recompute the kernel in Isomap and LLE (reconstruct the graph).

- Semi-definite embedding (SDE)

In SDE, K is driven by data and solved by semi-definite programming.