Kernal PCA

$\underline{x}_1 \cdots x_n \in \mathbb{R}^d$ are observations
(aka samples)
as column vectors

Non-linear lifting: $\underline{z}_k = \phi(\underline{x}_k)$

Matrix $Z = [z_1 \cdots z_n]$

Mean $\bar{z} = \frac{1}{n}(z_1 + \cdots + z_n) = \frac{1}{n} Z \underline{e}$, $\underline{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

Centered data:

$$\bar{Z} = [z_1 - \bar{z}, \cdots, z_n - \bar{z}]$$

$$= Z - [\bar{z}\,\bar{z} \cdots \bar{z}] = Z - (\tfrac{1}{n} Z e) e^T$$

$$= Z \underbrace{\left(I - \tfrac{1}{n} e e^T\right)}_{P \text{ = orthogonal projector}} = ZP$$

$P\underline{e} = 0$; if $\underline{v} \perp \underline{e}$ then $P\underline{v} = \underline{v}$.

Covariance Matrix $Cov = \bar{Z}\bar{Z}^T$ scalar ignored

Kernal Matrix $\bar{K} = \bar{Z}^T \bar{Z}$

Take SVD of $\bar{Z} = ZP$

$$\begin{array}{c} \xleftrightarrow{\text{samples}} \\ \text{features} \updownarrow \; \bar{Z} = U \Sigma V^T = \end{array}$$

$$\begin{array}{c} \text{features} \\ \updownarrow \end{array} = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_k \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} \begin{array}{c} \xleftrightarrow{\text{samples}} \\ \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_k^T & - \end{bmatrix} \end{array}$$

$k = \operatorname{rank} K$

Properties

$$\bar{K} = V \Sigma^2 V^T$$

$$Cov = U \Sigma^2 U^T$$

$$\bar{Z} \underline{v}_\gamma = \sigma_\gamma \underline{u}_\gamma \qquad \bar{Z}^T \underline{u}_\gamma = \sigma_\gamma \underline{v}_\gamma$$

Projection of samples onto u_γ's
(principal directions)

$$u_\gamma^T \bar{z}_i = u_\gamma^T \bar{Z} e_i = \sigma_\gamma v_\gamma^T e_i$$

$$= \sigma_\gamma [V^T]_{\gamma i}$$

Aside projection of $\bar{z}$

2 Kernal Matrices

$$K = Z^T Z$$

$$\bar{K} = \bar{Z}^T \bar{Z} = P Z^T Z P$$

where $P = I - \frac{1}{n} e e^T$

$$\bar{K} e = P Z^T Z \underbrace{P e}_{0} = 0$$

so $e \perp v_1 \cdots v_K$, e is a trailing sing. vector

so $P v_j = v_j$ and $Z v_j = \bar{Z} v_j$

Projection of $\bar{z}$ onto u_j:

$$\bar{z}^T u_j = \frac{1}{\sigma_j} \underbrace{\frac{1}{n} e^T Z^T}_{\bar{z}^T} \underbrace{Z P v_j}_{u_j} = \frac{1}{\sigma_j} \frac{1}{n} e^T K v_j$$

Projection of a new sample

$$t = \phi(x_{new})$$

$\bar{t} = t - \bar{z}$ - account for center shift

$$t^T u_j = t^T Z v_j \frac{1}{\sigma_j}$$

$$= \frac{1}{\sigma_j} \left[K(x_{new}, x_1) \cdots K(x_{new}, x_n) \right] v_j$$

$$\bar{z}^T u_j = \frac{1}{\sigma_j} \frac{1}{n} e^T K v_j$$

if $x_{new} = x_i$ (one of original samples) then

$\bar{t} = \bar{z}_i$ and $\bar{z}_i^T u_j = \dfrac{e_i^T v_j}{\sigma_j}$