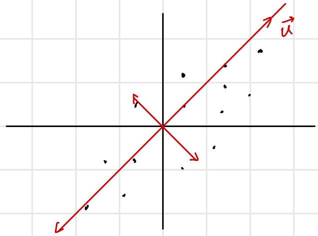
PCA

$$\max_{\|u\|_2 \le 1} \frac{1}{m} \sum_{i=1}^{m} \left(u^T \left(x_i - \frac{1}{m} \sum_{k=1}^{m} x_k\right) \right)^2$$



Kernel PCA: Approach I (makes more sense to me in optimization viewpoint)

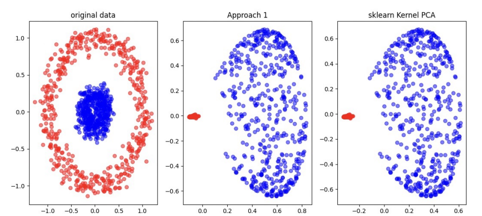
$$\max_{\|f\|_{\mathcal{H}} \le 1} \frac{1}{m} \sum_{i=1}^{m} \left\langle f, \phi(x_i) - \frac{1}{m} \sum_{k=1}^{m} \phi(x_k) \right\rangle_{\mathcal{H}}^2$$

$$\Rightarrow \max_{\|f\|_{\mathcal{H}} \le 1} \frac{1}{m} \sum_{i=1}^{m} \left\langle \sum_{j=1}^{m} \alpha_j \phi(x_j), \phi(x_i) - \frac{1}{m} \sum_{k=1}^{m} \phi(x_k) \right\rangle_{\mathcal{H}}^2$$

$$\Rightarrow \max_{\|f\|_{\mathcal{H}} \le 1} \frac{1}{m} \sum_{i=1}^{m} \left[\sum_{j=1}^{m} \alpha_j \left(k(x_j, x_i) - \frac{1}{m} \sum_{k=1}^{m} \phi(x_k, x_j) \right) \right]^2$$

$$\Rightarrow \max_{\|f\|_{\mathcal{H}} \le 1} \frac{1}{m} \sum_{i=1}^{m} \left[\alpha^T \tilde{K}_i \right] \qquad \tilde{K}_i[j] = k(x_i, x_j) - \frac{1}{m} \sum_{k=1}^{m} k(x_k, x_j)$$

$$\Rightarrow \max_{\|f\|_{\mathcal{H}} \le 1} \alpha^T \underbrace{\left(\frac{1}{m} \sum_{i=1}^{m} \tilde{K}_i \tilde{K}_i^T \right)}_{\tilde{K}} \alpha$$



$$\|f\|_{\mathcal{H}}^2 = \sum_i \sum_j \alpha_i \alpha_j k(x_i, x_j) = \alpha^T K \alpha$$

$$\Downarrow$$

$$\max_{\alpha} \alpha^T \tilde{K} \alpha \quad \text{s.t.} \quad \alpha^T K \alpha \le 1 \quad \underset{\alpha = K^{-1/2}\beta}{\overset{\beta = K^{1/2}\alpha}{\rightleftarrows}} \quad \max_{\beta} \beta^T K^{-1/2} \tilde{K} K^{-1/2} \beta \quad \text{s.t.} \quad \|\beta\|_2 \le 1$$

$\beta^* =$ eigenvector with maximum eigenvalue of $K^{-1/2} K K^{-1/2}$

$\alpha^* = K^{-1/2} \beta^*$

remark. if β_1, β_2 are orthogonal eigenvectors of $K^{-1/2} \tilde{K} K^{-1/2}$

then $\alpha_i = K^{-1/2} \beta_i$ are orthogonal eigenvectors of $\tilde{K}$

then the subspace $\sum_i \alpha_1[i] \tilde{\phi}(x_i, \cdot)$ and $\sum_i \alpha_2[i] \tilde{\phi}(x_i, \cdot)$ are orthogonal in $\mathcal{H}$

$$\left\langle \sum_i \alpha_1[i] \tilde{\phi}(x_i, \cdot), \sum_i \alpha_2[i] \tilde{\phi}(x_i) \right\rangle_{\mathcal{H}}$$

$$= \alpha_1^T \tilde{K} \alpha_2$$

$$= 0 \qquad \tilde{\phi}(x_i) = \phi(x_i) - \frac{1}{m} \sum_{k=1}^{m} \phi(x_k)$$

projection: $\langle \phi(z), \sum_i \alpha_i \phi(x_i) \rangle = \sum_i \alpha_i k(x_i, z)$

• **Kernel PCA : Approach II** (most online resource follows this approach)

$$\tilde{\phi}(x_i) = \phi(x_i) - \frac{1}{m}\sum_{j=1}^{m}\phi(x_j)$$

$$\begin{aligned}
\tilde{K}(x_i, x_j) &= \langle \tilde{\phi}(x_i), \tilde{\phi}(x_j) \rangle \\
&= \left\langle \phi(x_i) - \frac{1}{m}\sum_{k=1}^{m}\phi(x_k),\ \phi(x_j) - \frac{1}{m}\sum_{k=1}^{m}\phi(x_k) \right\rangle \\
&= \left. K(x_i, x_j) - \frac{1}{m}\sum_{k=1}^{m}K(x_k, x_j) - \frac{1}{m}\sum_{k=1}^{m}K(x_k, x_i) + \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m}K(x_i, x_j) \right\}
\end{aligned}$$

is this a valid kernel?

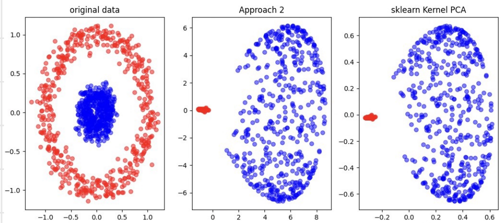
$$\max_{\|f\| \le 1} \frac{1}{m} \left\langle \sum_{j=1}^{m}\alpha_j \tilde{\phi}(x_j), \tilde{\phi}(x_i) \right\rangle^2$$

$$\Rightarrow \max_{\|f\| \le 1} \frac{1}{m} \sum_{i=1}^{m} \left(\sum_{j=1}^{m} \alpha_j \tilde{K}(x_i, x_j) \right)^2$$

$$\Rightarrow \max_{\|f\| \le 1} \frac{1}{m} \sum_{i=1}^{m} \left(\alpha^T \tilde{K}_i \right) \qquad \tilde{K}_i[j] = \tilde{K}(x_i, x_j)$$

$$\Rightarrow \max_{\|f\| \le 1} \alpha^T \left(\frac{1}{m} \sum_{i=1}^{m} \tilde{K}_i \tilde{K}_i^T \right) \alpha$$

$$\Rightarrow \max_{\|f\| \le 1} \alpha^T \tilde{K}\tilde{K} \alpha \qquad \tilde{K}_{ij} = \tilde{K}(x_i, x_j)$$



different scale than sklearn

$$\|f\|^2 = \sum_{ij} \alpha_i \alpha_j \tilde{K}(x_i, x_j) = \alpha^T \tilde{K} \alpha$$

$$\Downarrow$$

$$\left. \begin{aligned} \max_{\alpha} \quad & \alpha^T \tilde{K}\tilde{K} \alpha \\ \text{s.t.} \quad & \alpha^T \tilde{K} \alpha \le 1 \end{aligned} \right\}$$

α^* is the eigenvector with maximum eigenvalue

projection: $\langle \tilde{\phi}(z), \sum_i \alpha_i \tilde{\phi}(x_i) \rangle = \sum_i \alpha_i \tilde{K}(x_i, z)$

can we do this for $x \neq z$?

Anyway, I give more trust to Approach 1, even if it needs 2 eigenvalue decompositions :)