Kernel Based Principal Component Analysis

Yan PEI/裴岩 /ペイ　イエン

peiyan@u-aizu.ac.jp

Kernel-Based Principal Component Analysis

- Method: projecting samples into feature space

$$C = \frac{1}{N}\sum_{i=1}^{N}\varphi(x_i)\varphi(x_i)^T = \frac{1}{N}[\varphi(x_1),\ldots,\varphi(x_n)]\begin{bmatrix}\varphi(x_1)^T \\ \vdots \\ \varphi(x_n)^T\end{bmatrix}$$

$$C = \frac{1}{N}X^T X$$

- But, we do not know feature map, because it is implicit
- How to find eigenvectors of $X^T X$?

Kernel-Based Principal Component Analysis

- What is the kernel matrix?

$$K = \boxed{XX^T} = \begin{bmatrix} \varphi(x_1)^T \\ \vdots \\ \varphi(x_n)^T \end{bmatrix} [\varphi(x_1),\ldots,\varphi(x_n)]$$

- How we use K to calculate C?

$$C = \frac{1}{N}\sum_{i=1}^{N}\varphi(x_i)(x_i^T) = \frac{1}{N}[\varphi(x_1,\ldots,\varphi(x_n)]\begin{bmatrix} \varphi(x_1^T) \\ \vdots \\ \varphi(x_n^T) \end{bmatrix} = \frac{1}{N}\boxed{X^T X}$$

Kernel-Based Principal Component Analysis

- Let us calculate it
  - What is the eigenvalue of $K = XX^T = \begin{bmatrix} \varphi(x_1)^T \\ \vdots \\ \varphi(x_n)^T \end{bmatrix} [\varphi(x_1), \ldots, \varphi(x_n)]$

$$\boxed{\begin{gathered} (XX^T)u = \lambda u \\ X^T(XX^T)u = X^T \lambda u \\ (X^T X)X^T u = \lambda(X^T u) \end{gathered}}$$

$$C = \frac{1}{N} X^T X = [\varphi(x_1), \ldots, \varphi(x_n)] \begin{bmatrix} \varphi(x_1)^T \\ \vdots \\ \varphi(x_n)^T \end{bmatrix}$$

- Which are eigenvalue and eigenvector of C?
- However, eigenvector is not sure equal 1, how to normalize?

derivation

Kernel-Based Principal Component Analysis

- Eigen vector of C can be calculated by K

$$K = XX^T = \begin{bmatrix} \varphi(x_1)^T \\ \vdots \\ \varphi(x_n)^T \end{bmatrix} [\varphi(x_1), \ldots, \varphi(x_n)]$$

$$C = \frac{1}{N} X^T X = [\varphi(x_1), \ldots, \varphi(x_n)] \begin{bmatrix} \varphi(x_1)^T \\ \vdots \\ \varphi(x_n)^T \end{bmatrix}$$

$$\boxed{v = \frac{1}{|X^T u|} X^T u = \ldots = \frac{1}{\sqrt{\lambda}} X^T u}$$

Kernel-Based Principal Component Analysis

$$K = XX^T = \begin{bmatrix} \varphi(x_1)^T \\ \vdots \\ \varphi(x_n)^T \end{bmatrix} [\varphi(x_1), \ldots, \varphi(x_n)]$$

$$C = \frac{1}{N} X^T X = [\varphi(x_1), \ldots, \varphi(x_n)] \begin{bmatrix} \varphi(x_1)^T \\ \vdots \\ \varphi(x_n)^T \end{bmatrix}$$

$$\boxed{(XX^T)u = \lambda u}$$

$$X^T(XX^T)u = X^T \lambda u$$

$$\boxed{\begin{gathered} (X^T X) X^T u = \lambda (X^T u) \\ v = \frac{1}{\sqrt{\lambda}} X^T u \end{gathered}}$$

However, we do not know X, so we do not know v

derivation

Kernel-Based Principal Component Analysis

- However, we do not know X, so we do not know v!!!

$$v = \frac{1}{\sqrt{\lambda}} X^T u$$

- Question: how to know the projection of a sample coordination in the feature space?

$$v^T \varphi(x') = (\frac{1}{\sqrt{\lambda}} X^T u)^T \varphi(x') = \ldots = \frac{1}{\sqrt{\lambda}} u^T \begin{bmatrix} \kappa(x_1, x') \\ \vdots \\ \kappa(x_N, x') \end{bmatrix}$$

Kernel-Based Principal Component Analysis

- Solve the eigenvalue problem:

$$\kappa u_i = \lambda_i u_i, \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$$

- the projection of a sample coordination in the feature space?

$$v_i^T \varphi(x') = \frac{1}{\sqrt{\lambda_i}} u_i^T \begin{bmatrix} \kappa(x_1, x') \\ \vdots \\ \kappa(x_N, x') \end{bmatrix}$$

derivation

Kernel-Based Principal Component Analysis

- Relationship with Wahba theorem

$$v = \frac{1}{\sqrt{\lambda}} X^T u = \frac{1}{\sqrt{\lambda}} [\varphi(x_1), \ldots, \varphi(x_N)] \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix}$$

$$w = \sum_{i=1}^{N} \alpha_i \varphi(x_i)$$

Kernel-Based Principal Component Analysis

- Visual Example: PCA and KPCA