KPCA(Kernel based Principle Component Analysis)方法的推导

- 首先证明该样本集合协方差矩阵的特征向量处于样本所张成的空间，即：$v \in span\{x_1, \cdots, x_m\}$，设有样本集合：$X = \{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$，假设样本的均值为 0，即：$\sum_{i=1}^{m} \mathbf{x}_i = 0$，。

[证明]

样本集合的协方差矩阵为：$C = \frac{1}{m}\sum_{j=1}^{m} x_j \cdot x_j^t$，令 λ 为 C 的特征值，$\mathbf{v}$ 为对应特征矢量，则有：

$$\lambda \mathbf{v} = C\mathbf{v} = \left(\frac{1}{m}\sum_{j=1}^{m} \mathbf{x}_j \cdot \mathbf{x}_j^t\right)\mathbf{v} = \frac{1}{m}\sum_{j=1}^{m}\left(\mathbf{x}_j^t \mathbf{v}\right)\mathbf{x}_j \tag{1}$$

即：

$$\mathbf{v} = \frac{1}{m\lambda}\sum_{j=1}^{m}\left(\mathbf{x}_j^t \mathbf{v}\right)\mathbf{x}_j$$

因此：$v \in span\{x_1, \cdots, x_m\}$。 证毕

- 推导特征空间中协方差矩阵特征值和特征向量的求解方法。设有样本集合：$X = \{\mathbf{x}_1, \cdots, \mathbf{x}_m\}$，取非线性映射：$\phi: R^d \to \mathcal{F}$，$R^d$ 为样本所处的 d 维欧氏空间，称为输入空间，$\mathcal{F}$ 为一个 Hilbert 空间，称为特征空间，样本在特征空间中的内积可以用一个核函数来计算：$k(\mathbf{x}, \mathbf{y}) = (\phi(\mathbf{x}))^t \phi(\mathbf{y})$。假设样本集合在特征空间中的均值为 0，即：$\sum_{i=1}^{m} \phi(\mathbf{x}_i) = 0$。

  样本在特征空间中的协方差矩阵为：

$$\overline{C} = \frac{1}{m}\sum_{i=1}^{m} \phi(\mathbf{x}_i)\phi^t(\mathbf{x}_i) \tag{3}$$

令：λ 为 $\overline{C}$ 的特征值，$\mathbf{V}$ 为对应的特征矢量，则有：

$$\lambda \mathbf{V} = \overline{C}\mathbf{V} \tag{4}$$

利用之前证明的结果，特征矢量 $\mathbf{V} \in span\{\phi(\mathbf{x}_1), \cdots, \phi(\mathbf{x}_m)\}$，即存在 $\alpha_1, \cdots, \alpha_m$，使

得：

$$\mathbf{V}=\sum_{i=1}^{m}\alpha_i\,\phi\left(\mathbf{x}_i\right) \tag{5}$$

(5)式左乘$\phi^t\left(\mathbf{x}_j\right)$，则有：

$$\lambda\left[\phi^t\left(\mathbf{x}_j\right)\cdot\mathbf{V}\right]=\phi^t\left(\mathbf{x}_j\right)\cdot\overline{C}\mathbf{V},\quad j=1,\cdots,m \tag{6}$$

将(5)式代入(6)式左端：

$$\begin{aligned}\lambda\left[\phi^t\left(\mathbf{x}_j\right)\cdot\mathbf{V}\right]&=\lambda\left(\sum_{i=1}^{m}\alpha_i\phi^t\left(\mathbf{x}_j\right)\phi\left(\mathbf{x}_i\right)\right)\\&=\lambda\left(\sum_{i=1}^{m}\alpha_i k_{ji}\right)\\&=\lambda\left(\mathbf{K\alpha}\right)_j\end{aligned} \tag{7}$$

其中：

$$\mathbf{K}=\left[k\left(\mathbf{x}_i,\mathbf{x}_j\right)\right]_{m\times m}=\left(\phi^t\left(\mathbf{x}_i\right)\phi\left(\mathbf{x}_j\right)\right)_{m\times m}$$

$$\boldsymbol{\alpha}=\left(\alpha_1,\cdots,\alpha_m\right)^t$$

$\left(\mathbf{K\alpha}\right)_j$为矢量$\mathbf{K\alpha}$的第 j 个元素.

将(3)和(5)式代入(6)式得右端：

$$\begin{aligned}\phi^t\left(\mathbf{x}_j\right)\cdot\overline{C}\mathbf{V}&=\phi^t\left(\mathbf{x}_j\right)\cdot\frac{1}{m}\sum_{l=1}^{m}\phi\left(\mathbf{x}_l\right)\phi^t\left(\mathbf{x}_l\right)\cdot\sum_{i=1}^{m}\alpha_i\,\phi\left(\mathbf{x}_i\right)\\&=\frac{1}{m}\sum_{l=1}^{m}\sum_{i=1}^{m}\alpha_i\left(\phi^t\left(\mathbf{x}_j\right)\phi\left(\mathbf{x}_l\right)\cdot\phi^t\left(\mathbf{x}_l\right)\phi\left(\mathbf{x}_i\right)\right)\\&=\frac{1}{m}\sum_{l=1}^{m}\sum_{i=1}^{m}\alpha_i k_{jl}k_{li}\\&=\frac{1}{m}\left(\mathbf{K}^2\boldsymbol{\alpha}\right)_j\end{aligned} \tag{8}$$

(7)式(8)式相等，则有：

$$\lambda\mathbf{K\alpha}=\frac{1}{m}\mathbf{K}^2\boldsymbol{\alpha} \tag{9}$$

因此，求取$\overline{C}$的特征值和特征向量的问题可以转化为如下特征值问题：

$$m\lambda\boldsymbol{\alpha}=\mathbf{K\alpha} \tag{10}$$

用λ代替$m\lambda$，即为：

$$\lambda\boldsymbol{\alpha}=\mathbf{K\alpha} \tag{11}$$

令：$\lambda^1,\cdots,\lambda^m$为矩阵$\mathbf{K}$的特征值，$\boldsymbol{\alpha}^1,\cdots,\boldsymbol{\alpha}^m$为特征矢量，则对应的$\overline{C}$的第$i$个特征值和特征矢量为：

$$\lambda_i = \frac{\lambda^i}{m}, \quad \mathbf{V}_i = \sum_{j=1}^{m} \alpha_j^i \phi\left(\mathbf{x}_j\right) \tag{12}$$

其中α_j^i为$\boldsymbol{\alpha}^i$的第j个元素。

➢ 规范化，将$\mathbf{V}_i$转化为单位矢量。

$$\begin{aligned}
\left(\mathbf{V}_i\right)^t \cdot \mathbf{V}_i &= \sum_{j,l=1}^{m} \alpha_j^i \alpha_l^i \left(\phi^t\left(\mathbf{x}_j\right) \cdot \phi\left(\mathbf{x}_l\right)\right) \\
&= \sum_{j,l=1}^{m} \alpha_j^i \alpha_l^i \cdot k\left(\mathbf{x}_j, \mathbf{x}_l\right) \\
&= \left(\boldsymbol{\alpha}^i\right)^t \mathbf{K} \boldsymbol{\alpha}^i \\
&= \lambda^i \left(\boldsymbol{\alpha}^i\right)^t \boldsymbol{\alpha}^i \\
&= \lambda^i \left\|\boldsymbol{\alpha}^i\right\|^2 \\
&= 1
\end{aligned}$$

因此只须使$\left\|\boldsymbol{\alpha}^i\right\|^2 = \frac{1}{\lambda^i}$，即可使$\mathbf{V}_i$规范化。

➢ 计算样本$\mathbf{x}$在特征空间中第j个轴上的投影，利用(12)式，有：

$$\phi^t\left(\mathbf{x}\right)\mathbf{V}_j = \sum_{i=1}^{m} \alpha_i^j \phi^t\left(\mathbf{x}\right)\phi\left(\mathbf{x}_i\right) = \sum_{i=1}^{m} \alpha_i^j k\left(\mathbf{x}, \mathbf{x}_i\right) \tag{13}$$

➢ KPCA 算法：

1. 计算矩阵；$\mathbf{K} = \left[k\left(\mathbf{x}_i, \mathbf{x}_j\right)\right]_{m \times m}$；
2. 计算$\mathbf{K}$的特征值和特征向量：$\lambda^1, \cdots, \lambda^m$，$\boldsymbol{\alpha}^1, \cdots, \boldsymbol{\alpha}^m$，取最大的前$p$个特征值 $\lambda^1, \lambda^2, \ldots, \lambda^p$；
3. 规范化前p个特征向量，使得$\left\|\boldsymbol{\alpha}^i\right\|^2 = \frac{1}{\lambda^i}$；
4. 利用(13)式可以计算样本$\mathbf{x}$的第j个核主成分。