核PCA原理解析

【参考资料】

[核主成分分析（Kernel-PCA）](#)

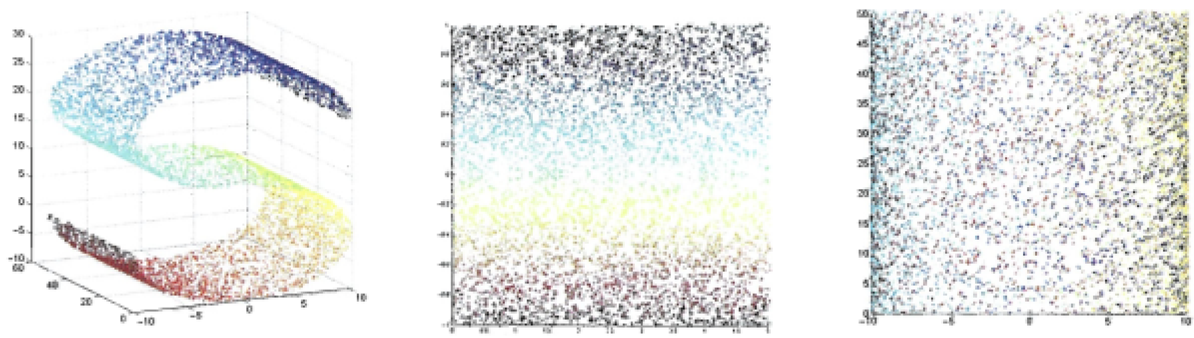
[Kernel PCA](#)

[Kernel Principal Component Analysis](#)

周志华 《机器学习》

1. PCA的缺陷

有些高维空间不适合PCA的线性降维方法，这时候就需要一些非线性降维的手段。



(a) 三维空间中的观察　(b) 本真二维结构　(c) PCA 降维结果

2. KPCA原理推导

Kernel PCA的基本思想是将原数据点非线性映射到高维空间，然后在高维空间中进行降维操作。

已知数据集为$\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M$，其中$\mathbf{x}_i \in R^N$，$i = 1, 2, \cdots, M$。存在一个非线性映射$\Phi : R^N \mapsto R^F$，映射后的数据集为$\Phi\left(\mathbf{x}_1\right), \Phi\left(\mathbf{x}_2\right), \cdots, \Phi\left(\mathbf{x}_n\right)$。PCA是在$R^N$中讨论的，那么KPCA就是在映射后的空间$R^F$中讨论的。一般情况下，映射$\Phi$是无法显式求解的，所以需要引入核函数$K : R^N \times R^N \mapsto R$，使得：

$$K\left(\mathbf{x}_i, \mathbf{x}_j\right) = \Phi(\mathbf{x}_i)^T \Phi\left(\mathbf{x}_j\right)$$

传统PCA中，我们要求解协方差矩阵及其特征值和特征向量，在KPCA中，我们需要求解的则是映射后的数据的协方差矩阵及其特征向量：

$$C = \frac{1}{M} \sum_{i=1}^{M} \left(\Phi\left(\mathbf{x}_i\right) - \frac{1}{M} \sum_{j=1}^{M} \Phi\left(\mathbf{x}_j\right)\right) \left(\Phi\left(\mathbf{x}_i\right) - \frac{1}{M} \sum_{j=1}^{M} \Phi\left(\mathbf{x}_j\right)\right)^T$$

为了讨论方便，令

$$\Psi\left(\mathbf{x}_i\right)=\Phi\left(\mathbf{x}_i\right)-\frac{1}{M}\sum_{j=1}^{M}\Phi\left(\mathbf{x}_j\right)$$

所以C可以表示为

$$C=\frac{1}{M}\sum_{i=1}^{M}\Psi\left(\mathbf{x}_i\right)\Psi\left(\mathbf{x}_i\right)^T$$

由于通常映射Φ是无法显式求解的，所以我们要从另一个角度来讨论C的特征向量的求解。设$\mathbf{v}$是C的特征向量，λ是对应$\mathbf{v}$的特征值，即

$$C\mathbf{v}=\lambda\mathbf{v}$$

$$\frac{1}{M}\sum_{i=1}^{M}\Psi\left(\mathbf{x}_i\right)\Psi\left(\mathbf{x}_i\right)^T\mathbf{v}=\lambda\mathbf{v}$$

显然

$$\mathbf{v}=\frac{1}{M\lambda}\sum_{i=1}^{M}\left(\Psi\left(\mathbf{x}_i\right)^T\mathbf{v}\right)\Psi\left(\mathbf{x}_i\right)$$

令$a_i=\frac{1}{M\lambda}\left(\Psi\left(\mathbf{x}_i\right)^T\mathbf{v}\right)$，$i=1,2,\cdots,M$，使得

$$\mathbf{v}=\sum_{i=1}^{M}a_i\Psi\left(\mathbf{x}_i\right)$$

于是我们便可以知道C的特征向量的性质，即$\mathbf{v}$是由$\Psi\left(\mathbf{x}_1\right),\Psi\left(\mathbf{x}_2\right),\cdots,\Psi\left(\mathbf{x}_M\right)$张成的。

因为

$$C\mathbf{v}=\lambda\mathbf{v}$$

所以

$$\Psi\left(\mathbf{x}_k\right)^T(C\mathbf{v})=\lambda\Psi\left(\mathbf{x}_k\right)^T\mathbf{v},k=1,2,\cdots,M$$

把$\mathbf{v}=\sum_{i=1}^{M}a_i\Psi\left(\mathbf{x}_i\right)$，带入上式得:

右式:

$$\lambda\Psi\left(\mathbf{x}_k\right)^T\mathbf{v}=\lambda\sum_{i=1}^{M}a_i\Psi\left(\mathbf{x}_k\right)^T\Psi\left(\mathbf{x}_i\right)=\lambda\left[\Psi\left(\mathbf{x}_k\right)^T\Psi\left(\mathbf{x}_1\right)\quad\Psi\left(\mathbf{x}_k\right)^T\Psi\left(\mathbf{x}_1\right)\cdots\quad\Psi\left(\mathbf{x}_k\right)^T\Psi\left(\mathbf{x}_M\right)\right]\begin{bmatrix}a_1\\a_2\\\vdots\\a_M\end{bmatrix}$$

左式:

$$\Psi(\mathbf{x}_k)^T(C\mathbf{v}) = \frac{1}{M}\sum_{i=1}^{M}\left(\Psi(\mathbf{x}_k)^T\left(\sum_{j=1}^{M}\Psi\left(\mathbf{x}_j\right)\Psi(\mathbf{x}_j)^T\right)\Psi\left(\mathbf{x}_i\right)\right)$$

$$= \frac{1}{M}\left[\Psi(\mathbf{x}_k)^T\left(\sum_{j=1}^{M}\Psi\left(\mathbf{x}_j\right)\Psi(\mathbf{x}_j)^T\right)\Psi\left(\mathbf{x}_1\right)\cdots \quad \Psi(\mathbf{x}_k)^T\left(\sum_{j=1}^{M}\Psi\left(\mathbf{x}_j\right)\Psi(\mathbf{x}_j)^T\right)\Psi\left(\mathbf{x}_M\right)\right]\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}$$

其中$k = 1, 2, 3, \cdots, M$。

我们用矩阵表示上式等式，即：

$$\lambda\overline{\mathrm{K}}\mathrm{a} = \frac{1}{M}\overline{\mathrm{K}}^2\mathrm{a}$$

其中：

$$\overline{\mathbf{K}} = \begin{bmatrix} \Psi(\mathbf{x}_1)^T\Psi\left(\mathbf{x}_1\right) & \Psi(\mathbf{x}_1)^T\Psi\left(\mathbf{x}_2\right) & \cdots & \Psi(\mathbf{x}_1)^T\Psi\left(\mathbf{x}_M\right) \\ \Psi(\mathbf{x}_2)^T\Psi\left(\mathbf{x}_1\right) & \Psi(\mathbf{x}_2)^T\Psi\left(\mathbf{x}_2\right) & \cdots & \Psi(\mathbf{x}_2)^T\Psi\left(\mathbf{x}_M\right) \\ \vdots & \vdots & \vdots & \vdots \\ \Psi(\mathbf{x}_M)^T\Psi\left(\mathbf{x}_1\right) & \Psi(\mathbf{x}_M)^T\Psi\left(\mathbf{x}_2\right) & \cdots & \Psi(\mathbf{x}_M)^T\Psi\left(\mathbf{x}_M\right) \end{bmatrix}$$

$$\overline{\mathbf{K}}_{ij} = \Psi(\mathbf{x}_i)^T\Psi\left(\mathbf{x}_j\right)$$

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}$$

那么

$$\lambda M\mathbf{a} = \overline{\lambda}\mathbf{a} = \overline{\mathbf{K}}\mathbf{a}$$

即$\mathbf{a}$是$\overline{\mathbf{K}}$的特征向量。

关于$\overline{\mathbf{K}}$的求法：

设$\boldsymbol{I} \in R^{M\times M}$，$\mathbf{I}_{ij} = 1$，$i = 1, 2, 3, \cdots, M$，$j = 1, 2, 3, \cdots, M$，即$\mathbf{I}$为全1矩阵，那么

$$\overline{\mathbf{K}}_{ij} = \Psi(\mathbf{x}_i)^T\Psi\left(\mathbf{x}_j\right) = \left(\Phi\left(\mathbf{x}_i\right) - \frac{1}{M}\sum_{m=1}^{M}\Phi\left(\mathbf{x}_m\right)\right)^T\left(\Phi\left(\mathbf{x}_j\right) - \frac{1}{M}\sum_{n=1}^{M}\Phi\left(\mathbf{x}_n\right)\right)$$

$$= \Phi(\mathbf{x}_i)^T\Phi\left(\mathbf{x}_j\right) - \frac{1}{M}\sum_{m=1}^{M}\Phi(\mathbf{x}_m)^T\Phi\left(\mathbf{x}_j\right) - \frac{1}{M}\sum_{n=1}^{M}\Phi(\mathbf{x}_i)^T\Phi\left(\mathbf{x}_n\right) + \frac{1}{M^2}\sum_{m,n=1}^{M}\Phi(\mathbf{x}_m)^T\Phi\left(\mathbf{x}_n\right)$$

$$= \mathbf{K}_{ij} - \frac{1}{M}\sum_{m=1}^{M}\mathbf{I}_{im}\mathbf{K}_{mj} - \frac{1}{M}\sum_{n=1}^{M}\mathbf{K}_{in}\mathbf{I}_{nj} + \frac{1}{M^2}\sum_{m,n=1}^{M}\mathbf{I}_{im}\mathbf{K}_{mn}\mathbf{I}_{nj}$$

其中：$\mathbf{K}_{ij}=\Phi(\mathbf{x}_i)^T\Phi\left(\mathbf{x}_j\right)$，$\mathbf{K}$为核矩阵，即$\mathbf{K}_{ij}=K\left(\mathbf{x}_i,\mathbf{x}_j\right)=\Phi(\mathbf{x}_i)^T\Phi\left(\mathbf{x}_j\right)$。

利用矩阵乘法，可以转化为

$$\overline{\mathbf{K}}=\mathbf{K}-\mathbf{I}_M\mathbf{K}-\mathbf{K}\mathbf{I}_M-\mathbf{I}_M\mathbf{K}\mathbf{I}_M \tag{1}$$

其中$\mathbf{I}_M=\frac{1}{M}\mathbf{I}$。

至此，$\overline{\mathbf{K}}$已经求出，那么特征向量$\mathbf{a}$也可以求出了。

求解新数据$\mathbf{t}_j$，$j=1,2,3,\cdots,L$的核主成分，即求：

$$\left(\Psi\left(\mathbf{t}_j\right),\mathbf{v}\right)=\sum_{i=1}^{M}a_i\left(\Psi\left(\mathbf{x}_i\right),\Psi\left(\mathbf{t}_j\right)\right)=\sum_{i=1}^{M}a_i\Psi(\mathbf{x}_i)^T\Psi\left(\mathbf{t}_j\right)$$

同$\overline{\mathbf{K}}$的求法

$$\begin{aligned}&\overline{\mathbf{K}}_{ij}^{\text{test}}=\Psi(\mathbf{x}_i)^T\Psi\left(\mathbf{t}_j\right)=\left(\Phi\left(\mathbf{x}_i\right)-\frac{1}{M}\sum_{m=1}^{M}\Phi\left(\mathbf{x}_m\right)\right)^T\left(\Phi\left(\mathbf{t}_j\right)-\frac{1}{M}\sum_{n=1}^{M}\Phi\left(\mathbf{x}_n\right)\right)\\&=\Phi(\mathbf{x}_i)^T\Phi\left(\mathbf{t}_j\right)-\frac{1}{M}\sum_{m=1}^{M}\Phi(\mathbf{x}_m)^T\Phi\left(\mathbf{t}_j\right)-\frac{1}{M}\sum_{n=1}^{M}\Phi(\mathbf{x}_i)^T\Phi\left(\mathbf{x}_n\right)+\frac{1}{M^2}\sum_{m,n=1}^{M}\Phi(\mathbf{x}_m)^T\Phi\left(\mathbf{x}_n\right)\\&=\mathbf{K}_{ij}^{test}-\frac{1}{M}\sum_{m=1}^{M}\mathbf{I}_{im}\mathbf{K}_{mj}^{test}-\frac{1}{M}\sum_{n=1}^{M}\mathbf{K}_{in}\bar{\mathbf{I}}_{nj}+\frac{1}{M^2}\sum_{m,n=1}^{M}\mathbf{I}_{im}\mathbf{K}_{mn}\mathbf{I}_{nj}\end{aligned}$$

其中$\mathbf{K}_{ij}^{\text{test}}=\Phi(\mathbf{x}_i)^T\Phi\left(\mathbf{t}_j\right)$。

转化为矩阵乘法，即

$$\overline{\mathbf{K}}^{test}=\mathbf{K}^{test}-\frac{1}{M}\mathbf{I}_M\mathbf{K}^{test}-\frac{1}{M}\mathbf{K}\mathbf{I}_{M\times L}+\frac{1}{M^2}\left(\sum_{m,n=1}^{M}\Phi(\mathbf{x}_m)^T\Phi\left(\mathbf{x}_n\right)\right)\mathbf{I}_{M\times L}$$

其中：$\mathbf{I}_M\in R^{M\times M}$和$\mathbf{I}_{M\times L}\in R^{M\times L}$每一元素都为1。

这里还有一个小问题，由于$\mathbf{v}$受到单位向量的约束，所以$\mathbf{a}$也要受到约束，即

$$\begin{aligned}(\mathbf{v},\mathbf{v})&=\left(\sum_{i=1}^{M}a_i\Psi\left(\mathbf{x}_i\right)\right)^T\left(\sum_{j=1}^{M}a_j\Psi\left(\mathbf{x}_j\right)\right)\\&=\sum_{i=1}^{M}\sum_{j=1}^{M}a_ia_j\left(\Psi\left(\mathbf{x}_i\right),\Psi\left(\mathbf{x}_j\right)\right)\\&=\sum_{i=1}^{M}\sum_{j=1}^{M}a_ia_j\overline{\mathbf{K}}_{ij}\\&=\overline{\lambda}\mathbf{a}^T\mathbf{a}\\&=1\end{aligned}$$

其中$\overline{\lambda}$为特征向量$\mathbf{a}$对应的$\overline{\mathrm{K}}$的特征值。

这里令$\mathbf{a}^T\mathbf{a}$也单位化为1，那么$\mathbf{a}^T\mathbf{a}$除以$\sqrt{\overline{\lambda}}$j即能保证$\|\mathbf{v}\|^2=1$。

3. KPCA算法步骤

1）选择一个核函数

2）计算核矩阵$\mathbf{K}$

3）按照$\overline{\mathbf{K}}=\mathbf{K}-\mathbf{I}_M\mathbf{K}-\mathbf{K}\mathbf{I}_M-\mathbf{I}_M\mathbf{K}\mathbf{I}_M$计算中心化的核矩阵$\overline{\mathbf{K}}$

4）求出中心化核矩阵$\overline{\mathbf{K}}$的特征向量和特征值

5）将特征向量按对应特征值大小从上到下按行排列成矩阵，取前k列组成矩阵α

6）将每一行特征向量单位化，每一行特征向量除去对应特征值的开方值

7）计算降维后数据的核主成分投影

$$y_j=\sum_{i=1}^{M}\alpha_{ji}K\left(x_i,x\right),j=1,\cdots k \tag{3.1}$$

注意这里y_j是一个scalar，代表第j个核主成分，α_j代表第j个特征向量，α_{ji}代表第j个特征向量的第i个元素（$i=1,\cdots,M$），每一个i对应了一个$K(x_i,x)$。所以式（3.1）实际上是利用α_{ji}对每一个训练样本x_i和测试样本的核$K(x_i,x)$加权求和。