

# KFDA

hgy

TongJi University

December 13, 2013

# 目录

## ① Fisher线性判别分析(C类)

## ② 核函数

- 定义
- 性质

## ③ 核Fisher判别法

- 定义
- 推理过程1
- 推理过程2
- 推理结果
- 疑惑

## ④ 参考

# Fisher线性判别分析(C类)

- 核心思想(详情参见课件PatternClassification\_8.ppt)  
选择使得Fisher准则函数达到极值的矢量作为最佳投影方向 $\vec{W}$ ,从而使得样本在该方向上投影后,达到最大的类间离散度 $S_b$ 和最小的类内离散度 $S_w$ .
- 表达式

$$\vec{W}^* = \arg \max_{\vec{W}} J(\vec{W}) = \arg \max_{\vec{W}} \left( \frac{\vec{W}^T S_b \vec{W}}{\vec{W}^T S_w \vec{W}} \right)$$

$$S_b = \sum_{i=1}^c N_i (\vec{v}_i - v)^T (\vec{v}_i - v)$$

$$S_w = \sum_{i=1}^c \sum_{j=1}^{N_i} (\vec{X}_i^j - \vec{v}_i)^T (\vec{X}_i^j - \vec{v}_i)$$

# 定义

- 核函数

设 $x, z \in X$ ,  $X$ 属于 $R(n)$ 空间, 非线性映射函数 $\phi$ 实现输入空间 $X$ 到特征空间 $F$ 的映射, 其中 $F$ 属于 $R(m)$ ,  $n \ll m$ 。根据核函数技术有:

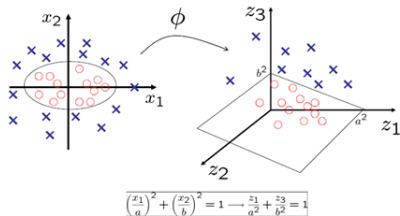
$$\kappa(\vec{X}, \vec{Z}) = \langle \phi(\vec{X}), \phi(\vec{Z}) \rangle$$

其中:  $\langle, \rangle$  为内积,  $\kappa(x, z)$  为核函数。

从上式可以看出, 核函数将 $m$ 维高维特征空间的内积运算转化为 $n$ 维低维输入空间的核函数计算, 巧妙地解决了在高维特征空间中计算中可能出现的“维数灾难”等问题, 从而为在高维特征空间中解决复杂的分类或回归问题奠定了理论基础。

# 性质

接受低维空间的输入值 $\vec{X}$ 和 $\vec{Z}$ ，却能算出高维空间的内积值 $\langle \phi(\vec{X}), \phi(\vec{Z}) \rangle$ .

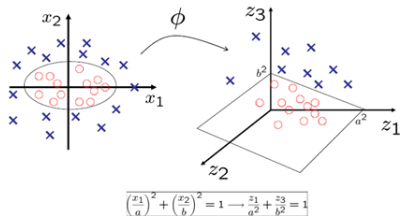


特征映射:

$$\phi(x_1, x_2) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$$

# 性质

接受低维空间的输入值 $\vec{X}$ 和 $\vec{Z}$ ，却能算出高维空间的内积值 $\langle\phi(\vec{X}),\phi(\vec{Z})\rangle$ 。



## Definition

$$\vec{X} = (x_1, x_2) \quad \vec{Z} = (z_1, z_2)$$

$$\begin{aligned}\kappa(\vec{X}, \vec{Z}) &= \langle \phi(\vec{X}), \phi(\vec{Z}) \rangle \\ &= (\vec{X} \cdot \vec{Z})^2\end{aligned}$$

特征映射:

$$\phi(x_1, x_2) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$$

$$\begin{aligned}\kappa(\vec{X}, \vec{X}) &= (\vec{X} \cdot \vec{X})^2 \\ &= (x_1z_1 + x_2z_2)^2 \\ &= (x_1^2, \sqrt{2}x_1x_2, x_2^2) * (z_1^2, \sqrt{2}z_1z_2, z_2^2) \\ &= \langle \phi(\vec{X}), \phi(\vec{Z}) \rangle\end{aligned}$$

# 定义

基于核的Fisher线性鉴别分析方法能够在原始输入空间中抽取非线性(在高维空间呈线性)最佳鉴别特征. 对于一个数据集 $\chi$ 和一个特征映射函数 $\phi$ ,在F空间类间散布矩阵 $S_b^\phi$ 和类内散布矩阵 $S_w^\phi$ 可以被定义为:

$$S_b^\phi = \sum_{i=1}^c N_i (\vec{v}_i^\phi - v^\phi)(\vec{v}_i^\phi - v^\phi)^T$$

$$S_w^\phi = \sum_{i=1}^c \sum_{j=1}^{N_i} (\phi(\vec{X}_i^j) - \vec{v}_i)(\phi(\vec{X}_i^j) - \vec{v}_i)^T$$

在特征空间F中利用最大化线性准则函数获取最优投影方向:

$$\vec{V}^* = \arg \max_{\vec{V}} J(\vec{V}) = \arg \max_{\vec{V}} \left( \frac{\vec{V}^T S_b^\phi \vec{V}}{\vec{V}^T S_w^\phi \vec{V}} \right)$$

## 推理过程1

易知  $\vec{V} \in F$ , 即  $\vec{V}$  一定在由  $\phi(X_i^j)$  所伸展的空间中, 因此它可以表示为:

$$\vec{V} = \sum_{i=1}^c \sum_{j=1}^{N_i} w_i^j \phi(\vec{X}_i^j) = \phi(X)^T \vec{W}$$

注:

$$\begin{aligned} \vec{W} &= (w_1^1, \dots, w_1^{N_1}, \dots, w_c^1, \dots, w_c^{N_c})^T && \sum_{i=1}^c N_i * 1 \\ \phi(X) &= (\phi(\vec{X}_1^1), \dots, \phi(\vec{X}_1^{N_1}), \dots, \phi(\vec{X}_c^1), \dots, \phi(\vec{X}_c^{N_c}))^T && \sum_{i=1}^c N_i * m \end{aligned}$$

带入到上式得到:

$$\begin{aligned} \vec{V}^T S_b^\phi \vec{V} &= \vec{W}^T \{ \phi(X)^T S_b^\phi \phi(X) \} \vec{W} \\ \vec{V}^T S_w^\phi \vec{V} &= \vec{W}^T \{ \phi(X)^T S_w^\phi \phi(X) \} \vec{W} \end{aligned}$$



## 推理过程2

上式{}中分别用 $K_b$ 和 $K_w$ 表示[ $K_b$ 和 $K_w$ 都是 $\sum_{i=1}^c N_i * \sum_{i=1}^c N_i$ ],则问题转变为:

$$\vec{W}^* = \arg \max_{\vec{W}} J(\vec{W}) = \arg \max_{\vec{W}} \left( \frac{\vec{W}^T K_b \vec{W}}{\vec{W}^T K_w \vec{W}} \right)$$

由广义的瑞利商得:

$\vec{W}^*$ 为 $K_w^{-1}K_b$ 的 $\sum_{i=1}^c N_i$ 个特征向量中的一个。  
 $\vec{W}^*$ 为 $\sum_{i=1}^c N_i * 1$ 维向量。

# 推理结果

$$W = (\vec{W}_1, \dots, \vec{W}_m)$$

注:

$\vec{W}_i (1 \leq i \leq m)$  为  $K_w^{-1} K_b$  的  $m$  个最大的特征向量

对于任何一个输入矢量  $\vec{X}$ , 它的低维特征  $\vec{Y} = (y_1, \dots, y_m)$  可以表示为:

$$\vec{Y} = (\vec{W}_1, \dots, \vec{W}_m)^T *$$

$$(\kappa(\vec{X}_1^1, \vec{X}), \dots, \kappa(\vec{X}_1^{N_1}, \vec{X}), \dots, \kappa(\vec{X}_c^1, \vec{X}), \dots, \kappa(\vec{X}_c^{N_c}, \vec{X}))$$

# 疑惑

- ① 推理结果:低维特征 $\vec{Y}$ 代表了什么意义,为嘛它是 $m$ 维(高维),而不是 $n$ 维(低维)?

根据模式分类课件PatternClassification\_8.ppt第18,19页的推理可知:权重向量 $\vec{W}$ 取的是 $K_w^{-1}K_b$ 的最大特征值对应的特征向量. ( $\vec{W}$ 是 $\sum_{i=1}^c N_i * 1$ 的列向量)

假定将训练样本集 $X$ (低维 $n$ )映射到 $F$ 空间 $\phi(X)$ (高维 $m$ ),并且用FDA方法求解出最优投影方向为 $\vec{V}$ ,对于任何一个测试样本 $\vec{T}$ ,它被映射成高维空间的 $\phi(\vec{T})$ ,计算的标量 $y = \vec{V}^T * \phi(\vec{T})$ ,根据 $y$ 的具体数值区分 $\phi(\vec{T})$ 属于哪一类.[2]

核函数的定义:  $\kappa(\vec{X}, \vec{W}) = \langle \phi(\vec{X}), \phi(\vec{W}) \rangle$

$$y = \vec{V}^T * \phi(\vec{X}) = \vec{W}^T * \phi(X) * \phi(\vec{T})$$

$$= \vec{W}^T (\kappa(\vec{X}_1^1, \vec{X}), \dots, \kappa(X_1^{N_1}, \vec{X}), \dots, \kappa(\vec{X}_c^1, \vec{X}), \dots, \kappa(X_c^{N_c}, \vec{X}))$$

# 参考



李璇and 杨杰.

基于核主元分析方法的kfda 人脸识别研究.  
2010.



Jasper.

Svm入门（七）为何需要核函数<http://www.blogjava.net/zhenandaci/archive/2009/03/06/258288.html>.  
In *BlogJava*.



Jasper.

Svm入门（六）线性分类器的求解——问题的转化，直观角度<http://www.blogjava.net/zhenandaci/archive/2009/03/01/257237.html>.  
In *BlogJava*.