媒体与认知 第四次作业

无 81 马啸阳 2018011054

2020 年 5 月 27 日

1 线性判别分析

$$\begin{aligned}\mu_1 &= \left(-1, \frac{5}{3}\right)^T \\ \mu_2 &= \left(\frac{5}{3}, -\frac{2}{3}\right)^T \\ \mu &= \left(\frac{1}{3}, -\frac{1}{2}\right)^T\end{aligned}$$

从而有类内散布矩阵

$$\begin{aligned}S_w &= \frac{1}{6}\left(\sum_{k=1}^{3}(x_k-\mu_1)(x_k-\mu_1)^T + \sum_{k=4}^{6}(x_k-\mu_2)(x_k-\mu_2)^T\right) \\ &= \begin{pmatrix} \frac{10}{9} & \frac{8}{9} \\ \frac{8}{9} & \frac{8}{9} \end{pmatrix}\end{aligned}$$

则由 Fisher 判据，LDA 投影方向为

$$\begin{aligned}w &= S_w^{-1}(\mu_1-\mu_2) \\ &= \left(-\frac{45}{2}, \frac{201}{8}\right)^T\end{aligned}$$

2 隐含马尔可夫模型

2.1 前向变量法

$$
\begin{aligned}
&\alpha_1(1) = P(A|S_1)P(S_1) = 0.2 \\
&\alpha_1(2) = P(A|S_2)P(S_2) = 0.05 \\
&\alpha_2(1) = P(C|S_1)\left[\alpha_1(1)P(S_1|S_1) + \alpha_1(2)P(S_1|S_2)\right] = 0.015 \\
&\alpha_2(2) = P(C|S_2)\left[\alpha_1(1)P(S_2|S_1) + \alpha_1(2)P(S_2|S_2)\right] = 0.04 \\
&\alpha_3(1) = P(T|S_1)\left[\alpha_2(1)P(S_1|S_1) + \alpha_2(2)P(S_1|S_2)\right] = 0.00185 \\
&\alpha_3(2) = P(T|S_2)\left[\alpha_2(1)P(S_2|S_1) + \alpha_2(2)P(S_2|S_2)\right] = 0.0146
\end{aligned}
$$

从而有产生观测序列 x 的概率

$$P(x|\lambda) = \alpha_3(1) + \alpha_3(2) = 0.01645$$

2.2 Viterbi 算法

$$
\begin{aligned}
\delta_1(1) &= P(A|S_1)P(S_1) = 0.2 \\
\phi_1(1) &= 0 \\
\delta_1(2) &= P(A|S_2)P(S_2) = 0.05 \\
\phi_1(2) &= 0 \\
\delta_2(1) &= P(C|S_1)\max\left[\alpha_1(1)P(S_1|S_1), \alpha_1(2)P(S_1|S_2)\right] = 0.014 \\
\phi_2(1) &= 1 \\
\delta_2(2) &= P(C|S_2)\max\left[\alpha_1(1)P(S_2|S_1), \alpha_1(2)P(S_2|S_2)\right] = 0.024 \\
\phi_2(2) &= 1 \\
\delta_3(1) &= P(T|S_1)\max\left[\alpha_2(1)P(S_1|S_1), \alpha_2(2)P(S_1|S_2)\right] = 0.00098 \\
\phi_3(1) &= 1 \\
\delta_3(2) &= P(T|S_2)\max\left[\alpha_2(1)P(S_2|S_1), \alpha_2(2)P(S_2|S_2)\right] = 0.00768 \\
\phi_3(2) &= 2 \\
P^* &= \max[\delta_3(1), \delta_3(2)] = 0.00768 \\
q_3^* &= 2 \\
q_2^* &= \phi_3(2) = 2 \\
q_1^* &= \phi_2(2) = 1
\end{aligned}
$$

即最可能的隐含状态序列是 $1, 2, 2$。

3 支持向量机

3.1

变换后负样本为 $Z_1 = (1, -2), Z_2 = (4, -5), Z_3 = (4, -1)$，正样本为 $Z_4 = (5, -2), Z_5 = (7, -7), Z_6 = (7, 1), Z_7 = (7, 1)$。

使用 libSVM 求解得 SVM 参数为

```
svm_type c_svc
kernel_type linear
nr_class 2
```

```
total_sv 3
rho 8.9969222083333396
label 1 −1
nr_sv 1 2
SV
1.9992897403846175 1:5 2:−2
−0.49982243509615443 1:4 2:−5
−1.4994673052884633 1:4 2:−1
```

即 Z 空间上支持向量有

$$
\begin{aligned}
\mathbf{z_1} &= (5,-2), \quad y_1 = 1, \quad \alpha_1 = 2 \\
\mathbf{z_1} &= (4,-5), \quad y_1 = -1, \quad \alpha_1 = 0.5 \\
\mathbf{z_1} &= (4,-1), \quad y_1 = -1, \quad \alpha_1 = 1.5
\end{aligned}
$$

从而 Z 空间上，$\mathbf{w} = 2(5,-2) - 0.5(4,-5) - 1.5(4,-1) = (2,0), b = -9$，从而决策函数为

$$f_Z(\mathbf{z}) = \mathbf{w}^T\mathbf{z} + b = 2z_1 - 9$$

还原到 X 空间上决策函数为

$$g_m(x_1,x_2) = 2\phi_1(\mathbf{x}) - 9 = 2x_2^2 - 4x_1 - 3$$

3.2

Lagrangian 对偶问题的目标函数如下，其中 $K(\mathbf{x_i},\mathbf{x_j}) = (1 + X_i^T X_j)^2$。

$$
\begin{aligned}
\text{maximize} \sum_{i=1}^{7}\alpha_i - \frac{1}{2}\sum_{i=1}^{7}\sum_{j=1}^{7}\alpha_i\alpha_j y_i y_j K(\mathbf{x_i},\mathbf{x_j}) \\
\text{s.t.} \quad 0 \le \alpha_i \le C, \quad \sum_{i=1}^{7}\alpha_i y_i = 0
\end{aligned}
$$

使用 libSVM 求解得 SVM 参数为

```
svm_type c_svc
kernel_type polynomial
degree 2
```

```
gamma 1
coef0 1
nr_class 2
total_sv 5
rho 1.6665309232321686
label 1 -1
nr_sv 3 2
SV
0.88871643479870632 1:-1 2:0
0.15028522296992985 1:0 2:2
0.36817043235512975 1:0 2:-2
-0.4857048695492317 1:0 2:1
-0.92146722057453434 1:0 2:-1
```

即有 X_2, X_3, X_4, X_5, X_6 为支持向量，α_i 取值如下，$b = -1.6665309232321686$

$$
\begin{aligned}
\alpha_1 &= 0 \\
\alpha_2 &= 0.4857048695492317 \\
\alpha_3 &= 0.92146722057453434 \\
\alpha_4 &= 0.88871643479870632 \\
\alpha_5 &= 0.15028522296992985 \\
\alpha_6 &= 0.36817043235512975 \\
\alpha_7 &= 0
\end{aligned}
$$

决策函数有

$$
\begin{aligned}
g_k(x_1, x_2) &= \sum_{j \in SV} y_j \alpha_j K(X_j, X) + b \\
&= \sum_{j=2}^{6} y_j \alpha_j (1 + x_{j1} x_1 + x_{j2} x_2)^2 + b \\
&\approx \frac{8}{9} x_1^2 - \frac{16}{9} x_1 + \frac{2}{3} x_2^2 - \frac{5}{3}
\end{aligned}
$$

3.3

如图 1所示，其中蓝色点为负样本，红色点为正样本，蓝色椭圆为 $g_k(x_1, x_2)$，黄色抛物线为 $g_m(x_1, x_2)$。二者虽然都区分了样本点，但区分方式不同。核函数可以实现非线性映射的作

用。实际上，映射 ϕ_1, ϕ_2 即对应核函数 $K(X_i, X_j) = \phi_1(X_i)\phi_1(X_j) + \phi_2(X_i)\phi_2(X_j)$。而且实际上，不需要显式非线性映射即可用核函数完成 SVM。核函数相当于投影至高维非线性映射，而不必显式写出非线性映射，能够非线性超平面划分。

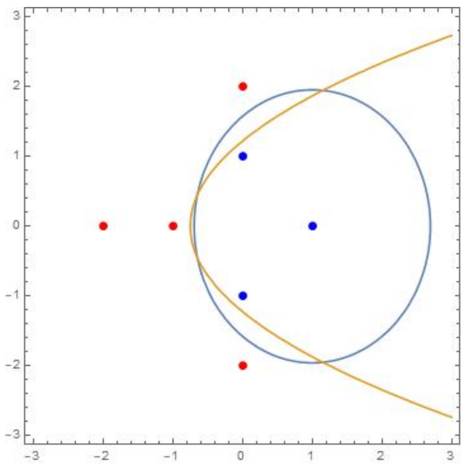

图 1: 决策函数比较图

3.4

$$
\begin{aligned}
g_m(X_8) &= -3 < 0 \\
g_k(X_8) &= -\frac{15}{9} < 0 \\
g_m(X_9) &= -3 < 0 \\
g_k(X_9) &= 1 > 0
\end{aligned}
$$

因此 $Y_{8m} = -1, Y_{8k} = -1, Y_{9m} = -1, Y_{9k} = 1$。