机器学习导论
习题五

14120132, 银琦, 141220132@smail.nju.edu.cn

2017 年 5 月 30 日

1 [25pts] Bayes Optimal Classifier

试证明在二分类问题中，但两类数据同先验、满足高斯分布且协方差相等时，LDA可产生贝叶斯最优分类器。

Solution.
在二分类问题中，由贝叶斯定理有

$$P(c|x) = \frac{P(c)P(x|c)}{P(x)}$$

设两个类分别为c_1, c_2由于两类数据同先验，所以$P(c_1) = P(c_2)$，而P(x)与类别无关，所以下面只考虑$P(x|c)$。因为两类数据均满足高斯分布，所以

$$P(c_1|x) = \frac{1}{\sqrt{2\pi \cdot \Sigma_{c_1}}} \cdot exp(-\frac{(\mathbf{x} - \mu_{c_1})^2}{2\Sigma_{c_1}})$$

$$P(c_2|x) = \frac{1}{\sqrt{2\pi \cdot \Sigma_{c_2}}} \cdot exp(-\frac{(\mathbf{x} - \mu_{c_2})^2}{2\Sigma_{c_2}})$$

因为两类数据的协方差相等，所以$\Sigma_{c_1} = \Sigma_{c_2}$，所以下面仅考虑$(\mathbf{x} - \mu_{c_1})^2$与$(\mathbf{x} - \mu_{c_2})^2$。
为了找出最优贝叶斯分类器，令$P(c_1|x) = P(c_2|x)$，即：

$$(\mathbf{x} - \mu_{c_1})^2 = (\mathbf{x} - \mu_{c_2})^2$$

$$\Leftrightarrow (\mathbf{x} - \mu_{c_1})^2 - (\mathbf{x} - \mu_{c_2})^2 = 0$$

$$\Leftrightarrow \mathbf{x}^2 + \mu_{c_1}^2 - 2\mu_{c_1}\mathbf{x} - \mathbf{x}^2 - \mu_{c_2}^2 + 2\mu_{c_2}\mathbf{x} = 0$$

$$\Leftrightarrow (2\mu_{c_2} - 2\mu_{c_1})\mathbf{x} + (\mu_{c_1}^2 - \mu_{c_2}^2) = 0$$

上式可以化为$\mathbf{w^T x} + b = 0$的格式，其中$w_i = 2(\mu_{c_2} - \mu_{c_1})$，于是LDA可求出$\mathbf{w}$的最优解，所以LDA产生了贝叶斯最优分类器。

2 [25pts] Naive Bayes

考虑下面的400个训练数据的数据统计情况，其中特征维度为2（$\mathbf{x} = [x_1, x_2]$），每种特征取值0或1，类别标记$y \in \{-1, +1\}$。详细信息如表1所示。

根据该数据统计情况，请分别利用直接查表的方式和朴素贝叶斯分类器给出$\mathbf{x} = [1, 0]$的测试样本的类别预测，并写出具体的推导过程。

Table 1: 数据统计信息

x_1	x_2	$y = +1$	$y = -1$
0	0	90	10
0	1	90	10
1	0	51	49
1	1	40	60

Solution.

(1)查表方式：在表1中，$\mathbf{x} = [1, 0]$时，+1类别有51个，-1类别有49个，所以此时$\mathbf{x}$的类别应为$y = +1$。

(2)朴素贝叶斯分类器：首先估计类先验概论$P(c)$，显然有

$$P(y = +1) = \frac{90 + 90 + 51 + 40}{400} = \frac{271}{400} = 0.6775,$$

$$P(y = -1) = \frac{10 + 10 + 49 + 60}{400} = \frac{129}{400} = 0.3225,$$

然后为每个属性估计条件概率$P(x_i|c)$:

$$P(x_1 = 1|y = +1) = \frac{51 + 40}{90 + 90 + 51 + 40} = \frac{91}{271} \approx 0.3358$$

$$P(x_1 = 1|y = -1) = \frac{60 + 49}{60 + 49 + 10 + 10} = \frac{109}{129} \approx 0.8450$$

$$P(x_2 = 0|y = +1) = \frac{90 + 51}{90 + 90 + 51 + 40} = \frac{141}{271} \approx 0.5203$$

$$P(x_2 = 0|y = -1) = \frac{49 + 10}{10 + 10 + 49 + 60} = \frac{59}{129} \approx 0.4574$$

于是有：

$$P(y = +1) \times P(x_1 = 1|y = +1) \times P(x_2 = 0|y = +1) = 0.1183$$

$$P(y = -1) \times P(x_1 = 1|y = -1) \times P(x_2 = 0|y = -1) = 0.1246$$

由于$0.1183 < 0.1246$,所以朴素贝叶斯分类器将测试样本$\mathbf{x} = [1, 0]$判别为$y = -1$。

3 [25pts] Bayesian Network

贝叶斯网(Bayesian Network)是一种经典的概率图模型，请学习书本7.5节内容回答下面的问题：

(1) **[5pts]** 请画出下面的联合概率分布的分解式对应的贝叶斯网结构：

$$\Pr(A,B,C,D,E,F,G) = \Pr(A)\Pr(B)\Pr(C)\Pr(D|A)\Pr(E|A)\Pr(F|B,D)\Pr(G|D,E)$$

(2) **[5pts]** 请写出图1中贝叶斯网结构的联合概率分布的分解表达式。

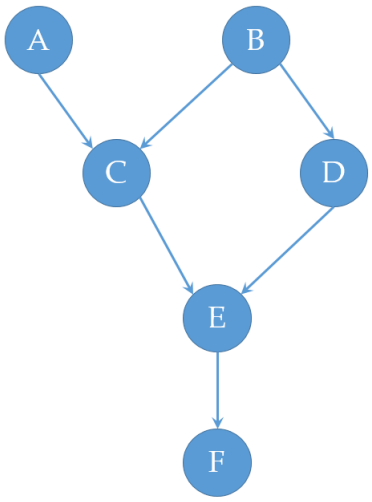


Figure 1: 题目3-(2)有向图

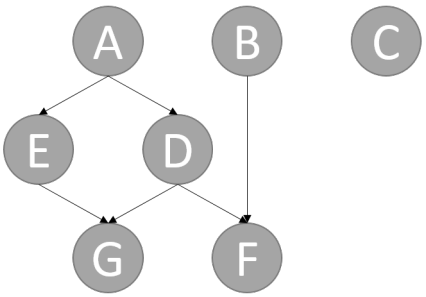
(3) **[15pts]** 基于第(2)问中的图1, 请判断表格3中的论断是否正确，只需将下面的表格填完整即可。

Table 2: 判断表格中的论断是否正确

序号	关系	True/False	序号	关系	True/False
1	$A \perp\!\!\!\perp B$		7	$F \perp B \mid C$	
2	$A \perp B \mid C$		8	$F \perp B \mid C, D$	
3	$C \perp\!\!\!\perp D$		9	$F \perp B \mid E$	
4	$C \perp D \mid E$		10	$A \perp\!\!\!\perp F$	
5	$C \perp D \mid B, F$		11	$A \perp F \mid C$	
6	$F \perp\!\!\!\perp B$		12	$A \perp F \mid D$	

Solution.

(1)贝叶斯网络结构图如下：



(2)图1中贝叶斯网络结构的联合概率分布分解表达式如下：

$$\Pr(A, B, C, D, E, F) = \Pr(A)\Pr(B)\Pr(C|A, B)\Pr(D|B)\Pr(E|C, D)\Pr(F|E)$$

(3)

序号	关系	True/False	序号	关系	True/False
1	$A \perp\!\!\!\perp B$	True	7	$F \perp B\|C$	False
2	$A \perp B\|C$	False	8	$F \perp B\|C, D$	True
3	$C \perp\!\!\!\perp D$	False	9	$F \perp B\|E$	True
4	$C \perp D\|E$	False	10	$A \perp\!\!\!\perp F$	False
5	$C \perp D\|B, F$	False	11	$A \perp F\|C$	False
6	$F \perp\!\!\!\perp B$	False	12	$A \perp F\|D$	False

4 [25pts] Naive Bayes in Practice

请实现朴素贝叶斯分类器，同时支持离散属性和连续属性。详细编程题指南请参见链接：`http://lamda.nju.edu.cn/ml2017/PS5/ML5_programming.html`.