机器学习导论
习题五

141220120, 徐世坚, xsj13260906215@gmail.com

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. 此处用于写证明(中英文均可)
已知两类数据同先验，满足高斯分布且协方差相同，设方差均为Σ
贝叶斯最优分类器要求选择使后验概率$P(c|\mathbf{x})$最大的类别标记。
$P(c|(x)) = \frac{P(c)P(\mathbf{x}|c)}{P(\mathbf{x})}$,又$P(\mathbf{x})$对所有类别都相同，所以舍去。
对上述化简后的式子取对数，同时，因为先验分布相同，所以舍去$log(P(c))$项，则，对于贝叶斯最优分类器来说，第i类的决策函数为：
$f_i = \mathbf{x}^T\Sigma^{-1}\mu_i - \frac{1}{2}\mu_i^T\Sigma^{-1}\mu_i$
所以，贝叶斯最优分类器的判别函数为：
$f = f_0 - f_1 = \mathbf{x}^T\Sigma^{-1}\mu_0 - \frac{1}{2}\mu_0^T\Sigma^{-1}\mu_0 - \mathbf{x}^T\Sigma^{-1}\mu_1 + \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1$
$= \mathbf{x}^T\Sigma^{-1}(\mu_0 - \mu_1) + \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 - \frac{1}{2}\mu_0^T\Sigma^{-1}\mu_0$
下面考虑FLD。由FLD的计算可知，$\omega = (\Sigma_0 + \Sigma_1)^{-1}(\mu_0 - \mu_1) = \frac{1}{2}\Sigma^{-1}(\mu_0 - \mu_1)$
所以，FLD的判别函数为：
$\omega(\mathbf{x} - \frac{1}{2}(\mu_0 + \mu_1)$
$= \mathbf{x}^T\omega - \frac{1}{2}(\mu_0 + \mu_1)^T\omega$
$= \frac{1}{2}\mathbf{x}^T\Sigma^{-1}(\mu_0 - \mu_1) - \frac{1}{4}(\mu_0 + \mu_1)\Sigma^{-1}(\mu_0 - \mu_1)$
$= \frac{1}{2}(\mathbf{x}^T\Sigma^{-1}(\mu_0 - \mu_1) + \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 - \frac{1}{2}\mu_0^T\Sigma^{-1}\mu_0)$
得到的形式和贝叶斯最优分类器的形式是一样的，只是多了常数项$\frac{1}{2}$.
所以，二分类任务中两类数据满足高斯分布，且方差相同、先验相同时，线性判别分析产生贝叶斯最优分类器。

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。

表 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. 此处用于写解答(中英文均可)

直接查表可知：

$P(y=+1|x=[1,0])=\frac{51}{100}, P(y=-1|x=[1,0])=\frac{49}{100}$

$\therefore$样本$\mathbf{x}=[1,0]$的预测类别是$y=+1$.

$P(y=+1)=\frac{271}{400} \quad P(y=-1)=\frac{129}{400}$

$P(x_1=0|y=+1)=\frac{180}{271} \quad P(x_1=1|y=+1)=\frac{91}{271}$

$P(x_2=0|y=+1)=\frac{141}{271} \quad P(x_2=1|y=+1)=\frac{130}{271}$

$P(x_1=0|y=-1)=\frac{20}{129} \quad P(x_1=1|y=-1)=\frac{109}{129}$

$P(x_2=0|y=-1)=\frac{59}{129} \quad P(x_2=1|y=-1)=\frac{70}{129}$

$P(y=+1)\prod_{i=1}^{2}P(x_i|y=+1)=P(y=+1)P(x_1=1|y=+1)P(x_2=0|y=+1)=0.1184$

$P(y=-1)\prod_{i=1}^{2}P(x_i|y=-1)=P(y=-1)P(x_1=1|y=-1)P(x_2=0|y=-1)=0.1246$

$\therefore$样本$\mathbf{x}=[1,0]$的预测类别是$y=-1$

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。

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

Solution. 此处用于写解答(中英文均可)

(1)

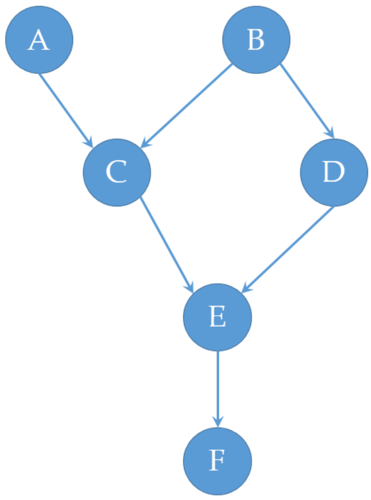
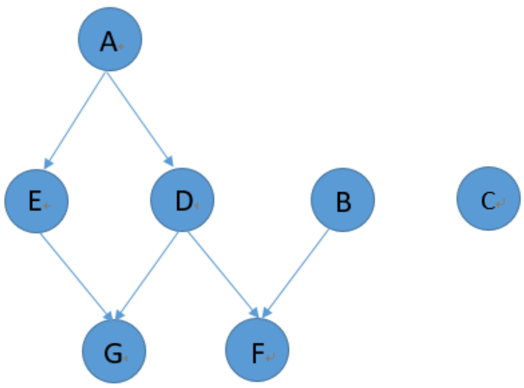


图 1: 题目3-(2)有向图

表 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	$A \perp\!\!\!\perp B$	T	7	$F \perp B \vert C$	F
2	$A \perp B \vert C$	F	8	$F \perp B \vert C, D$	T
3	$C \perp\!\!\!\perp D$	F	9	$F \perp B \vert E$	T
4	$C \perp D \vert E$	F	10	$A \perp\!\!\!\perp F$	F
5	$C \perp D \vert B, F$	F	11	$A \perp F \vert C$	F
6	$F \perp\!\!\!\perp B$	F	12	$A \perp F \vert D$	F



(2)$P_r(A,B,C,D,E,F) = P_r(A)P_r(B)P_r(C|A,B)P_r(D|B)P_r(E|C,D)P_r(F|E)$
(3)见填表。

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.