(1) 一个可以判别各个模式所属类别的线性函数
(2) 有些模式不是线性可分的，需要非线性函数将其分开
(3) $n=4 \quad r=1 \quad C_{n+r}^{r}=C_5^1=5$
$n=4 \quad r=2 \qquad C_6^2=\frac{6\times5}{1\times2}=15$

2017-2018 学年秋季学期　试题专用纸

中国科学院大学
试 题 专 用 纸

课程编号：091M4042H
课程名称：模式识别与机器学习
任课教师：黄庆明、兰艳艳、郭嘉丰、山世光

注意事项：
1.考试时间为 120 分钟，考试方式 闭 卷；
2.全部答案写在答题纸上；
3.考试结束后，请将本试卷和答题纸、草稿纸一并交回。

1. （8 分）试阐述线性判别函数的基本概念，并说明既然有线性判别函数，为什么还需要非线性判别函数？假设有两类模式，每类包括 6 个 4 维不同的模式，且良好分布。如果它们是线性可分的，问权向量至少需要几个系数分量？假如要建立二次的多项式判别函数，又至少需要几个系数分量？（设模式的良好分布不因模式变化而改变）

2. （8 分）简述 SVM 算法的原理。如果使用 SVM 做二分类问题得到如下结果，分别应该采取什么措施以取得更好的结果？并说明原因。
(1) 训练集的分类准确率 90%，验证集的分类准确率 90%，测试集的分类准确率 88%；欠拟合，增加惩罚系数C
(2) 训练集的分类准确率 98%，验证集的分类准确率 90%，测试集的分类准确率 88%。过 〃，减小 〃

3. （8 分）请从两种角度解释主成分分析（PCA）的优化目标。
① 最小化重建误差：样本点到超平面的距离足够近，样本点重建尽可能小，丢失更少信息，$\|x-\hat{x}\|$ 尽量小
② 最大化方差：寻找一个方向向量，使投影到上面的数据尽可能分散，即方差最大，而后继续寻找下一个正交于前一向量且使方差最大的向量

4. （8 分）请给出卷积神经网络 CNN 中卷积、Pooling、RELU 等基本层操作的含义。然后从提取特征的角度分析 CNN 与传统特征提取方法（例如 Gabor 小波滤波器）的异同。

5. （10 分）用线性判别函数的感知器赏罚训练算法求下列模式分类的解向量，并给出相应的判别函数。
ω_1: $\{(0\ 0)^T,\ (0\ 1)^T\}$
ω_2: $\{(1\ 0)^T,\ (1\ 1)^T\}$

将 ω_2 乘 (-1) 并写成增广形式：　$W(1)=(0\ 0\ 0)^T$　C 取 1
$x_1=(0\ 0\ 1)^T \quad x_2=(0\ 1\ 1)^T \quad x_3=(-1\ 0\ -1)^T \quad x_4=(-1\ -1\ -1)^T$
$W^T(1)x_1=(0\ 0\ 0)(0\ 0\ 1)=0 \not> 0 \quad W(2)=W(1)+x_1=(0\ 0\ 1)$
$W^T(2)x_2=(0\ 0\ 1)(0\ 1\ 1)=1>0 \quad W(3)=W(2)$

6. （10 分）试述 K-L 变换的基本原理，并将如下两类样本集的特征维数降到一维，同时画出样本在该空间中的位置。
ω_1: $\{(-5\ -5)^T,\ (-5\ -4)^T,\ (-4\ -5)^T,\ (-5\ -6)^T,\ (-6\ -5)^T\}$
ω_2: $\{(5\ 5)^T,\ (5\ 6)^T,\ (6\ 5)^T,\ (5\ 4)^T,\ (4\ 5)^T\}$，
其中假设其先验概率相等，即 $P(\omega_1)=P(\omega_2)=0.5$。

① 首先应保证 $E[x]=0$　$m=\frac{1}{2}\left(\frac{1}{5}\sum x_{1j}+\frac{1}{5}\sum x_{2j}\right)=0$
∴ 不需要平移
② 计算自相关矩阵 R
$R=P(\omega_1)E(x_1x_1^T)+P(\omega_2)E(x_2x_2^T)$
$=\frac{1}{2}\left[\frac{1}{5}\sum_{j=1}^{5}x_{1j}x_{1j}^T\right]+\frac{1}{2}[\quad]$
③ 计算 R 特征值 λ，$Rv=\lambda v \Rightarrow |R-\lambda I|v=0$
④ 取最大的 λ 对应的 v 单位化
⑤ $y=\Phi^T x$ 求得变换后的 ω_1'，ω_2'

（12 分）请解释 AdaBoost 的基本思想和工作原理，写出 AdaBoost 算法

① 选正交函数集

$\varphi_1(x)=\varphi_1(x_1,x_2)=H_0(x_1)H_0(x_2)=1$

$\varphi_2(x)=\varphi_2(x_1,x_2)=H_0(x_1)H_1(x_2)=$

$=H_0(x_1)H_2(x_2)=$

$=H_1(x_1)H_0(x_2)=$

$=H_1(x_1)H_1(x_2)=$

$\vdots$

$=H_2(x_1)H_0(x_2)=$

$\vdots$

② 生成势函数：

$K(x,x_k)=\sum_{i=1}^{9}\varphi_i(x)\varphi_i(x_k)$

③ 训练样本

取 $x_{①}=(0\ 1)^T\in\omega_1$

$K_1(x)=K(x_1,x_{①})$

$=|x_1|+2x_2(2\times1)+($

8.（12分）选择埃尔米特多项式，其前面几项的表达式为

$H_0(x)=1,\quad H_1(x)=2x,\quad H_2(x)=4x^2-2,$

$H_3(x)=8x^3-12x,\quad H_4(x)=16x^4-48x^2+12$

试用二次埃尔米特多项式的势函数算法求解以下模式的分类问题

ω_1: $\{(0\ 1)^T,\ (0\ -1)^T\}$

ω_2: $\{(1\ 0)^T,\ (-1\ 0)^T\}$

9. （12分）已知以下关于垃圾邮件的 8 条标注数据，A、B 为邮件的 2 个特征，Y 为类别，其中 Y=1 表示该邮件为垃圾邮件，Y=0 表示该邮件为正常邮件。请依此训练一个朴素贝叶斯分类器，并预测特征为“A=0，B=1”的邮件是否为垃圾邮件。

序号	1	2	3	4	5	6	7	8
A	0	0	1	1	1	1	1	1
B	0	0	0	0	0	0	1	1

10. （12分）假设有 3 个罐子，每个罐子里都装有红、黑两种颜色的弹珠。按照下面的方法取弹珠：开始，以概率 π 随机选取 1 个罐子，从这个罐子以概率 B 随机取出一个弹珠，记录其颜色后，放回；然后，从当前盒子以概率 A 随机转移到下一个盒子，再从这个盒子里以概率 B 随机抽出一个球，记录其颜色，放回；如此重复 3 次，得到一个弹珠的颜色观测序列：O=（红，黑，红）。请用前向传播算法计算生成该序列的概率 $P(O|\{A,B,\pi\})$。

$$\pi=[0.4,0.4,0.2]^T \qquad A=\begin{array}{c|ccc} & \text{罐子1} & \text{罐子2} & \text{罐子3} \\ \hline \text{罐子1} & 0.3 & 0.5 & 0.2 \\ \text{罐子2} & 0.2 & 0.3 & 0.5 \\ \text{罐子3} & 0.1 & 0.4 & 0.5 \end{array} \qquad B=\begin{array}{c|cc|c} & \overset{O_1}{\text{红}} & \overset{O_2}{\text{黑}} & \\ \hline \text{罐子1} & 0.7 & 0.3 & b_1 \\ \text{罐子2} & 0.5 & 0.5 & b_2 \\ \text{罐子3} & 0.4 & 0.6 & b_3 \end{array}$$

$\alpha_1(1)=\pi_1 b_1(O_1)=0.4\times0.7=0.28$ 选①号 拿红球

$\alpha_1(2)=\pi_2 b_2(O_1)=0.4\times0.5=0.2$ ②

$\alpha_1(3)=\pi_3 b_3(O_1)=0.2\times0.4=0.08$ ③

第一轮 $(0.4\ \ 0.4\ \ 0.2)\cdot b^T(O_1)=$

$\alpha_2(1)=\underbrace{\alpha_1(1)}_{\text{第一轮①}}\cdot\underbrace{a_{11}}_{\text{第一轮至①}}\cdot\underbrace{b_1(O_2)}_{\text{拿黑球}}+\underbrace{\alpha_1(2)}_{\text{第一轮②}}\cdot\underbrace{a_{21}}_{\text{第二轮转①}}\cdot b_1(O_2)+\underbrace{\alpha_1(3)}_{\text{第一轮③}}\cdot\underbrace{a_{13}}_{\text{第2轮转①}}\cdot b_1(O_2)=(0.28\times0.3+0.2\times0.2+0.08\times0.1)\times0.3=0.0369$

$\alpha_2(2)=(0.28\ \ 0.2\ \ 0.08)\begin{pmatrix}0.5\\0.3\\0.4\end{pmatrix}\cdot0.5=0.116$

第二轮 $(0.28\ \ 0.2\ \ 0.08)\cdot A\cdot b^T(O_2)=$

$\alpha_2(3)=(0.28\ \ 0.2\ \ 0.08)\begin{pmatrix}0.2\\0.5\\0.5\end{pmatrix}\cdot0.6=0.1176$

$\alpha_3(1)=\alpha_2(1)\cdot a_{11}\cdot b_1(O_3)+\alpha_2(2)\cdot a_{21}\cdot b_1(O_3)+\alpha_2(3)\cdot a_{31}\cdot b_1(O_3)=(0.0369\ \ 0.116\ \ 0.1176)\begin{pmatrix}0.3\\0.2\\0.1\end{pmatrix}\cdot0.7=0.03221$

$\alpha_3(2)=(0.0369\ \ 0.116\ \ 0.1176)\begin{pmatrix}0.5\\0.3\\0.4\end{pmatrix}\cdot0.5=0.050145$

第三轮 $(0.0369\ 0.116\ 0.1176)\cdot A\cdot\underbrace{b(O_1)}_{\begin{pmatrix}0.7\\0.5\\0.4\end{pmatrix}}=0.132038$

$\alpha_3(3)=(\qquad)\begin{pmatrix}0.2\\0.5\\0.5\end{pmatrix}\cdot0.4=0.049672$

“+”

$\alpha_3(1)+\alpha_3(2)+\alpha_3(3)=0.03221+0.050145+0.049672=0.132038$

试求最优序列

第一轮

$\delta_1(1) = 0.4 \times 0.7 = 0.28$

$\delta_1(2) = 0.4 \times 0.5 = 0.2$

$\delta_1(3) = 0.2 \times 0.4 = 0.08$

第一轮通过①拿红球概率最大

"对应"

第二轮

$\delta_2(1) = \max\{\underbrace{0.28 \times 0.3}_{\text{第一轮拿①，第2轮还拿①}}, \underbrace{0.2 \times 0.2}_{\text{第一轮②，第二轮①}}, 0.08 \times 0.1\} \times 0.7 = 0.28 \times 0.3 \times 0.7 = 0.0588$

第一轮确定能拿到红球之后（通过任何盒子）第2轮在①中拿黑球概率最大

$\psi_2(1) = 1$ 第一轮通过① 之后第2轮再通过①拿球的最大

$\delta_2(2) = \max\{\underbrace{0.28 \times 0.5}_{\text{第一轮通过①拿两红球，第二轮从②拿}}, 0.2 \times 0.3, 0.08 \times 0.4\} \times 0.5 = 0.28 \times 0.5 \times 0.5 = 0.07$

第2轮在②中拿黑球最大概率

$\psi_2(2) = 1$

$(\delta(1)\ \delta(2)\ \delta(3)) \cdot A \cdot \vec{b}(0) = (\ \ \ \ )$

$\delta_2(3) = \max\{0.28 \times 0.2, 0.2 \times 0.5, 0.08 \times 0.5\} \times 0.6 = 0.2 \times 0.5 \times 0.6 = 0.06$

$\psi_2(3) = 2$

第三轮

$\delta_3(1) = \max\{0.0588 \times 0.3, 0.07 \times 0.2, 0.06 \times 0.1\} \times 0.7 = 0.0588 \times 0.3 \times 0.7 = 0.012348$ ✓

$\delta_3(2) = \max\{\ \text{〃} \times 0.5, \text{〃} \times 0.3, \text{〃} \times 0.4\} \times 0.5 = 0.0588 \times 0.5 \times 0.5 = 0.0147$

$\delta_3(3) = \max\{\ \text{〃}\ 0.2\ \ 0.5\ \ 0.5\} \times 0.4 = 0.07 \times 0.5 \times 0.4 = 0.014$

$(\delta(1)\ \delta(2)\ \delta(3)) \cdot A \cdot \vec{b}(0) = (\ \ \ \ )$