作业一

1.最小错误率贝叶斯决策

- 已知条件:
  - $类先验概率p(\omega_i), i=1,2,...,c$
  - $类条件概率p(x|\omega_i), i=1,2,...,c$
- 求解任务:
  - 观测到样本x，将其分到哪一类错误率最小。
- 计算步骤:
  - 首先已知 $p(\omega_i|x)=\frac{p(x|\omega_i)p(\omega_i)}{p(x)}, i=1,2,...,c$
  - 然后找到最小的 $p(\omega_i|x)$，将x分类为 i
- 两类情形下的决策规则
  - 如果$p(x|\omega_1)p(\omega_1)>p(x|\omega_2)p(\omega_2)$ 将其分到类别1，否则分到类别2

2.最小风险贝叶斯决策

- 已知条件:
  - $类先验概率p(\omega_i), i=1,2,...,c$
  - $类条件概率p(x|\omega_i), i=1,2,...,c$
  - 损失函数：$\lambda(\alpha_i|\omega_j), 将类别j错误区分成i的损失，记为\lambda_{ij}$
- 求解任务:
  - 观测到样本x，将其分到哪一类损失最小。
- 计算步骤:
  - 首先已知 $R_i(x)=\sum_{j=1}^{c}\lambda_{ij}p(\omega_j|x)=\frac{1}{p(x)}\sum_{j=1}^{c}\lambda_{ij}p(x|\omega_j)p(\omega_j)$
  - 找到最小的 $R_i(x)$，将x分类为 i
- 两类情形下的决策规则:
  - $R_1(x)=A(\lambda_{11}p(x|\omega_1)p(\omega_1)+\lambda_{12}p(x|\omega_2)p(\omega_2))$
    $R_2(x)=A(\lambda_{21}p(x|\omega_1)p(\omega_1)+\lambda_{22}p(x|\omega_2)p(\omega_2))$
  - 如果$R_1(x)<R_2(x)$ 将其分到类别1否则分到类别2

3

- 已知条件:
  - $p(x|\omega_i)\sim N(\mu_i,\Sigma_i), i=1,2,...,c$

- $P(\omega_i), i=1,2,\ldots,c$
- 判别函数
  - $$g_i(x) = \ln(p(x|\omega_i) + \ln(P(\omega_i))) \\ = -\frac{1}{2}(x-\mu_i)^T\Sigma^{-1}(x-\mu_i) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln(|\Sigma_i|) + \ln(P(\omega_i)) \quad (i=1,2,\ldots,c)$$
- 最小距离分类器的情况
  - $P(\omega_i) = P(\omega_1) \quad i=1,2,\ldots,c$
  - $\Sigma_i = \sigma^2\mathrm{I} \quad \mathrm{i=1,2,\ldots,c}$
- 线性判别函数的情况
  - $\Sigma_i = \Sigma_1 \quad i=1,2,\ldots,c$

4.概率密度函数参数估计

- 已知条件:
  - $p(x|\omega_i)$具有确定的函数形式，只是某些参数θ未知
  - 独立同分布假设：每类样本都是从类条件概率密度$p(x|\omega_i)$中独立抽取出来的
  - 每类样本只包含本类的信息，不同类别的待估计参数是独立的。可以独立地分别处理c类问题。
- 求解任务:
  - 估计参数θ的取值。
- 计算步骤:
  - 对样本集$D=\{x_1,x_2,\ldots,x_n\}$。计算$H(\theta)=\ln\prod_{i=1}^{n}p(x|\theta)$，找到使$H(\theta)$最大的$\theta$

5.类条件概率密度函数估计

- 已知条件:
  - $p(x|\omega_i)$具有确定的函数形式，只是某些参数θ未知
  - 参数先验分布$p(\theta)$
  - 数据独立采样，对于样本集$D=\{x_1,x_2,\ldots,x_n\}$。$P(D|\theta)=p(x_1,x_2,\ldots,x_n|\theta)=\prod_{i=1}^{n}p(x_i|\theta)$
  - 每类样本只包含本类的信息，不同类别的待估计参数是独立的。可以独立地分别处理c类问题。
- 求解任务:
  - 估计类条件概率密度$p(x|\omega_i)$
- 计算步骤:
  - 先根据样本集和贝叶斯公式估计$p(\theta|D)=\frac{p(D|\theta)p(\theta)}{\int_\theta p(D|\theta)p(\theta)d\theta}=\frac{\prod_{i=1}^{n}p(x_i|\theta)p(\theta)}{\int_\theta\prod_{i=1}^{n}p(x_i|\theta)p(\theta)d\theta}=\alpha\prod_{i=1}^{n}p(x_i|\theta)p(\theta)$
  - 再根据关于θ的后验分布估计$p(x|D)=\int_\theta p(x,\theta|D)d\theta=\int_\theta p(x|\theta)p(\theta|D)d\theta$

6.不同之处

- 最大似然估计是将待估计的参数当作未知但固定的变量，其任务是根据观测数据估计其在参数空间的取值。
- 贝叶斯估计是将待估计的参数视为一个随机变量，其一个核心任务是根据观测数据对参数的分布进行估计