6.780 Notes

Lecturer:

Andrew Liu

Spring 2024

Last updated on Thursday 8th February, 2024.

Contents

1 February 8, 2024

Model family $\mathcal{H} \in \{H_0, H_1, \dots, H_{M-1}\}$. We can think of $\mathcal{H}$ as a set of class labels, and we want to determine the correct label given some test data.

1.1 Bayesian Binary Hypothesis Testing

In this case, $M = 2$, so there are only two hypotheses. Our model has two major components. The first is some a priori information

$$\begin{aligned} P_0 &= \mathbb{P}[H = H_0] \\ P_1 &= \mathbb{P}[H = H_1] = 1 - P_0. \end{aligned}$$

We also have the observation model, which is given by likelihood functions

$$\begin{aligned} H_0 &: p_{Y|H}(\cdot|H_0) \\ H_1 &: p_{Y|H}(\cdot|H_1). \end{aligned}$$

Bayes risk:

$$\varphi(f) \triangleq \mathbb{E}[C(H, f(y))]$$

The optimal decision rule takes form

$$\hat{H}(\cdot) = \arg\min_{f(\cdot)} \varphi(f).$$

$$\begin{aligned} \varphi(f) &= \mathbb{E}_{p_{y,H}}[C(H, f(y))] \\ &= \mathbb{E}_{p_y}[\mathbb{E}_{p_{H|y}}[C(H, f(y))|Y = y]] \\ &= \int p_Y(y)\mathbb{E}[C(H, f(y))|Y = y]\mathrm{d}y. \end{aligned}$$

Notice that we have control over the expected risk for each point, so to minimize $\varphi(f)$, we only have to solve a solution for individual points.

Given $y = y^*$, there are two possibilities; if $f(y^*) = H_0$, then

$$\mathbb{E}[C(H, f(y^*))|y = y^*] = C_{00}\mathbb{P}[H = H_0|y = y^*] + C_{01}\mathbb{P}[H = H_1|y = y^*],$$

otherwise

$$\mathbb{E}[C(H, f(y^*))|y = y^*] = C_{10}\mathbb{P}[H = H_0|y = y^*] + C_{11}\mathbb{P}[H = H_1|y = y^*].$$

For any y, we can compare both and choose as $\hat{H}$ the hypothesis that corresponds to the smaller one. Since

$$\mathbb{P}[H = H_i|Y = y] = \frac{p_{Y|H}(y|H_i)p_H(H_i)}{p_Y(y)},$$

we can substitute into the above expressions:

$$C_{00}p_{Y|H}(y|H_0)P_0 + C_{01}p_{Y|H}(y|H_1)P_1 \gtreqless C_{10}p_{Y|H}(y|H_0)P_0 + C_{11}p_{Y|H}(y|H_1)P_1$$

We can rewrite this expression in terms of the ratios

$$L(y) \triangleq \frac{p_{Y|H}(y|H_1)}{p_{Y|H}(y|H_0)} \gtreqless \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \triangleq \eta,$$

where we say that $L(y)$ is the **likelihood ratio**.

Theorem 1.1 (Likelihood Ratio Test)
Given a priori probabilities P_0, P_1, data y, observation models $p_{Y|H}(\cdot|H_0), p_{Y|H}(\cdot|H_1)$, and costs $C_{00}, C_{01}, C_{10}, C_{11}$, the Bayesian decision rule form

$$L(y) \triangleq \frac{p_{Y|H}(y|H_1)}{p_{Y|H}(y|H_0)} \underset{\hat{H}_0}{\overset{\hat{H}_1}{\gtreqless}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \triangleq \eta,$$

meaning that the decision is $\hat{H}(y) = H_1$ when $L(y) > \eta$, $\hat{H}(y) = H_0$ when $L(y) < \eta$, and it is indifferent when $L(y) = \eta$.

Note that the optimal rule is simple and deterministic.

1.2 Special Cases

In the case of “0-1 loss”, i.e., $C_{00} = C_{11} = 0$, $C_{01} = C_{10} = 1$, in which case our test simplifies to

$$p_{H|Y}(H_1|y) \gtrless p_{H|Y}(H_0|y).$$

This is the **maximum a posteriori** (MAP) decision rule.

If we additionally assume that $P_0 = P_1$, i.e., that our prior belief is indifferent, then our test further simplifies to

$$p_{Y|H}(y|H_1) \gtrless p_{Y|H}(y|H_0).$$

This is the **maximum likelihood** (MLE) decision rule.