ECE531 Screencast 9.3: Bayesian Detection with Finite Possible Observations

D. Richard Brown III

Worcester Polytechnic Institute

The Bayesian Approach

We assume a prior state distribution $\pi \in \mathcal{P}_N$ such that

$$\mathrm{Prob}(\text{state is } x_j) \;\; = \;\; \pi_j$$

Like Bayesian estimation, this prior reflects our belief of the state probabilities **prior to the observation**.

The Bayesian Approach

We assume a prior state distribution $\pi \in \mathcal{P}_N$ such that

$$\mathrm{Prob}(\text{state is } x_j) \quad = \quad \pi_j$$

Like Bayesian estimation, this prior reflects our belief of the state probabilities **prior to the observation**.

We denote the (scalar) Bayes Risk of the decision rule ρ as

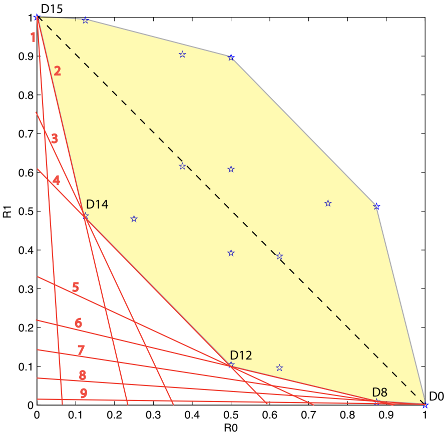
$$r(\rho, \pi) = \sum_{j=0}^{N-1} \pi_j R_j(\rho)$$

This is simply the weighted overall risk, or average risk, given our prior belief of the state probabilities. A decision rule that minimizes this risk is called a Bayes decision rule for the prior π.

Geometric Intuition

The prior π weights the conditional risks and establishes a family of "level sets". Given a constant $c \in \mathbb{R}$, the **level set** of value c is defined as

$$L_c^{\pi} \quad := \quad \{x \in R^N : \pi^\top x = c\}$$



Solving Bayesian HT Problems: Finite Observations

Given a decision matrix D, we can write the Bayes Risk as

$$
\begin{aligned}
r(D,\pi) &= \sum_{j=0}^{N-1} \pi_j R_j(D) = \sum_{j=0}^{N-1} \pi_j c_j^\top D p_j \\
&= \sum_{j=0}^{N-1} \pi_j \sum_{i=0}^{M-1} C_{ij} \sum_{\ell=0}^{L-1} D_{i\ell} P_{\ell j} \\
&= \sum_{\ell=0}^{L-1} \left(\sum_{i=0}^{M-1} D_{i\ell} \left[\sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j} \right] \right) \\
&= \sum_{\ell=0}^{L-1} \left(\sum_{i=0}^{M-1} D_{i\ell} G_{i\ell} \right) = \sum_{\ell=0}^{L-1} d_\ell^\top g_\ell
\end{aligned}
$$

Solving Bayesian HT Problems: Finite Observations

Given a decision matrix D, we can write the Bayes Risk as

$$
\begin{aligned}
r(D,\pi) &= \sum_{j=0}^{N-1} \pi_j R_j(D) = \sum_{j=0}^{N-1} \pi_j c_j^\top D p_j \\
&= \sum_{j=0}^{N-1} \pi_j \sum_{i=0}^{M-1} C_{ij} \sum_{\ell=0}^{L-1} D_{i\ell} P_{\ell j} \\
&= \sum_{\ell=0}^{L-1} \left(\sum_{i=0}^{M-1} D_{i\ell} \left[\sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j} \right] \right) \\
&= \sum_{\ell=0}^{L-1} \left(\sum_{i=0}^{M-1} D_{i\ell} G_{i\ell} \right) = \sum_{\ell=0}^{L-1} d_\ell^\top g_\ell
\end{aligned}
$$

Note that we can minimize each term in this sum separately.

Example Part 1

Let $G_{i\ell} := \sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j}$ and $G \in \mathbb{R}^{M \times L}$ be the matrix composed of elements $G_{i\ell}$. Suppose

$$G = \begin{bmatrix} 0.3 & 0.5 & 0.2 & 0.8 \\ 0.4 & 0.2 & 0.1 & 0.5 \\ 0.5 & 0.1 & 0.7 & 0.6 \end{bmatrix}$$

What decision rule minimizes the Bayes Risk?

Example Part 1

Let $G_{i\ell} := \sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j}$ and $G \in \mathbb{R}^{M \times L}$ be the matrix composed of elements $G_{i\ell}$. Suppose

$$G = \begin{bmatrix} 0.3 & 0.5 & 0.2 & 0.8 \\ 0.4 & 0.2 & 0.1 & 0.5 \\ 0.5 & 0.1 & 0.7 & 0.6 \end{bmatrix}$$

What decision rule minimizes the Bayes Risk?

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

This is a deterministic decision rule and it is unique. Note that the Bayes Risk of this decision rule is then simply

$$r(\pi, D) = 0.3 + 0.1 + 0.1 + 0.5 = 1.0.$$

Example Part 2

- What happens if

$$G = \begin{bmatrix} 0.3 & 0.5 & 0.2 & 0.8 \\ 0.4 & 0.2 & 0.2 & 0.5 \\ 0.5 & 0.1 & 0.7 & 0.6 \end{bmatrix}?$$

In this case, both

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \text{ or } D = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

achieve the same Bayes Risk $r(\pi, D) = 1.1$.

Example Part 2

- What happens if

$$G = \begin{bmatrix} 0.3 & 0.5 & 0.2 & 0.8 \\ 0.4 & 0.2 & 0.2 & 0.5 \\ 0.5 & 0.1 & 0.7 & 0.6 \end{bmatrix}?$$

In this case, both

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \text{ or } D = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

achieve the same Bayes Risk $r(\pi, D) = 1.1$.

- In fact, any decision matrix of the form

$$D = \begin{bmatrix} 1 & 0 & \alpha & 0 \\ 0 & 0 & 1-\alpha & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

for $\alpha \in [0, 1]$ will also achieve $r(\pi, D) = 1.1$.

Summary: Finite Observations

To minimize the Bayes Risk for finite $\mathcal{Y}$, we just find the index

$$m_\ell = \arg \min_{i \in \{0,\dots,M-1\}} \underbrace{\sum_{j=0}^{N-1} \pi_j C_{ij} P_{\ell j}}_{G_{i\ell}}$$

for each $\ell = 0, \dots, L-1$ and set $D^{B\pi}_{m_\ell,\ell} = 1$ and $D^{B\pi}_{i,\ell} = 0$ for all $i \neq m_\ell$.