❊ Lecture 17

17.1 Decision Theory

Example

An investor is deciding whether or not to purchase \$1000 of risky ZZZ bonds. If the investor buys the bonds, they can be redeemed at maturity for a net gain of \$500. There could, however, be a default on the bonds, in which case the original \$1000 investment would be lost. If the investor doesn't buy the bonds, she will put her money in a "safe" investment, for which she will be guaranteed a net gain of \$300 over the same time period. She estimates the probability of a default to be 0.1.

Solution.

$$\mathcal{A} = \{a_1, a_2\} = \{\text{buy ZZZ bonds}, \text{don't buy ZZZ bonds}\}$$

$$\Theta = \{\theta_1, \theta_2\} = \{\text{default}, \text{no default}\}$$

$$R(\theta, a_1(x)) = \mathbb{E}_{x|\theta} L(\theta, a_1(x)) = \int L(\theta, a_1(x)) f(x \mid \theta)\, dx = L(\theta, a_1) \int f(x \mid \theta)\, dx = L(\theta, a_1)$$

$$r(f, a_1) = \mathbb{E}_\theta R(\theta, a_1) = 150 \qquad r(f, a_2) = \mathbb{E}_\theta R(\theta, a_2) = 300$$

■

17.2 Minimax

We want to choose an action that minimizes the worst-case risk. The maximum Risk:

$$\bar{R}(a) = \sup_\theta R(\theta, a) = \max\{R(\theta_1, a_1), R(\theta_2, a_2)\}$$

Example ***Cont'd***

$$\bar{R}(a_1) = \sup_\theta R(\theta, a_1) = \max\{R(\theta_1 . a_1), R(\theta_2, a_1)\} = 1500$$

$$\bar{R}(a_2) = \sup_\theta R(\theta, a_2) = \max\{R(\theta_1, a_2), R(\theta_2, a_2)\} = 300$$

$$a_2 := \text{minimax}$$

In the estimation context, our possible actions are estimators $\hat{\theta}$. Then the maximum risk is

$$R(\hat{\theta}) = \sup_{\theta} R(\theta, \hat{\theta})$$

Example

$$X \sim N(\theta, 1) \qquad R(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2 \qquad \hat{\theta}_c(x) = cx$$

1. Find $R(\theta, \hat{\theta}_c)$.
2. Find minimax rule $\hat{\theta}_{c^*}$.
3. Let prior $\theta \sim N(a, b)$. Determine Bayes rule θ.

Solution.

$$R(\hat{\theta}_c) = \sup_{\theta} R(\theta, \hat{\theta}_c) = \sup_{\theta} (\theta - cx)^2 = \begin{cases} +\infty & c \neq 1 \\ 1 & c = 1 \end{cases}$$

For minimax rule, choose $c = 1$.

$$f(\theta \mid x) \propto \exp\{-\frac{(x-\theta)^2}{2}\} \cdot \exp\{-\frac{(\theta - a)^2}{2b}\} \propto \exp\{-\frac{1+b}{2b}\left(\theta - \frac{bx+a}{1+b}\right)^2\}$$

Bayes rule:

$$\implies \mathbb{E}_{\theta|x}[\theta] = \frac{b}{1+b}x + \frac{a}{1+b} \neq \hat{\theta}_{c^*} \text{ which is of the form } cx$$

But the above is not in the form cx. So we compute the Bayes risk:

$$r(f, \hat{\theta}_c) = \mathbb{E}_\theta R(\theta, \hat{\theta}_c) = \mathbb{E}_\theta[(c-1)^2\theta^2 + c^2] = (a^2+b+1)(c - \frac{a^2+b}{a^2+b+1})^2 + a^2 + b - \frac{(a^2+b)^2}{a^2+b+1}$$

$$c = (a^2+b)/(a^2+b+1), \qquad \hat{\theta}_c \text{ minimizes } r(f, \hat{\theta}_c)$$

■

17.3 Geometry of Bayes and Minimax Rules for Finite Ω

Given a finite parameter space $\Omega = \{\theta_1, \cdots, \theta_k\}$, we define the risk set as $S \subseteq \mathbb{R}^k$ such that

$$S = \{(y_1, \cdots, y_k) : y_i = R(\theta_i, \delta) \text{ for } \delta \in \mathcal{A}\}.$$

We can visualize S in $\mathbb{R}^k$. Each decision rule δ corresponds to a point in S. The goal of decision theory is to find optimal points in S.

And by allowing randomized estimators, we can form convex combinations of points in S.

Lemma. The risk set S is always convex when $\mathcal{A}$ has randomized estimators.

In this setting, a prior of θ can be considered as a finite vector

$$\lambda(\theta) = (\lambda_1, \cdots, \lambda_k) = (\lambda(\theta_1), \cdots, \lambda(\theta_k)),$$

with $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda \geq 0$. The Bayes risk is

$$r(\lambda, \delta) = \sum_{i=1}^{k} \lambda_i R(\theta_i, \delta) = (\lambda_1, \cdots, \lambda_k) \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}.$$

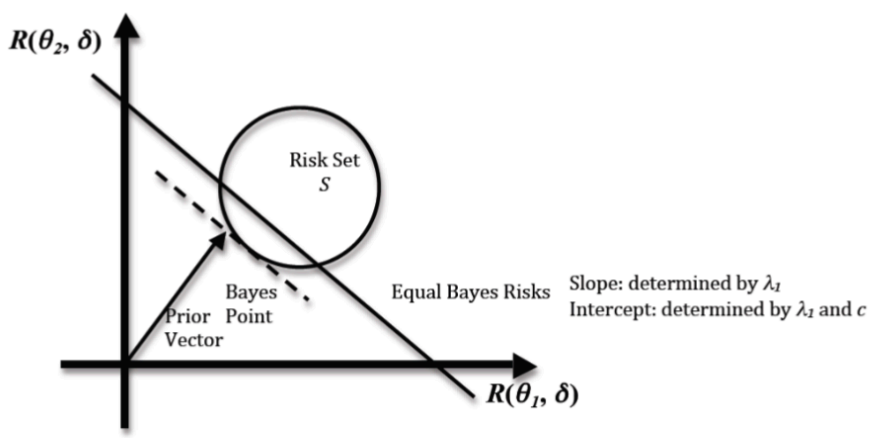

Figure 4: Geometry of a Bayes Point for $k = 2$

The tangent line with slope $-\lambda_1/\lambda_2$ corresponds to the Bayes rule with prior $\lambda = (\lambda_1, \lambda_2)$.

$$\lambda_1 \cdot R(\theta_1, \delta) + \lambda_2 \cdot R(\theta_2, \delta) = c$$

References

[1] Larry Wasserman *All of Statistics*. Section 2 & 3

[2] Morris H. DeGroot *Probability and Statistics*.