Lecture Notes 1 – Bayes Decision Theory

Probabilities and Measurements

- State of nature $\omega_1, \omega_2, \ldots$
- Prior probability $P(\omega_1), P(\omega_2)$
- Measurements: $\boldsymbol{x} \in \mathbb{R}^d$
- Likelihood function: $p(\boldsymbol{x} \mid \omega_j)$

Bayes Theorem

- Bayes rule: $P(\omega_j \mid \boldsymbol{x}) = \dfrac{P(\boldsymbol{x} \mid \omega_j)P(\omega_j)}{p(\boldsymbol{x})} = \dfrac{\text{likelihood} \times \text{prior}}{\text{evidence}}$

Classification Error

- $P(\text{error} \mid \boldsymbol{x}) = \begin{cases} P(\omega_1 \mid \boldsymbol{x}) & \text{if we decide } \omega_2 \\ P(\omega_2 \mid \boldsymbol{x}) & \text{if we decide } \omega_1 \end{cases}$
- Error rate: $R = \displaystyle\int_{\boldsymbol{x}} P(\text{error} \mid \boldsymbol{x}) \cdot p(\boldsymbol{x})\, d\boldsymbol{x}$

Optimal Classifiers

- Bayes decision rule: $\begin{cases} \omega_1 & \text{if} \;\; P(\omega_1 \mid \boldsymbol{x}) > P(\omega_2 \mid \boldsymbol{x}) \\ \omega_2 & \text{else} \end{cases}$
- Bayes error rate: $R = \displaystyle\int_{\boldsymbol{x}} \min\{P(\omega_1 \mid \boldsymbol{x}), P(\omega_2 \mid \boldsymbol{x})\} \cdot p(\boldsymbol{x})\, d\boldsymbol{x}$

Multiclass

- Bayes decision rule: $\{\; \omega_i \quad \text{if} \;\; \forall_{j \neq i} : \; P(\omega_i \mid \boldsymbol{x}) > P(\omega_j \mid \boldsymbol{x})$
- Bayes error rate: $R = \displaystyle\int_{\boldsymbol{x}} \min_i \; (1 - P(\omega_i \mid \boldsymbol{x})) \cdot p(\boldsymbol{x})\, d\boldsymbol{x}$

Discriminant functions

- Bayes decision rule: $\begin{cases} \omega_1 & \text{if} \;\; \frac{P(\boldsymbol{x}|\omega_1)}{P(\boldsymbol{x}|\omega_2)} > \theta \\ \omega_2 & \text{else} \end{cases}$

  with $\theta = P(\omega_2)/P(\omega_1)$

Discriminant functions (multiclass)

- Bayes decision rule: $\{\; \omega_i \quad \text{if} \;\; \forall_{j \neq i} : \; g_i(\boldsymbol{x}) > g_j(\boldsymbol{x})$

  with $g_i(\boldsymbol{x}) = f_{\text{incr}}(P(\omega_i \mid \boldsymbol{x}))$ and f_{incr} a monotonically increasing function.

Actions and Cost

- Actions $\alpha_1, \dots, \alpha_a$
- Loss $\lambda(\alpha_i \mid \omega_j)$ caused by taking action α_i given state of nature ω_j.
- Expected cost of an action: $R(\alpha_i \mid \boldsymbol{x}) = \displaystyle\sum_{j=1}^{c} \lambda(\alpha_i \mid \omega_j)\, P(\omega_j \mid \boldsymbol{x})$.
- Risk of optimal decider: $R = \displaystyle\int_{\boldsymbol{x}} \min_i R(\alpha_i \mid \boldsymbol{x}) \cdot p(\boldsymbol{x})\, d\boldsymbol{x}$

Practical likelihood functions

Multivariate Density ($\boldsymbol{x} \in \mathbb{R}^d$)

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]$$

Discrete Density ($\boldsymbol{x} \in \{0,1\}^d$)

$$p(\boldsymbol{x}) = \prod_{i=1}^{d} \underbrace{\Pr[x_i = 0]}_{1-p_i} \cdot 1_{x_i=0} + \underbrace{\Pr[x_i = 1]}_{p_i} \cdot 1_{x_i=1} = \prod_{i=1}^{d} p_i^{x_i}(1-p_i)^{1-x_i}$$
