Lecture 8, Feb 2, 2024

Bayesian Hypothesis Testing

- We switch to using a MAP approach instead of ML if we have prior knowledge of the hypotheses
- Assume we know a priori that $P[H_0] = \pi_0$ and $P[H_1] = 1 - \pi_0 = \pi 1$
- MAP rule: compare $p(H_1|x)$ with $p(H_0|x)$
  - $P(H_j|x) = \dfrac{p_x(x|H_j)P[H_j]}{p_x(x|H_0)P[H_0] + p_x(x\ H_1)P[H_1]}$
  - Therefore this is equivalent to $p(x|H_1)\pi_1 \underset{H_0}{\overset{H_1}{\lessgtr}} p(x|H_0)\pi_0$
- Example: binary communications
  - If a 0 was sent then $X \sim \mathcal{N}(-1, \sigma^2)$; if a 1 was sent then $X \sim \mathcal{N}(1, \sigma^2)$
  - Compare $\dfrac{1}{\sqrt{2\pi}\sigma} e^{-(x-1)^2} 2\sigma^2 \pi_1$ and $\dfrac{1}{\sqrt{2\pi}\sigma} e^{-(x+1)^2} 2\sigma^2 \pi_0$
  - The threshold is $e^{\frac{4x}{2\sigma^2}} \lessgtr \dfrac{\pi_0}{\pi 1}$ or $x \lessgtr \dfrac{\sigma^2}{2} \log \dfrac{\pi_0}{\pi_1}$
  - If we assume that the bits are balanced, i.e. $\pi_0 = \pi_1$, then we're simply seeing if x is positive or negative
- Assign different costs, C_{ij} being the cost of accepting H_i if H_j is true
  - The expected cost is $C_{00}P[\text{Accept } H_0|H_0]\pi_0 + C_{01}P[\text{Reject } H_0|H_0]\pi_0 + C_{10}P[\text{Accept } H_1|H_1]\pi_1 + C_{11}P[\text{Reject } H_1|H_1]\pi_1$
  - Define the likelihood $\Lambda(x) = \dfrac{f_X(x|H_1)}{f_X(x|H_0)}$
  - Optimal decision rule: accept H_0 if $\Lambda(x) < \dfrac{\pi_0(C_{01} - C_{00})}{\pi_1(C_{10} - C_{11})}$
    * We can derive this using the same procedure as above
    * A special case is to minimize the probability of error in which case $C_{00} = C_{11} = 0, C_{01} = C_{10} = 1$
    * Equivalent to minimizing $P_E = \int_R f_X(x|H_0)\pi_0 \,\mathrm{d}x + \int_{R^c} f_X(x|H_1)\pi_1 \,\mathrm{d}x$
- Consider the Gaussian example with minimum error probability as the cost
  - $P_E = \int_R \dfrac{1}{\sqrt{2\pi}\sigma} e_0^{\frac{-(x+1)^2 2\sigma^2}{\pi}} \,\mathrm{d}x + \int_{R^c} \dfrac{1}{\sqrt{2\pi}\sigma} e_1^{\frac{-(x-1)^2 2\sigma^2}{\pi}} \,\mathrm{d}x = 1 + \int_{R^c} \dfrac{1}{\sqrt{2\pi}\sigma} \left(e^{-\frac{(x-1)^2}{2\sigma^2}} \pi_1 - e^{-\frac{(x+1)^2}{2\sigma^2}} \pi_0 \right) \mathrm{d}x$
  - To minimize this we want to pick the expression inside brackets so it is always negative
  - Making the expression negative gives us back the MAP rule!
- For multiple hypotheses:
  - Probability of correct is $\sum_{k=1}^{K} \int_{R_k} P[X = x, H_k] \,\mathrm{d}x$
  - Min cost: $\sum_j \int_{R_j} \sum_{k=1}^{K} L_{kj} P[X = x, H_k] \,\mathrm{d}x$
    * We need a cost L_{kj} between every pair of hypotheses
    * Pick j to minimize the sum inside the integral
- Naive Bayes assumption: all measurements are independent given θ

Significance Testing

- Sometimes we know what H_0 is but we don't have a clear alternative hypothesis
- We want to bound α, the probability of type I error
- Example: testing if a unit variance Gaussian is zero mean
  - Let $S = \dfrac{1}{\sqrt{n}}(X_1, \ldots, X_n)$
  - Then if H_0 holds then $S \sim \mathcal{N}(0, 1)$
  - Suppose our decision rule is to accept H_0 if $S \in [-\gamma, \gamma]$

- $\alpha = \int_{-\infty}^{\gamma} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, \mathrm{d}x + \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, \mathrm{d}x = 2Q(\gamma)$
- If we want to restrict the type I error probability to 5%, then $Q(\gamma) = 0.05 \implies \gamma = 1.96$