Machine Learning 1 - Exercise 1

Fränz Beckius (374057)
Ivan David Aranzales Acero (399364)
Janek Tichy (584200)
Jeremias Eichelbaum (358685)
Johannes Krause (395469)

October 22, 2018

1 Estimate the Bayes Error

1.a

We need to show that the P(error) is upper bounded by the given term.

$$P(error) \leq \int \frac{2}{\frac{1}{P(w_1 \mid x)} + \frac{1}{P(w_2 \mid x)}} p(x)dx \tag{1}$$

$$P(error) \leq \int \frac{2P(w_1 \mid x)P(w_2 \mid x)}{P(w_1 \mid x) + P(w_2 \mid x)} p(x)dx \tag{2}$$

With given equation 3 we can substitute in equation 4.

$$P(error) = \int P(err \mid x)p(x)dx \tag{3}$$

$$\int P(err \mid x)p(x)dx \leq \int \frac{2P(w_1 \mid x)P(w_2 \mid x)}{P(w_1 \mid x) + P(w_2 \mid x)} p(x)dx \tag{4}$$

Both sides of inequation integrate over same range and variable, hence we can transform to inequation 5.

$$P(err \mid x) \leq \frac{2P(w_1 \mid x)P(w_2 \mid x)}{P(w_1 \mid x) + P(w_2 \mid x)} \tag{5}$$

With given equation 6 we can substitute in inequation 7.

$$P(err \mid x) = min[P(w_1 \mid x), P(w_2 \mid x)] \tag{6}$$

$$min[P(w_1 \mid x), P(w_2 \mid x)] \leq \frac{2P(w_1 \mid x)P(w_2 \mid x)}{P(w_1 \mid x) + P(w_2 \mid x)} \tag{7}$$

Since given $P(w_1 \mid x) + P(w_2 \mid x) = 1$ and $P(w_2 \mid x) = 1 - P(w_1 \mid x)$.

$$min[P(w_1 \mid x), P(w_2 \mid x)] \leq 2P(w_1 \mid x)P(w_2 \mid x) \tag{8}$$

Assume the inequation 8

$$min[P(w_1 \mid x), P(w_2 \mid x)] > 2P(w_1 \mid x)P(w_2 \mid x) \tag{9}$$

With $P(w_1 \mid x) = P(w_2 \mid x) = 0.5$ we get a contradiction, hence we showed inequation 8.

1.b

We need to show that Bayes error can be upper-bounded. Starting by given inequation (a) and substitute with Bayes formula.

$$P(error) \leq \int \frac{2}{\frac{1}{p(x \mid w_1)P(w_1)} + \frac{1}{p(x \mid w_2)P(w_2)}} dx \tag{10}$$

Now we can substitute with the univariate probability distributions.

$$\leq \int \frac{2}{\frac{1+x^2-2x\mu+\mu^2}{\pi^{-1}P(w_1)} + \frac{1+x^2-2x\mu+\mu^2}{\pi^{-1}P(w_2)}} dx \tag{11}$$

$$\leq \int \frac{1}{\frac{1}{2}[\frac{x^2}{\pi^{-1}P(w_1)} + \frac{x^2}{\pi^{-1}P(w_2)} + \frac{-2x\mu}{\pi^{-1}P(w_1)} + \frac{2x\mu}{\pi^{-1}P(w_2)} + \frac{1+\mu^2}{\pi^{-1}P(w_1)} + \frac{1+\mu^2}{\pi^{-1}P(w_2)}]} dx \tag{12}$$

$$\leq \int \frac{1}{\frac{\pi(P(w_1)+P(w_2))}{2P(w_1)P(w_2)}x^2 + \frac{\mu\pi(P(w_1)-P(w_2))}{P(w_1)P(w_2)}x + \frac{\pi(1+\mu^2)(P(w_1)+P(w_2))}{2P(w_1)P(w_2)}} dx \tag{13}$$

Using the given hint we can transform inequation 12.

$$\leq \frac{2\pi}{\sqrt{\frac{\pi^2(P(w_1)+P(w_2))^2(1+\mu^2)}{P(w_1)^2P(w_2)^2} - \frac{\mu^2\pi^2(P(w_1)-P(w_2))^2}{P(w_1)^2P(w_2)^2}}} dx \tag{14}$$

$$= \frac{2P(w_1)P(w_2)}{\sqrt{(1+\mu^2)[P(w_1)^2 + 2P(w_1)P(w_2) + P(w_2)^2] - \mu^2[P(w_1)^2 - 2P(w_1)P(w_2) + P(w_2)^2]}} dx \tag{15}$$

$$= \frac{2P(w_1)P(w_2)}{\sqrt{(P(w_1) + P(w_2))^2 + 4\mu^2 P(w_1)P(w_2)}} dx \tag{16}$$

Since $P(w_1) + P(w_2)$ reduces to equal 1 we get the equation 16.

$$= \frac{2P(w_1)P(w_2)}{\sqrt{1 + 4\mu^2 P(w_1)P(w_2)}} dx \tag{17}$$

1.c

2 Bayes Decision Boundaries

2.a

We define a function g(x) that classifies x as w1 if g(x) > 0 otherwise w2

$$g(x) = P(w_1 \mid x) - P(w_2 \mid x) \tag{18}$$

$P(w_i \mid x)$ can also be rewritten as

$$\begin{aligned} P(w_i \mid x) &= \frac{p(x \mid w_i)P(w_i)}{\sum_{j=i}^{c} p(x \mid w_j)P(w_j)} \\ &= p(x \mid w_i)P(w_i) \\ &= ln(p(x \mid w_i)) + ln(P(w_i)) \end{aligned} \tag{19}$$

Therefore we can rewrite g(x) as

$$\begin{aligned} g(x) &= P(w_1 \mid x) - P(w_2 \mid x) \\ &= ln(\frac{p(x \mid w_1)}{p(x \mid w_2)}) + ln(\frac{P(w_1)}{P(w_2)}) \\ &= ln(\frac{1}{2\sigma} exp(\frac{-|x-\mu|}{\sigma})) - ln(\frac{1}{2\sigma} exp(\frac{-|x+\mu|}{\sigma})) + ln(\frac{P(w_1)}{P(w_2)}) \\ &= \frac{-|x-\mu| + |x+\mu|}{\sigma} + ln(\frac{P(w_1)}{P(w_2)}) \end{aligned} \tag{20}$$

2.b

2.c

Similar to 2.a, we define a function g(x) that classifies x as w1 if g(x) > 0 otherwise w2

$$\begin{aligned} g(x) &= P(w_1 \mid x) - P(w_2 \mid x) \\ &= ln(\frac{p(x \mid w_1)}{p(x \mid w_2)}) + ln(\frac{P(w_1)}{P(w_2)}) \\ &= ln(\frac{1}{2\sigma} exp(\frac{-(x-\mu)^2}{2\sigma^2})) - ln(\frac{1}{2\sigma} exp(\frac{-(x+\mu)^2}{2\sigma^2})) + ln(\frac{P(w_1)}{P(w_2)}) \\ &= \frac{-(x-\mu)^2 + (x+\mu)^2}{2\sigma^2} + ln(\frac{P(w_1)}{P(w_2)}) \\ &= \frac{-(x^2 - 2x\mu + \mu^2) + (x^2 + 2x\mu + \mu^2)}{2\sigma^2} + ln(\frac{P(w_1)}{P(w_2)}) \\ &= \frac{2x\mu}{\sigma^2} + ln(\frac{P(w_1)}{P(w_2)}) \end{aligned} \tag{21}$$