Elements of Machine Learning

Assignment 1

Xingrong Zong, tnd179

February 14th, 2022

Exercise 8.3

Show that $p(a,b) \neq p(a)p(b)$

$p(a=0) = \sum_b \sum_c p(a=0,b,c) = 0.192 + 0.144 + 0.048 + 0.216 = 0.6$

$p(b=0) = \sum_a \sum_c p(a,b=0,c) = 0.192 + 0.144 + 0.192 + 0.064 = 0.592$

$p(a=0,b=0) = \sum_c p(a=1,b=1,c) = 0.192 + 0.144 = 0.336$

$p(a=0)p(b=0) = 0.6 \cdot 0.592 = 0.3552$

$p(a=0,b=0) \neq p(a=0)p(b=0)$

$p(a,b) \neq p(a)p(b)$

Show that $p(a,b|c) = p(a|c)p(b|c)$

$p(c=0) = 0.192 + 0.048 + 0.192 + 0.048 = 0.48$

$p(a=0|c=0) = \frac{p(a=0,c=0)}{p(c=0)} = \frac{\sum_b p(a=0,b,c=0)}{p(c=0)} = \frac{0.192+0.048}{0.48} = 0.5$

$p(a=1|c=0) = 1 - p(a=0|c=0) = 0.5$

$p(b=0|c=0) = \frac{p(b=0,c=0)}{p(c=0)} = \frac{\sum_a p(a,b=0,c=0)}{p(c=0)} = \frac{0.192+0.192}{0.48} = 0.8$

$p(b=1|c=0) = 1 - p(b=0|c=0) = 0.2$

$p(a,b|c=0) = \frac{p(a,b,c=0)}{p(c=0)}$

$p(a=0,b=0|c=0) = \frac{p(a=0,b=0,c=0)}{p(c=0)} = \frac{0.192}{0.48} = 0.4$

$p(a=0,b=1|c=0) = \frac{p(a=0,b=1,c=0)}{p(c=0)} = \frac{0.048}{0.48} = 0.1$

$p(a=1,b=0|c=0) = \frac{p(a=1,b=0,c=0)}{p(c=0)} = \frac{0.192}{0.48} = 0.4$
$p(a=1,b=1|c=0) = \frac{p(a=1,b=1,c=0)}{p(c=0)} = \frac{0.048}{0.48} = 0.1$

$p(a=0|c=0)p(b=0|c=0) = 0.5 \cdot 0.8 = 0.4 = p(a=0,b=0|c=0)$

$p(a=0|c=0)p(b=1|c=0) = 0.5 \cdot 0.2 = 0.1 = p(a=0,b=1|c=0)$

$p(a=1|c=0)p(b=0|c=0) = 0.5 \cdot 0.8 = 0.4 = p(a=1,b=0|c=0)$

$p(a=1|c=0)p(b=1|c=0) = 0.5 \cdot 0.2 = 0.1 = p(a=1,b=1|c=0)$

$p(a,b|c) = p(a|c)p(b|c)$

Exercise 8.4

$p(a=0) = 0.192 + 0.144 + 0.048 + 0.216 = 0.6$

$p(a=1) = 0.192 + 0.064 + 0.048 + 0.096 = 0.4$

$p(b|c) = \frac{p(b,c))}{p(c)}$

$p(c=0) = 0.192 + 0.048 + 0.192 + 0.048 = 0.48$

$p(c=1) = 0.144 + 0.216 + 0.064 + 0.096 = 0.52$

$p(b=0|c=0) = \frac{p(b=0,c=0)}{p(c=0)} = \frac{0.192+0.192}{0.48} = 0.8$

$p(b=0|c=1) = \frac{p(b=0,c=1)}{p(c=0)} = \frac{0.144+0.064}{0.52} = 0.4$

$p(b=1|c=0) = \frac{p(b=1,c=0)}{p(c=0)} = \frac{0.048+0.048}{0.48} = 0.2$

$p(b=1|c=1) = \frac{p(b=1,c=1)}{p(c=0)} = \frac{0.216+0.096}{0.52} = 0.6$

$p(c|a) = \frac{p(a,c)}{p(a)}$ $p(a=0) = 0.192 + 0.144 + 0.048 + 0.216 = 0.6$

$p(a=1) = 0.192 + 0.064 + 0.048 + 0.096 = 0.4$

$p(c=0|a=0) = \frac{p(c=0,a=0)}{p(a=0)} = \frac{0.192+0.048}{0.6} = 0.4$

$p(c=0|a=1) = \frac{p(c=0,a=1)}{p(a=1)} = \frac{0.192+0.048}{0.4} = 0.6$

$p(c=1|a=0) = \frac{p(c=1,a=0)}{p(a=0)} = \frac{0.144+0.216}{0.6} = 0.6$

$p(c=1|a=1) = \frac{p(c=1,a=1)}{p(a=1)} = \frac{0.064+0.096}{0.4} = 0.4$

$p(a=0)p(c=0|a=0)p(b=0|c=0) = 0.6 \cdot 0.4 \cdot 0.8 = 0.192 = p(a =$

$0, b = 0, c = 0)$

$p(a = 0)p(c = 1|a = 0)p(b = 0|c = 1) = 0.6 \cdot 0.6 \cdot 0.4 = 0.144 = p(a = 0, b = 0, c = 1)$

$p(a = 0)p(c = 0|a = 0)p(b = 1|c = 0) = 0.6 \cdot 0.4 \cdot 0.2 = 0.048 = p(a = 0, b = 1, c = 0)$

$p(a = 0)p(c = 0|a = 0)p(b = 1|c = 1) = 0.6 \cdot 0.4 \cdot 0.6 = 0.216 = p(a = 0, b = 1, c = 1)$

$p(a = 1)p(c = 0|a = 1)p(b = 0|c = 0) = 0.4 \cdot 0.6 \cdot 0.8 = 0.192 = p(a = 1, b = 0, c = 0)$

$p(a = 1)p(c = 1|a = 1)p(b = 0|c = 1) = 0.4 \cdot 0.4 \cdot 0.4 = 0.064 = p(a = 1, b = 0, c = 1)$

$p(a = 1)p(c = 0|a = 1)p(b = 1|c = 0) = 0.4 \cdot 0.6 \cdot 0.2 = 0.048 = p(a = 1, b = 1, c = 0)$

$p(a = 1)p(c = 1|a = 1)p(b = 1|c = 1) = 0.4 \cdot 0.4 \cdot 0.6 = 0.096 = p(a = 1, b = 1, c = 1)$

The directed graph looks like $a \to c \to b$

Exercise 8.10

$p(a, b, c, d) = p(a)p(b)p(c|a, b)p(d|c)$

$p(a, b) = p(a)p(b) \sum_c \sum_d p(c|a, b)p(d|c)$

$= p(a)p(b) \sum_c p(c|a, b)[\sum_d p(d|c)]$

$= p(a)p(b) \sum_c p(c|a, b) \cdot 1$

$= p(a)p(b) \cdot 1$

$= p(a)p(b)$

Prove a and b are dependent conditioned on d: $p(a, b|d) = p(a|d)p(b|d)$

Multiply both sides by p(d): $p(a, b, d) = p(a)p(b|d)$

Since we have, $p(a, b, c, d) = p(a)p(b)p(c|a, b)p(d|c)$

Summation on both sides with respect to c: $p(a,b,d) = p(a)p(b)\sum_c p(c|a,b)p(d|c)$
So, $p(a,b,d) = p(a)p(b|d) = p(a)p(b)\sum_c p(c|a,b)p(d|c)$
$p(b|d) = p(b)\sum_c p(c|a,b)p(d|c)$, the left hand side depends on b and d, the right hand side depends on a,b and d. Therefore, a and b are not dependent conditioned on d.

Exercise 9.5

Since, $\mu, \sum, \pi$ can be omitted, we can write, $p(X,Z) = p(x_1,z_1)p(z_1)\cdots p(x_N,z_N)p(z_N) = p(x_1,z_1)\cdots p(x_n,z_n)$
Also, there is no link from z_m to z_n, from x_m to x_n, and from z_m to x_n $(m \neq n)$, so, $p(Z) = p(z_1)\cdots p(z_N)$, $p(X) = p(x_1)\cdots p(x_N)$ According to Bayes' Theorem, $p(Z|X) = \frac{p(X|Z)p(Z)}{p(X)}$
$= \frac{\prod_{n=1}^{N} p(x_n|z_n) \cdot \prod_{n=1}^{N} p(z_n)}{\prod_{n=1}^{N} p(x_n)}$
$= \prod_{n=1}^{N} \frac{p(x_n|z_N)p(z_n)}{p(x_n)}$
$= \prod_{n=1}^{N} p(z_n|x_n)$
In the directed graph, there are only links from z_n to x_n, and we assume the x_n is independent and identically distributed and thus there is no link from x_m to x_n. $p(Z|X,\mu,\sum,\pi) = \prod_{n=1}^{N} p(Z_n|X_n,\mu,\sum,\pi)$

Old Faithful

Exercise 1

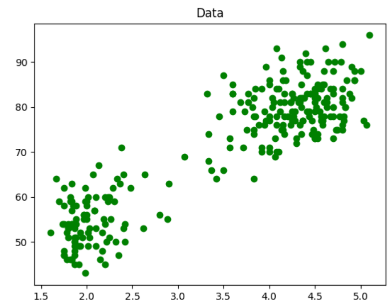

Figure 1: Data

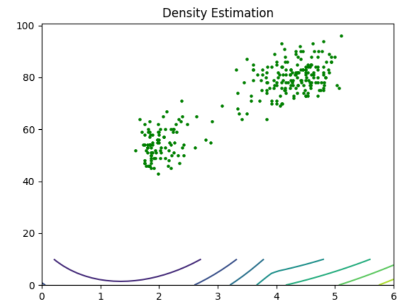

Figure 2: Density Estimation

Exercise 2

$eruptions_{mean} = 3.4877830882352936$

$waiting_{mean} = 70.8970588235294$

$eruptions_{var} = 1.2979388904492861$

$waiting_{var} = 184.14381487889273$

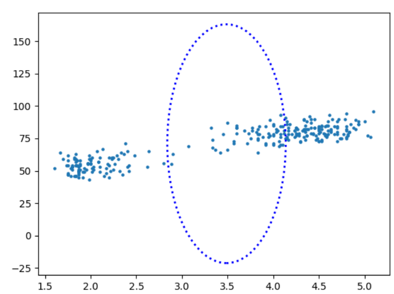

Figure 3: EM

Exercise 3

Exercise 4

If use more than two features, it does not give any results.