CS4100/5100: Fundamental Artificial Intelligence — Robert Platt
Siyong Ma — Dec. 8, 2014

Solution for Homework 3

1 Problems

1. **Problem 1**.
   Answer:
   a) $K(x,z) = (x_1^2,\ x_1x_2,\ x_2^2) \cdot (z_1^2,\ z_1z_2,\ z_2^2)^T = x_1^2z_1^2 + x_1x_2z_1z_2 + x_2^2z_2^2$
   b) $F(x) = (x, x_1x_2)$

2. **Problem 2**.
   Answer:
   a)
   $P_{C|D} = \underset{C}{normalize}\ P_{CD}$
   $P_{CD} = \sum_{A,B} P_{ABCD}$
   $P_{ABCD} = P_A P_{B|A} P_{C|B} P_{D|B}$
   $P_{C|D} = \underset{C}{normalize}\ \sum_{A,B} P_A P_{B|A} P_{C|B} P_{D|B}$

   b)
   A, B will be eliminated.
   2 ways. first A then B, or in reverse order.
   A, B:
   size of the largest factor $2^3 = 8$, since 3 free binary variables BCD in P_{BCD} before SUM
   $\sum_B[(\sum_A P_A P_{B|A}) P_{C|B} P_{D|B}] = \sum_B P_B P_{C|B} P_{D|B} = \sum_B P_{BCD} = P_{CD}$
   B, A:
   size of the largest factor $2^4 = 16$, since 4 free binary variables ABCD in $P_{BCD|A}$ before SUM
   $\sum_A[P_A(\sum_B P_{B|A} P_{C|B} P_{D|B})] = \sum_A P_A \sum_B P_{BCD|A} = \sum_A P_A P_{CD|A} = \sum_A P_{ACD} = P_{CD}$

3. **Problem 3**.
   Answer:
   a)
   $sign((2,0,1,1) \cdot (1,1,0,-1)) = sign(1) = +1$
   $sign((-1,0,1,1) \cdot (1,1,0,-1)) = sign(-2) = -1$
   b)
   $W_{new} = W_{old} + \alpha(desired\_output - prediction) \cdot input\_vector$
   So new weight vector will be $(1,1,0,-1) + 0.1 * (-1-1) * (0,0,0,-1) = (1,1,0,-0.8)$

4. **Problem 4.**
Answer:
1. From the problem we know that all training examples are the same, the only difference it the label. When we do the training, our object is to minimize the loss function which in this case is the squared error or error rate function. logistic regression Predictor is deterministic, that is, given an input sample, the prediction is deterministic not by random guess. In another word, the final prediction will always be either 0 or 1 for all training examples, since they are the same. Predicting the training example as 1 will minimize the squared error, since the error rate in this case is 20/100 while predicting it as 0, the error rate is 80/100.

5. **Problem 5.**
Answer:
a)
#Occurrences of Y=0: 5
#Occurrences of $F_1 = 1$ when Y=0: 3
So $P(F_1 = 1|Y = 0) = \frac{3}{5}$

b)
Adding 1 occurrence for $F_1 = 1$ when Y=0, and 1 occurrence for $F_1 = 0$ when Y=0
#Occurrences of Y=0: 5+2
#Occurrences of $F_1 = 1$ when Y=0: 3+1
So $P(F_1 = 1|Y = 0) = \frac{4}{7}$