AML Assignment 1 **Due Date: August 19th, 2019** **Total Marks: 50**

Instructions: You are allowed to discuss but the final answer should be your own. Any instance of cheating will be considered as academic dishonesty and penalty will be applied.

Ans 1 [6 Marks]. (a) The rank is 2 and its inverse is $[-1, 2; 1, -1]$
(b) The rank is 3 and its inverse is $[-1/2,\ 3, -1; 3/2, -4, 2; 1, -2, 1]$

(c) The rank is 4 and its inverse is $[-51, 15, 7, 12; 31, -9, -4, -7; -10, 3, 1, 2; -3, 1, 1, 1]$

Ans 2 [4 Marks]. Observations: $x_1,\ x_2,\ x_3,\ x_4,\ x_5$
Mean = $(x_1 + x_2 + x_3 + x_4 + x_5)/5 = 3$
$x_1^2 + x_2^2 + x_3^2 = 14$ (1)
$x_3^2 + x_4^2 + x_5^2 = 50$ (2)
$x_1^2 + x_2^2 = 5$ (3)

Sample standard deviation = $\sqrt{\frac{1}{n-1}\ [x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 2.3.(x_1 + x_2 + x_3 + x_4 + x_5) + 5.9]}$

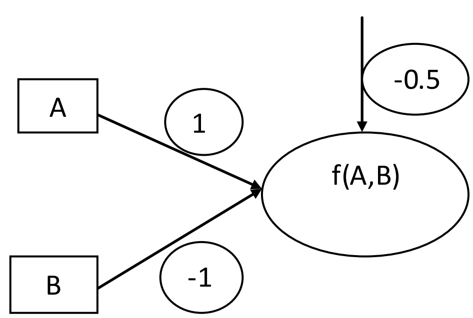
$$= \sqrt{\frac{1}{4}\ [5 + 50 - 2.3.(5.3) + 5.9]}$$

$$= \sqrt{\frac{5}{2}} \qquad = 1.581$$

Ans 3 [10 Marks]. (a) True

(b) True

(c) Yes, you can represent this function with a single logistic threshold unit, since it is linearly separable. Here is one example: F(A,B) = 1{A- B- 0.5 > 0)



(d) (i) $P(A = 0) = P(A = 0, B = 0) + P(A = 0, B = 1) = 0.32 + 0.48 = 0.8$
$P(A = 1) = 1 - P(A = 0) = 0.2$
$P(B = 1) = P(B = 1, A = 0) + P(B = 1, A = 1) = 0.48 + 0.12 = 0.6$
$P(B = 0) = 1 - P(B = 1) = 0.4$

$P(A = 1|B = 0) = P(A = 1,B = 0)/P(B = 0) = 0.08/0.4 = 0.2$

(ii) Yes,
$P(A = 0)P(B = 0) = 0.8 * 0.4 = 0.32 = P(A = 0, B = 0)$
$P(A = 0)P(B = 1) = 0.8 * 0.6 = 0.48 = P(A = 0, B = 1)$
$P(A = 1)P(B = 0) = 0.2 * 0.4 = 0.08 = P(A = 1, B = 0)$
$P(A = 1)P(B = 1) = 0.2 * 0.6 = 0.12 = P(A = 1, B = 1)$

Ans 4 [5 Marks]. (a) SVD of matrix $C = U \sum V'$ (' represents the transpose)

where $\sum$ is a diagonal matrix containing Eigen values
U and V are the ortho-normal matrices

To compute the SVD of C: compute the following two matrices

$C'*C = V \sum' \sum V$

$CV = U \sum$

hence, $U = CV \sum^{-1}$ (-1 represents the inverse)

V is ortho-normal matrix and hence make the eigen vectors a unit vector.

Singular Values ($\sum_{6*5}$)=

6.480	0	0	0	0
0	3.464	0	0	0
0	0	6.682e-16	0	0
0	0	0	4.344e-16	0
0	0	0	0	5.669e-34
0	0	0	0	0

Left Singular vectors (U_{6*6}) =

-0.267	0	-0.948	0	-0.171	0
0.8017	0	-0.317	0	0.506	0
-0.534	0	-0.0025	0	0.845	0
0	-0.408	0	-0.912	0	-7.817e-17
0	0.816	0	-0.365	0	0.447
0	-0.408	0	0.182	0	0.894

Left Singular vectors (V_{5*5}) =

-0.577	0	0.816	0	0
-0.577	0	-0.408	0	-0.707
-0.577	0	-0.408	0	0.707
0	0.707	0	-0.707	0
0	0.707	0	0.707	0.

Verify: $C= U*\sum*V'$

(b) $\pm[0.577, 0.577, 0.577, 0, 0]$

Ans 5 [15 Marks].

Nearest Neighbors

1. How does it overfit?
(a) Every point in dataset (including noise) defines its own decision boundary.
(b) The distance function can be chosen to do well on training set but less well on new data.

2. How can you reduce overfitting?
(a) Use k-NN for larger k (b) Use cross-validation to choose k and the distance function

Decision Trees

1. How does it overfit?
By adding new tests to the tree to correctly classify every data point in the training set.

2. How can you reduce overfitting?
By pruning the resulting tree based on performance on a validation set.

Ans 6 [10 Marks].