

Name: Lianhan Huang 3700459; Qu Wang 3700666

1. True or False

1.1 Yes

1.2 Yes

1.3 No. LSM asks for linear structure, but only the logit transformation $\log(P/(1-P))$ is linear. But P can not be observed, only y, which is either 0 or 1, can be observed.

1.4 No. Predicted probabilities $\in [0,1]$.

1.5 Yes.

2. From Linear Regression to Classification

2.1(a)

$X = [[1,1,3],[1,-2,2],[1,0.3,1],[1,5,-1],[1,3,4],[1,7,3]]$

$X^T = [[1,1,1,1,1,1],[1,-2,0.3,5,3,7],[3,2,1,-1,4,3]]$

$Y = [1,0,0,1,1,1]$

$X^T * X = \begin{bmatrix} 6 & 14.3 & 12 \\ 14.3 & 88.09 & 27.3 \\ 12 & 27.3 & 40 \end{bmatrix}$

$X^T * Y = [4,16,9]$

$(X^T * X)^{-1} = \begin{bmatrix} 0.53690677 & -0.04723016 & -0.12883745 \\ -0.04723016 & 0.01855194 & 0.00150735 \\ -0.12883745 & 0.00150735 & 0.06262247 \end{bmatrix}$

$\beta = [0.23240754 \ 0.12147658 \ 0.07236997]$

$y = 0.23 + 0.12x_1 + 0.07x_2$

2.1(b)

$\hat{Y} = [0.571, 0.134, 0.341, 0.767, 0.886, 1.3]$

Classifier: If $\hat{Y} > 0.5$, $Y=1$

If $\hat{Y} < 0.5$, $Y=0$

2.2

Discriminative function models the decision boundary between different classes.