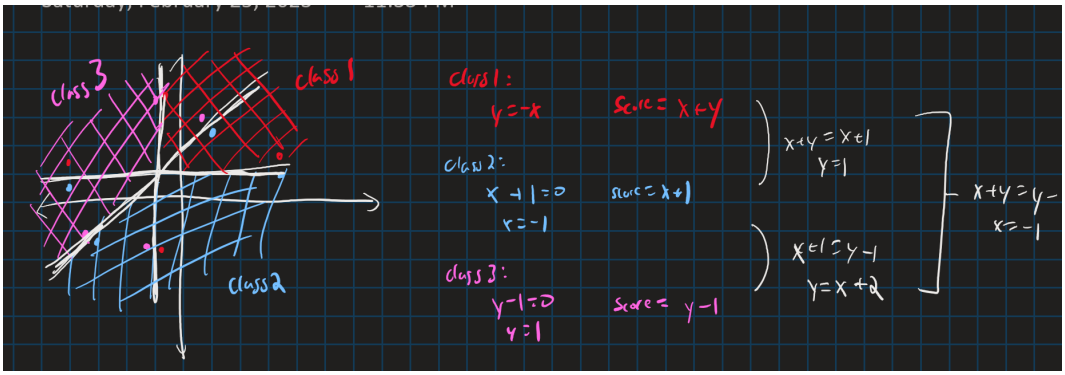
James Zhao
HW5
A15939512

1 1



2 2

$x_1 + 3x_1x_2 - 6x_2^2 - 8 = 0$
$= (x_1, x_2, x_1^2, x_2^2, x_1x_2) \cdot (1, 0, 0, -6, 3) - 8 = 0$
$w = (1, 0, 0, -6, 3)$

3 3

3.1 a

Yes. If the data is linearly separable with just order-1 terms, it can learn a weight vector $\Phi(x)$ that has 0 in which all the terms with degree > 1 are 0.

3.2 b

It will not necessarily return a vector with all quadratic terms being 0. It is possible that the data has some wiggle room where some terms can be nonzero, but still be linearly separable.

4 4

4.1 a

$\alpha = (0, 2, 2, 1)$

4.2 b

$b = 2 * 1 + 2 * -1 + -1 = -1$

5 5

5.1 a

Dimension of α is the number of training points, or 9*4=36.

5.2 b

Number of entries in α that are > 0 is the number of points that used updates. There are only 6 points that used updates (the larger dots)

5.3 c

None are. The constraints of kernel SVM are that all α_i are non-negative.

6 6

7 7

8 8

8.1 a

8.2 b

C	Train Err	Test Err
0.01	0.1544	0.1579
0.1	0.1359	0.1429
1.0	0.1204	0.1219
10	0.1331	0.1326
100	0.1500	0.1626

8.3 c

C	Train Err	Test Err
1.0	0.0137	0.0244

n_support=9158