THE CHINESE UNIVERSITY OF HONG KONG, SHENZHEN

DDA 3020

MACHINE LEARNING

Assignment2 Report

Author:
Zhao Rui

Student Number:
121090820

March 12, 2023

Contents

1   Written Questions

1.1   Question 1

Q1: (1) F. As $l(w, D)$ is concave function, $\pi \|w\|_2^2$ is convex function. So $J(w)$ is convex function. Has only global optimal solution.

(2) F. For all feature set. We can find a linearly separable set of data required all feature to separate the classes.

(3) T. To minimize $J(\hat{w})$, we have $J'(\hat{w}) = 0$, then $\hat{w}^2 = \frac{1}{2\pi|D|}$, when $\pi = 0$, some w_j might become infinite.

(4) F. When π increase, we will have $\hat{w}$ decrease. So at this time, $l(\hat{w}, D_{train})$ will also decrease

(5) F. $l(\hat{w}, D_{test})$ will decrease as $\hat{w}$ decrease.

1.2   Question 2

Q_2:

(1) ξ_i means the distant from hyperplane to the point at this time. if ξ_i taken negative value, the point have no error but $\frac{C}{2}\sum_{i=1}^{m}\xi_i^2$ will be larger, so at this time ξ_i will be 0 to minimize $\frac{C}{2}\sum_{i=1}^{m}\xi_i^2$ so the optimal value will be the same.

(2) $L(w,b,\xi,a) = \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{m}\xi_i^2 - \sum_{i=1}^{m}a_i\left[y^{(i)}(w^Tx^{(i)}+b)-1+\xi_i\right]$

$i=1,\cdots,m \qquad a_i \geq 0$

(3) $\frac{\partial L}{\partial w} = w - \sum_{i=1}^{m}a_iy^{(i)}x^{(i)} = 0 \Rightarrow w = \sum_{i=1}^{m}a_iy^{(i)}x^{(i)} \qquad i=1,\cdots,m$

$\frac{\partial L}{\partial b} = -\sum_{i=1}^{m}a_iy^{(i)} = 0 \Rightarrow \sum_{i=1}^{m}a_iy^{(i)} = 0 \qquad i=1,\cdots,m$

$\frac{\partial L}{\partial \xi} = C\xi - a = 0 \Rightarrow \xi_i = \frac{a_i}{C} \qquad i=1,\cdots,m$

$\min L(w,b,\xi,a)$

$= \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{m}\frac{a_i^2}{C^2} - \sum_{i=1}^{m}a_i\left[y^{(i)}(w^Tx^{(i)}+b)-1\right] - \sum_{i=1}^{m}a_i\frac{a_i}{C}$

$= \frac{1}{2}\|w\|^2 - \frac{1}{2C}\sum_{i=1}^{m}a_i^2 - \sum_{i=1}^{m}a_iy^{(i)}\left(\sum_{i=1}^{m}a_iy^{(i)}x^{(i)}\right)^Tx^{(i)} - b\sum_{i=1}^{m}a_iy^{(i)} + \sum_{i=1}^{m}a_i$

$= \frac{1}{2}\|w\|^2 - \frac{1}{2C}\sum_{i=1}^{m}a_i^2 - \sum_{i,j}^{m}a_ia_jy^{(i)}y^{(j)}x^{(i)T}x^{(j)} + \sum_{i=1}^{m}a_i$

$= \sum_{i=1}^{m}a_i - \frac{1}{2C}\sum_{i=1}^{m}a_i^2 - \frac{1}{2}\sum_{i,j}^{m}a_ia_jy^{(i)}y^{(j)}x^{(i)T}x^{(j)}$
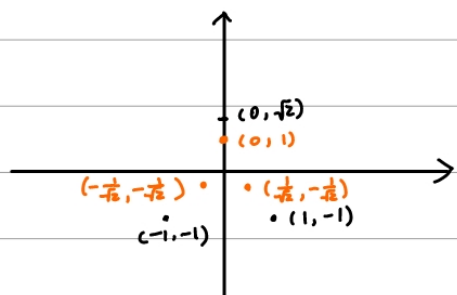
(4) $\max_{a} \sum_{i=1}^{m}a_i - \frac{1}{2C}\sum_{i=1}^{m}a_i^2 - \frac{1}{2}\sum_{i,j}^{m}a_ia_jy^{(i)}y^{(j)}x^{(i)T}x^{(j)}$

$s.t. \sum_{i=1}^{m}a_iy^{(i)} = 0 \qquad a_i \geq 0 \qquad i=1,\cdots,m$

1.3   Question 3

Q3:

(1)

$(0, \sqrt{2})$
$(0, 1)$
$(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ ·   · $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$
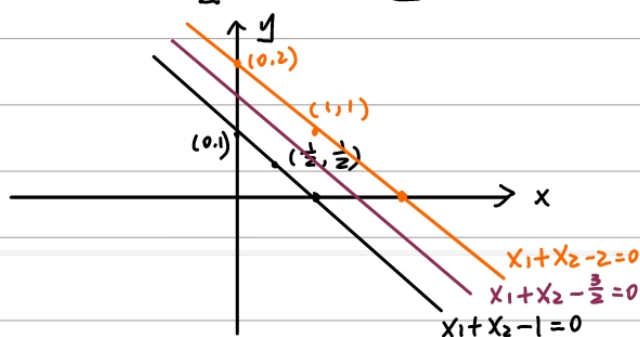$(-1, -1)$ · $(1, -1)$

we can not find a svm
classifier (without slack variable)
for this data set. as for the
left figure (the black points in
class +1, orange points in class -1)
they can not be seperated
linear, but SVM can only handle
linearly separable data.

(2) $\phi(x) = [x_1^2 ; x_2^2]$

$$\text{Class } -1: \begin{bmatrix} (0, 1) \\ (\frac{1}{2}, \frac{1}{2}) \\ (\frac{1}{2}, \frac{1}{2}) \end{bmatrix} \qquad \text{class } +1: \begin{bmatrix} (0, 2) \\ (1, 1) \\ (1, 1) \end{bmatrix}$$

y
$(0,2)$
$(1,1)$
$(0,1)$ $(\frac{1}{2}, \frac{1}{2})$
x
$x_1 + x_2 - 2 = 0$
$x_1 + x_2 - \frac{3}{2} = 0$
$x_1 + x_2 - 1 = 0$

from the right figure we
can see $w = [1 ; 1]$ $b = -\frac{3}{2}$
for $x_0 = [(-\frac{1}{\sqrt{2}}, \sqrt{2})]$
$\phi(x_0) = [\frac{1}{2}, 2]$
$w^T \phi(x_0) + b$

$= \frac{1}{2} + 2 - \frac{3}{2} = 1$

it's label will be +1

4

1.4   Question 4

Q4:

(1) $H(\text{passed}) = H(\frac{2}{3}, \frac{1}{3}) = -\frac{2}{3} \log_2 \frac{2}{3} - \frac{1}{3} \log_2 \frac{1}{3}$

$$= -\frac{2}{3}(1-1.6) + \frac{1}{3} \times 1.6 = \frac{14}{15}$$

(2) $H(\text{Passed} \mid \text{GPA}) = \frac{1}{3} H(\frac{1}{2}, \frac{1}{2}) + \frac{1}{3} H(\frac{1}{2}, \frac{1}{2}) + \frac{1}{3} H(0,1)$

$$= \frac{1}{3}(-\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2}) \times 2 + \frac{1}{3} \times 0$$
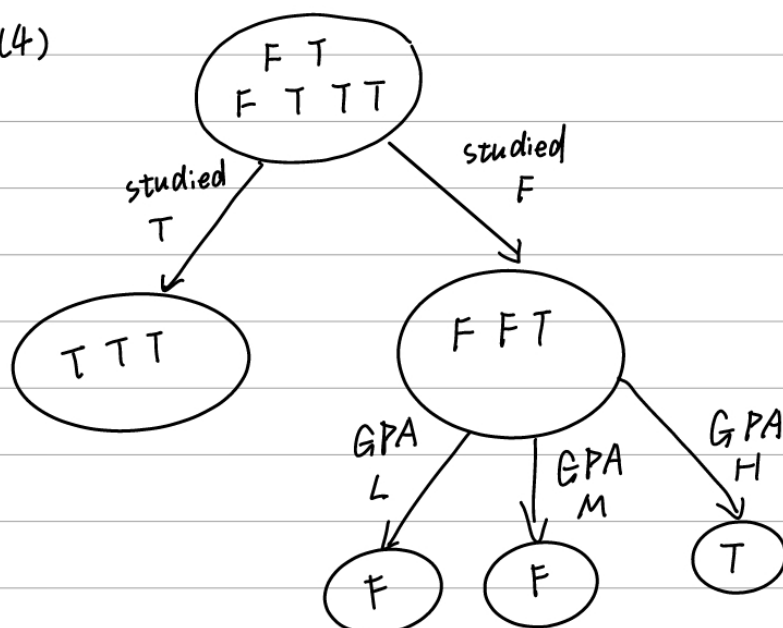
$$= \frac{2}{3}$$

(3) $H(\text{Passed} \mid \text{Studied}) = \frac{1}{2} H(0,1) + \frac{1}{2} H(\frac{1}{3}, \frac{2}{3})$

$$= \frac{1}{2} \times 0 + \frac{1}{2}(-\frac{1}{3} \log_2 \frac{1}{3} - \frac{2}{3} \log_2 \frac{2}{3})$$

$$= \frac{1}{2} \times 0 + \frac{1}{2} \times \frac{14}{15}$$

$$= \frac{7}{15}$$

(4)

2 Programming Question

2.1 Subproblem 1

In this problem, we need to solve the question:

$$\min_{w,b} \tfrac{1}{2}||w||$$

$$s.t. 1 - y_i(w^\top x_i + b) \le 0, \forall i$$

The Lagrange function is:

$$L(w, b, \alpha) = \tfrac{1}{2}||w||^2 + \sum_i^m \alpha_i(1 - y_i(w^\top x_i + b))$$

And to derive the problem, we need to transform it into its dual problem:

$$\max_\alpha \sum_{i=0}^m \alpha_i - \tfrac{1}{2}\sum_{i,j} \alpha_i\alpha_j y_i y_j x_i^\top x_j$$

$$s.t. \sum_{i=0}^m \alpha_i y_i = 0, \alpha_i \ge 0, \forall i$$

We set the parameter C = 1e5 and use sklearn.svm.svc() function. Finally, after running the program. And the error means the percentage of wrong prediction.

The training error is 0.0

The testing error is 0.05555555555555555

w of wine type 0 is [1.24726288 0.51082573 1.89779761 -0.18493888 0.00632474 -0.21005112 0.26584414 0.48027629 -0.10075874 0.38592801 -0.36948271 0.24884132 0.00423172]

b of wine type 0 is -24.0502332548531

The support vector indices of wine type0 is [13, 23, 38, 79, 81]

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901]

b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 106, 119, 131, 136]

w of wine type 2 is [-0.23960522 -0.20251868 -0.11563171 0.01773131 -0.02291777 0.27459377 0.51489556 -0.06029894 0.22303912 -0.94764994 0.28980052 0.5664039 -0.00119474]

b of wine type 2 is 7.538778443710943

The support vector indices of wine type2 is [10, 25, 32, 58, 77, 107]

the number of positive class slack variable of wine type 0 is 2 ,the number of negative class slack variable of wine type 0 is 42

the number of positive class slack variable of wine type 1 is 49 ,the number of negative class slack variable of wine type 1 is 49

the number of positive class slack variable of wine type 2 is 40 ,the number of negative class slack variable of wine type 2 is 2

Besides, as we can see the training error is 0, so the data is linear separable.

2.2   Subproblem 2

In this problem, we need to solve the question for any two classes. Each time, we think of one class as +1 and others as -1, so the separate plane can separate one class from the other for three-time:

$$\min_{w,b,\xi} \frac{1}{2}||w|| + C\sum_{i=0}^{m}\xi_i$$

$$s.t. 1 - \xi_i - y_i(w^\top x_i + b) \le 0, -\xi_i \le 0, \forall i$$

Its Lagrange function is:

$$L(w,b,\xi,\alpha,\mu) = \frac{1}{2}||w||^2 + C\sum_{i}^{m}\xi_i + \sum_{i}^{m}[\alpha_i(1 - \xi_i - y_i(w^\top x_i + b)) + \mu_i(-\xi_i)], \text{ and}$$
$$\alpha_i, \mu_i \ge 0, \forall i$$

$$\frac{\partial L}{\partial w} = 0, \frac{\partial L}{\partial b} = 0, \frac{\partial L}{\partial \xi_i}$$

$$w = \sum_{i}^{m}\alpha_i y_i x_i, \sum_{i}^{m}\alpha_i y_i = 0, \alpha_i = C - \mu_i, \forall i$$

$$\alpha_i \ge 0, 1 - \xi_i - y_i(w^\top x_i + b) \le 0, \xi_i \ge 0, \mu_i \ge 0, \forall i$$

$$\alpha_i(1 - \xi_i - y_i(w^\top x_i + b)) = 0, \mu_i\xi_i = 0, \forall i$$

$$L(\alpha,\mu) = \frac{1}{2}||w||^2 + \sum_{i}^{m}[\alpha_i(1 - y_i(w^\top x_i + b))] + \sum_{i}^{m}(C - \alpha_i - \mu_i)\xi_i$$

And the dual problem should be:

$$\max_{\alpha} \sum_{i=0}^{m}\alpha_i - \frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j x_i^\top x_j$$

$$s.t. \sum_{i=0}^{m}\alpha_i y_i = 0, \alpha_i \ge 0, \forall i$$

For the positive class, the slack variables appear if and only if when this class separates from other class, but $w^\top x_i + b < 1$, at this time ξ_i must not equal to zero. And the ξ_i equal to the distance from the point to the hyperplane:

$$\frac{|w^\top x_i + b - 1|}{||w||}$$

For the negative class, the slack variables appear if and only if when this class separates from other class, but $w^\top x_i + b > -1$, at this time ξ_i must not equal to zero. And the ξ_i equal to the distance from the point to the hyperplane:

$$\frac{|w^\top x_i + b + 1|}{||w||}$$

We set the $C = 0.1 \times t, t = 1, 2, ..., 10$, and after the fitting, we get the training error, testing error, weight vector w, bias b, the indices of support vectors, and the number of non-zero slack variables as follow:

FOR C = 0.1 :

The training error is 0.014084507042253521

The testing error is 0.08333333333333333

w of wine type 0 is [0.41730962 0.4659169 0.20545096 -0.12578337 -0.00803297 0.00966485 0.12604541 0.04662531 -0.08045934 0.34891393 -0.08128877 0.17649093 0.00501105], b of wine type 0 is -9.701331395615595

The support vector indices of wine type0 is [2, 5, 13, 38, 50, 79, 81]

the number of positive class slack variable of wine type 0 is 6

the number of negative class slack variable of wine type 0 is 42

w of wine type 1 is [5.20387035e-02 6.57172969e-02 9.57968063e-02 -5.76790932e-02 7.84359865e-03 2.87734709e-01 4.56259915e-01 -3.73332649e-04 1.70952928e-01 -1.01922702e-01 6.69604841e-02 3.66106417e-01 1.65560319e-03], b of wine type 1 is -4.213634315697107

The support vector indices of wine type1 is [0, 11, 31, 33, 68, 78, 83, 94, 97, 106, 119, 125, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.1809895 -0.22375927 -0.0724963 -0.03149182 0.00114604 0.11634512 0.54432367 -0.04697161 0.13243704 -0.54655004 0.11338837 0.32426399 -0.00246963], b of wine type 2 is 5.481803497191667

The support vector indices of wine type2 is [10, 25, 32, 55, 58, 77, 107, 124]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 5

FOR C = 0.2 :

The training error is 0.007042253521126761

The testing error is 0.08333333333333333

w of wine type 0 is [0.44911146 0.52875126 0.36384757 -0.14849315 -0.006408 -0.00081434 0.14949645 0.08041011 -0.17507929 0.37461792 -0.07698614 0.21466524 0.00544455], b of wine type 0 is -10.673897706083247

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 4

the number of negative class slack variable of wine type 0 is 42

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 97, 106, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.22002586 -0.22276202 -0.06542054 -0.05559606 -0.00196552 0.17691697 0.58886902 -0.02306193 0.09022018 -0.60763075 0.18653495 0.39813105 -0.00237621], b of wine type 2 is 6.634707091427711

The support vector indices of wine type2 is [10, 25, 32, 55, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 1

FOR C = 0.3 :

The training error is 0.007042253521126761

The testing error is 0.05555555555555555

w of wine type 0 is [0.4982509 0.5046472 0.52776239 -0.15722481 -0.00650179 0.01217359 0.15016586 0.11026875 -0.16757049 0.33217157 -0.09385907 0.26794495 0.00530959], b of wine type 0 is -11.37445019753535

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 4

the number of negative class slack variable of wine type 0 is 40

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 97, 106, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.21267044 -0.2298503 -0.0533564 -0.05141952 -0.00663904 0.18390736 0.54289168 -0.02625746 0.10632363 -0.67963334 0.2511058 0.42020337 -0.00193539], b of wine type 2 is 6.868183237841175

The support vector indices of wine type2 is [10, 25, 32, 55, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 2

FOR C = 0.4 :

The training error is 0.007042253521126761

The testing error is 0.05555555555555555

w of wine type 0 is [0.55043096 0.48569619 0.69200335 -0.16712588 -0.00665837 0.02495195 0.15215734 0.14030686 -0.16168884 0.29374693 -0.11112818 0.32219135 0.00522635], b of wine type 0 is -12.147452059394817

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 3

the number of negative class slack variable of wine type 0 is 38

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 97, 106, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.22589421 -0.20682077 -0.09290671 -0.0226983 -0.01426946 0.2505764 0.54839716 -0.04390481 0.18274595 -0.82041559 0.28217024 0.47700208 -0.00152428], b of wine type 2 is 7.197438130272816

The support vector indices of wine type2 is [10, 32, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 2

FOR C = 0.5 :

The training error is 0.0070422253521126761

The testing error is 0.05555555555555555

w of wine type 0 is [0.62190362 0.56419535 0.84081 -0.1944432 -0.00574664 0.06787711 0.15423401 0.16771333 -0.22240904 0.35096805 -0.14472737 0.3613965 0.00584758], b of wine type 0 is -13.825989206336398

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 3

the number of negative class slack variable of wine type 0 is 38

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 97, 106, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.23964913 -0.20255183 -0.11562884 0.01773382 -0.02292411 0.27457019 0.51480452 -0.06030505 0.22305472 -0.94768468 0.28976967 0.56644938 -0.00119466], b of wine type 2 is 7.540179157008514

The support vector indices of wine type2 is [10, 25, 32, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 2

FOR C = 0.6 :

The training error is 0.0

The testing error is 0.05555555555555555

w of wine type 0 is [7.09421700e-01 6.05540144e-01 1.00003103e+00 -2.08871772e-01 7.62040176e-04 9.47471776e-02 1.68852875e-01 1.99375801e-01 -2.87768413e-01

4.52050159e-01 -1.68374739e-01 3.89885196e-01 5.49302786e-03], b of wine type 0 is -15.97953053142151

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 3

the number of negative class slack variable of wine type 0 is 40

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 106, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.23960522 -0.20251868 -0.11563171 0.01773131 -0.02291777 0.27459377 0.51489556 -0.06029894 0.22303912 -0.94764994 0.28980052 0.5664039 -0.00119474], b of wine type 2 is 7.538778443710943

The support vector indices of wine type2 is [10, 25, 32, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 2

FOR C = 0.7 :

The training error is 0.0

The testing error is 0.05555555555555555

w of wine type 0 is [0.77939181 0.61705376 1.12616102 -0.21585168 0.00509638 0.09885689 0.18483038 0.22695069 -0.31620008 0.50546216 -0.18257052 0.41005302 0.00512433], b of wine type 0 is -17.503241569044828

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 3

the number of negative class slack variable of wine type 0 is 41

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 106, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.23960522 -0.20251868 -0.11563171 0.01773131 -0.02291777 0.27459377 0.51489556 -0.06029894 0.22303912 -0.94764994 0.28980052 0.5664039 -0.00119474], b of wine type 2 is 7.538778443710943

The support vector indices of wine type2 is [10, 25, 32, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 2

FOR C = 0.8 :

The training error is 0.0

The testing error is 0.05555555555555555

w of wine type 0 is [0.80361874 0.58367817 1.19144035 -0.21201501 0.00528336 0.09075537 0.18970539 0.24371629 -0.29298287 0.47231808 -0.18900121 0.40874674 0.00483402], b of wine type 0 is -17.70020855459486

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 3

the number of negative class slack variable of wine type 0 is 41

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 106, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.23960522 -0.20251868 -0.11563171 0.01773131 -0.02291777 0.27459377 0.51489556 -0.06029894 0.22303912 -0.94764994 0.28980052 0.5664039 -0.00119474], b of wine type 2 is 7.538778443710943

The support vector indices of wine type2 is [10, 25, 32, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 2

FOR C = 0.9 :

The training error is 0.0

The testing error is 0.05555555555555555

w of wine type 0 is [0.82998124 0.55282597 1.25777093 -0.20844499 0.00562087 0.08114077 0.19253464 0.26098357 -0.27424745 0.43813031 -0.19544902 0.40925303 0.00455341], b of wine type 0 is -17.93753550602574

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 3

the number of negative class slack variable of wine type 0 is 41

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.23960522 -0.20251868 -0.11563171 0.01773131 -0.02291777 0.27459377 0.51489556 -0.06029894 0.22303912 -0.94764994 0.28980052 0.5664039 -0.00119474], b of wine type 2 is 7.538778443710943

The support vector indices of wine type2 is [10, 25, 32, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 2

FOR C = 1.0 :

The training error is 0.0

The testing error is 0.05555555555555555

w of wine type 0 is [0.88715225 0.55897693 1.3395635 -0.20878282 0.00813466 0.04971187 0.16595902 0.28554012 -0.32016366 0.38894324 -0.20221663 0.43616668 0.0044169], b of wine type 0 is -18.762155810258577

The support vector indices of wine type0 is [2, 13, 38, 79]

the number of positive class slack variable of wine type 0 is 3

the number of negative class slack variable of wine type 0 is 39

w of wine type 1 is [0.06994742 0.10415368 0.07834312 -0.07731134 0.01131957 0.3191686 0.49566427 0.00532707 0.18240291 -0.11562862 0.05264111 0.37415947 0.00176901], b of wine type 1 is -4.649588106026545

The support vector indices of wine type1 is [0, 11, 68, 83, 94, 119, 131, 136]

the number of positive class slack variable of wine type 1 is 49

the number of negative class slack variable of wine type 1 is 49

w of wine type 2 is [-0.23960522 -0.20251868 -0.11563171 0.01773131 -0.02291777 0.27459377 0.51489556 -0.06029894 0.22303912 -0.94764994 0.28980052 0.5664039 -0.00119474], b of wine type 2 is 7.538778443710943

The support vector indices of wine type2 is [10, 25, 32, 58, 77, 107]

the number of positive class slack variable of wine type 2 is 40

the number of negative class slack variable of wine type 2 is 2

2.3   Subproblem 3

For a 2nd-order polynomial kernel:

In this question, we define the kernel:

$$k(x_i, x_j) = (1 + \frac{x_i^\top x_j}{\sigma^2})^2$$

And when solving the dual problem of the original problem, we utilize this kernel to replace $x_i x_j$, we have the following dual problem:

$$\max_\alpha \sum_{i=0}^m \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j)$$
$$s.t. \sum_{i=0}^m \alpha_i y_i = 0, \alpha_i \geq 0, \forall i$$

And after solving this problem using sklearn, we have:

The training error is 0.2887323943661972

The testing error is 0.3611111111111111

b of wine type 0 is -2.378638088675321

The support vector indices of wine type0 is [2, 5, 13, 17, 35, 38, 47, 79, 81, 86, 103, 129]

b of wine type 1 is -2.6905503877417587

The support vector indices of wine type1 is [0, 1, 3, 4, 7, 11, 12, 15, 19, 29, 31, 33, 40, 57, 60, 63, 65, 67, 68, 71, 78, 83, 84, 87, 93, 97, 106, 108, 110, 111, 115, 119, 121, 122, 125, 127, 131, 137]

b of wine type 2 is 1.6291534909876972

The support vector indices of wine type2 is [6, 10, 14, 16, 18, 22, 25, 26, 27, 28, 32, 41, 44, 48, 49, 51, 52, 55, 58, 59, 61, 69, 70, 73, 76, 77, 85, 92, 95, 96, 99, 107, 112, 116, 117, 124, 128, 133, 134, 135]

For a 3nd-order polynomial kernel:

In this question, we define the kernel:

$$k(x_i, x_j) = (1 + \frac{x_i^\top x_j}{\sigma^2})^3$$

And when solving the dual problem of the original problem, we utilize this kernel to replace $x_i x_j$, we have the following dual problem:

$$\max_\alpha \sum_{i=0}^m \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j)$$

$$s.t. \sum_{i=0}^m \alpha_i y_i = 0, \alpha_i \geq 0, \forall i$$

And after solving this problem using sklearn, we have:

The training error is 0.2887323943661972

The testing error is 0.3611111111111111

b of wine type 0 is -2.183049132988348

The support vector indices of wine type0 is [2, 5, 13, 17, 35, 38, 79, 81, 86, 103, 129]

b of wine type 1 is -2.073776123047625

The support vector indices of wine type1 is [0, 1, 3, 4, 7, 11, 12, 15, 19, 29, 31, 33, 40, 57, 60, 63, 65, 67, 68, 71, 78, 83, 84, 87, 93, 97, 106, 108, 110, 111, 115, 119, 121, 122, 125, 127, 131, 137]

b of wine type 2 is 1.2656977404916434

The support vector indices of wine type2 is [6, 10, 14, 16, 18, 22, 25, 26, 27, 28, 32, 41, 44, 48, 49, 51, 52, 55, 58, 59, 61, 69, 70, 73, 76, 77, 85, 92, 95, 96, 99, 107, 112, 116, 117, 124, 128, 133, 134, 135]

For Radial Basis Function kernel with $\sigma = 1$:

In this question, we define the kernel:

$$k(x_i, x_j) = \exp\{-\frac{||x_i - x_j||^2}{2\sigma^2}\}$$

16

And when solving the dual problem of the original problem, we utilize this kernel to replace $x_i x_j$, we have the following dual problem:

$$\max_\alpha \sum_{i=0}^m \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j)$$
$$s.t. \sum_{i=0}^m \alpha_i y_i = 0, \alpha_i \geq 0, \forall i$$

And after solving this problem using sklearn, we have:

The training error is 0.0

The testing error is 0.5

b of wine type 0 is -0.0754715050294759

The support vector indices of wine type0 is [2, 5, 8, 13, 17, 20, 23, 24, 34, 35, 37, 38, 39, 42, 43, 45, 46, 47, 50, 53, 56, 62, 64, 66, 72, 75, 79, 80, 81, 86, 88, 89, 90, 91, 101, 102, 103, 104, 105, 109, 114, 118, 123, 126, 129, 130, 132, 139, 141]

b of wine type 1 is 0.1836671705558712

The support vector indices of wine type1 is [0, 1, 3, 4, 7, 9, 11, 12, 15, 19, 21, 29, 30, 31, 33, 36, 40, 54, 57, 60, 63, 65, 67, 68, 71, 74, 78, 82, 83, 84, 87, 93, 94, 97, 98, 100, 106, 108, 110, 111, 113, 115, 119, 120, 121, 122, 125, 127, 131, 136, 137, 138, 140]

b of wine type 2 is 0.24528286146413605

The support vector indices of wine type2 is [6, 10, 14, 16, 18, 22, 25, 26, 27, 28, 32, 41, 44, 48, 49, 51, 52, 55, 58, 59, 61, 69, 70, 73, 76, 77, 85, 92, 95, 96, 99, 107, 112, 116, 117, 124, 128, 133, 134, 135]

For Sigmoidal kernel with $\sigma = 1$:

In this question, we define the kernel:

$$k(x_i, x_j) = \frac{1}{1 + \exp\{-\frac{x_i^\top x_j + b}{\sigma^2}\}}$$

And when solving the dual problem of the original problem, we utilize this kernel to replace $x_i x_j$, we have the following dual problem:

$$\max_\alpha \sum_{i=0}^m \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i, x_j)$$
$$s.t. \sum_{i=0}^m \alpha_i y_i = 0, \alpha_i \geq 0, \forall i$$

And after solving this problem using sklearn, we have:

The training error is 0.6267605633802817

The testing error is 0.5

b of wine type 0 is -1.0

The support vector indices of wine type0 is [2, 5, 8, 13, 17, 20, 23, 24, 34, 35, 37, 38, 39, 42, 43, 45, 46, 47, 50, 53, 56, 62, 64, 66, 72, 75, 79, 80, 81, 86, 88, 89, 90, 91, 101, 102, 103, 104, 105, 109, 114, 118, 123, 126, 129, 130, 132, 139, 141]

b of wine type 1 is 1.0

The support vector indices of wine type1 is [7, 9, 11, 12, 15, 19, 21, 29, 30, 31, 33, 36, 40, 54, 57, 60, 63, 65, 67, 68, 71, 74, 78, 82, 83, 84, 87, 93, 94, 97, 98, 100, 106, 108, 110, 111, 113, 115, 119, 120, 121, 122, 125, 127, 131, 136, 137, 138, 140]

b of wine type 2 is 1.0

The support vector indices of wine type2 is [6, 10, 14, 16, 18, 22, 25, 26, 27, 28, 32, 41, 44, 48, 49, 51, 52, 55, 58, 59, 61, 69, 70, 73, 76, 77, 85, 92, 95, 96, 99, 107, 112, 116, 117, 124, 128, 133, 134, 135]