ASSIGNMENT-4

Sumedh Masulkar(11736)

3. The 5 fold cross validated accuracy is found to be 0.805.

Q 1(a). Consider a ~~[illegible]~~ potential solution to the above problem with some $\xi < 0$. Then the constraint $y^{(i)}(w^T x^{(i)} + w_0) \geq 1 - \xi_i$ would also be satisfied for $\xi_i = 0$, and the objective function would be lower, proving that this could be an optimal solution.

(b.) $$\mathcal{L}(w, w_0, \xi, \alpha) = \frac{1}{2} w^T w + \frac{C}{2} \sum_{i=1}^{n} \xi_i^2 - \sum_{i=1}^{n} \alpha_i \left[y^{(i)}(w^T x^{(i)} + w_0) - 1 + \xi_i \right]$$

where $\alpha_i \geq 0$ for $i = 1 \ldots m$.

(c.) Now minimizing it by taking the gradient with respect to w, we get,

$$\nabla_w \mathcal{L} = w - \sum_{i=1}^{m} \alpha_i y^{(i)} x^{(i)} = 0.$$

$$\Rightarrow w = \sum_{i=1}^{m} \alpha_i y^{(i)} x^{(i)}$$

Taking the derivative with respect to w_0,

$$\frac{\partial \mathcal{L}}{\partial w_0} = -\sum_{i=1}^{m} \alpha_i y^{(i)} = 0.$$

$$0 = \sum_{i=1}^{m} \alpha_i y^{(i)}$$

Now for dual,

The objective function for the dual is

$$W(\alpha) = \min_{w, w_0, \xi} \mathcal{L}(w, w_0, \xi, \alpha)$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} (\alpha_i y^{(i)} x^{(i)})^T (\alpha_j y^{(j)} x^{(j)})$$

$$+ \frac{1}{2}\sum_{i=1}^{n} \frac{\alpha_i}{\xi_i} \xi_i^2$$

$$- \sum_{i=1}^{n} \alpha_i \Big[y^{(i)} \Big(\Big(\sum_{j=1}^{n} \alpha_j y^{(j)} x^{(j)} \Big)^T x^{(i)} + w_0 \Big) - 1 + \xi_i \Big]$$

$$= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y^{(i)} y^{(j)} (x^{(i)})^T x^{(j)}$$

$$+ \frac{1}{2}\sum_{i=1}^{n} \alpha_i \xi_i - \Big(\sum_{i=1}^{n} \alpha_i y^{(i)} \Big) w_0 +$$

$$\sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \alpha_i \xi_i$$

$$= \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y^{(i)} y^{(j)} (x^{(i)})^T x^{(j)}$$

$$- \frac{1}{2}\sum_{i=1}^{n} \alpha_i \xi_i$$

$$= \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y^{(i)} y^{(j)} (x^{(i)})^T (x^{(j)})$$

$$- \frac{1}{2}\sum_{i=1}^{m} \frac{\alpha_i^2}{C}$$

Then the dual formation is:

$$\max_\alpha \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y^{(i)} y^{(j)} (x^{(i)})^T x^{(j)}$$

$$- \frac{1}{2}\sum_{i=1}^{n} \frac{\alpha_i^2}{C}.$$

$$\text{s.t. } \alpha_i \geq 0, \quad i = 1..n$$

$$\sum_{i=1}^{n} \alpha_i y^{(i)} = 0.$$

(a) To find w and w_0

using $\frac{\partial L}{\partial w} = 0$

$$\Rightarrow \bar{w} - \sum_{i=1}^{n} \alpha_i y_i x_i = 0$$

$$\bar{w} = \sum_{i=1}^{n} \alpha_i y_i \bar{x}_i$$

using complemantarity conditon,

$$\alpha_i [y_i (\bar{w}^T \bar{x} + w_0) - 1] = 0.$$

$$\Rightarrow \alpha_i \left[y_i \left(\left(\sum_{i=1}^{n} \alpha_i y_i \bar{x}_i \right)^T \bar{x} + \bar{w}_0 \right) - 1 \right] = 0$$

$$\bar{w}_0 = \frac{1}{y_i} - \left(\sum_{i=1}^{n} \alpha_i y_i \bar{x}_i \right)^T \bar{x}.$$

Q.2 ν - SUM

$w^T x + w_0 = \pm\rho$ (Margin plane)

$$\min f(w, w^0, \xi, \rho) = \frac{\|w\|^2}{2} - \nu\rho + \frac{1}{n}\sum_{i=1}^{n} \xi_i$$

where $0 \le \nu \le 1$.

constraints $y_i[w^T x_i + w_0] \ge \rho + \xi_j \quad i = 1(1)n$

$\xi_i \ge 0 \quad i = 1(1)n$

$\rho \ge 0$.

(a) Lagrangian for the problem is:

$$L(w, w_0, \xi, \rho, \alpha, \beta, \delta) = \frac{1}{2}\|w\|^2 - \nu\rho + \frac{1}{n}\sum_{i=1}^{n}\xi_i \cancel{- \sum_{i=1}^{n}(\alpha_i(y_i(w^T w_i}$$

$$- \sum_{i=1}^{n}(\alpha_i(y_i/ w^T x_i + w_0) - \rho + \xi_i) + \beta_i \xi_j) - \delta\rho$$

where $\alpha_i, \beta_i, \delta \ge 0$

Now, this fn. has to minimize w.r.t w_i, w_0, ξ_j, ρ.

$$\frac{\partial L}{\partial w} = \nabla w L = w - \sum_{i=1}^{n} \alpha_i y_i x_i = 0$$

$$\Rightarrow w = \sum_{i=1}^{n} \alpha_i y_i x_i \quad -(1)$$

$$\frac{\partial L}{\partial w_0} = -\sum_{i=1}^{n} \alpha_i y_i = 0.$$

$$\therefore \sum_{i=1}^{n} \alpha_i y_i = 0 \quad -(2)$$

Q.

$$\frac{\partial L}{\partial \xi_i} = \frac{1}{n} - \alpha_i - \beta_i = 0$$

$$\Rightarrow \boxed{\alpha_i + \beta_i = \frac{1}{n}} \quad \text{--- (3)}$$

$$\frac{\partial L}{\partial \rho} = -\nu + \sum_{i=1}^{n} \alpha_i - \delta = 0$$

$$\Rightarrow \nu = \sum_{i=1}^{n} \alpha_i - \delta \quad \text{--- (4)}$$

Now substituting values, we have

$$w(\alpha, \beta, \delta) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_i y_i x_i)^T (\alpha_j y_j x_j)$$
$$+ \frac{1}{n} \sum_{i=1}^{n} \xi_j - \nu\rho - \sum_{i=1}^{n} \alpha_i \Big[y_i \big((\sum_{j=1}^{n} \alpha_j y_j x_j)^T x_i + w_0 \big)$$
$$- \rho + \xi_i \Big] - \delta\delta - \sum_{i=1}^{n} \beta_i \xi_i$$

$$= -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i x_i^T x_j + \frac{1}{n} \sum_{i=1}^{n} \xi_i - \nu\rho$$
$$- \Big(\sum_{i=1}^{n} \alpha_i y_i \Big) w_0 + \sum_{i=1}^{n} \alpha_i \rho - \sum_{i=1}^{n} (\alpha_i + \beta_i) \xi_i - \delta\rho$$

$$\Rightarrow w(\alpha, \beta, \delta) = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

(b) The dual formation is

$$\max_{\alpha, \beta, \delta} w(\alpha, \beta, \delta) = \max_{\alpha, \beta, \delta} -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

subject to $0 \le \alpha_i \le \frac{1}{n}$, $\sum_{i=1}^{n} \alpha_i \ge \nu$.

Q3

(c) Now to calculate $\bar{w}$, we can use the optimal values from dual problem α_i^* &

$$\bar{w}^* = \sum_{i=1}^{n} \alpha_i y_i x_i$$

To find w_0^*, we use the complementarity slackness condition, of KKT

$$\alpha_i \left[(w^T x_i + w_0) y_i - \rho + \xi_i \right] = 0$$

$$\beta_i \xi_i = 0$$

$$w^T x + w_0 = \pm \rho$$

w_0^* & ρ can be found from these equations.

Q.4(a). $S \subseteq R^n$ is a convex set
Show that
$$\theta_1\bar{x}_1 + \theta_2 x_2 + \dots + \theta_k x_k \in S, \text{ where } x_i \in S,\ i = 1 \dots k$$
$$\& \sum_{i=1}^{k} \theta_i = 1.$$

We need to prove that S contains any convex combination of vectors $x_i \in S$.

Let $y = \sum_{i=1}^{m} \theta_i x_i \qquad [\sum \theta_i = 1]$ — (A)

By applying induction on m,

Basecase: $m = 1$.

In case $m = 1$, $y = x_1 \in S$ which is true ~~already~~ already.

Hypothesis: Assume that we know that any convex combination of $m-1$ vectors ($m \geq 2$) from S is present in S.

To prove: The above statement remains valid for convex combination of m vectors as well.

From (A),

y can be written as:
$$y = (1-\theta_m) \sum_{i=1}^{m-1} \frac{\theta_i}{1-\theta_i} x_i + \theta_m x_m.$$

Note that θ_m can be assumed to be < 1 because $\sum \theta_i = 1$, & if $\theta_m = 1$, then other $\theta_i = 1$ to $m-1 = 0$.

$\therefore y = x_m \in S$

$y = (1-\theta_m) z + \theta_m x_m$

with z & $x_m \in S$.

$\therefore$ $y \in S$, [acc. to definition of convex se

Hence proved.

4(b). Show that the set,

$$S = \{x = (x_1, x_2) \in R_+^2 \mid x_1 x_2 \geq 1\}$$ is convex.

Consider a convex combination z of 2 points $x = (x_1, x_2)$ & $y = (y_1, y_2)$ belonging to the set S.

Case 1: if $x \geq y$, then.

$$z = \theta x + (1-\theta) y \geq y$$

& $z_1 z_2 \geq y_1 y_2 \geq 1$

Similarly for the case $y \geq x$

Case 2.

If $y \ngeq x$ & $x \ngeq y$

i.e. $(y_1 - x_1)(y_2 - x_2) < 0$

$$z_1 z_2 = (\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2)$$

$$= (\theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta) x_1 y_2 + \theta(1-\theta) x_2 y_1$$

$$= \theta^2 x_1 x_2 + y_1 y_2 + \theta^2 y_1 y_2 - 2\theta y_1 y_2 + \theta(1-\theta)[x_1 y_2 - y_1 x_2]$$

$$= \theta x_1 x_2 + (1-\theta) y_1 y_2 - \theta(1-\theta)(y_1 - x_1)(y_2 - x_2).$$

$$\geq 1.$$

Hence $z \in S$.

Show that $\{x \in (x_1, x_2, \ldots, x_n) \in R_+^n \mid \prod_{i=1}^{n} x_i > 1\}$ is convex consider a convex combination x of 2 points,

$x = (x_1, x_2, \ldots, x_n)$ & $y = (y_1, y_2, \ldots y_n)$ belonging to S.

$\prod_{i=1}^{n} x_i \geq 1$ & $\prod_{i=1}^{n} y_n = 1$ [since x & y belong to S]

<u>case 1</u> $x \geq y$.

$z = \theta x + (1-\theta) y \geq y$.

& $z_1 z_2 \cdots z_n \geq y_1 y_2 \cdots y_n \geq 1$.

Similarly for case $y \geq x$.

<u>Case 2.</u> If $y \ngeq x$ & $x \ngeq y$

i.e.

$(y_1 - x_1)(y_2 - x_2)(y_3 - x_3) \cdots (y_n - x_n) < 0$.

Inequality 1: If $\alpha, \beta \geq 0$ & $0 \leq \theta \leq 1$
then $\alpha^{\theta} \beta^{1-\theta} \leq \theta\alpha + (1-\theta)\beta$.

Using the above inequality,

$z_1 z_2 \ldots z_n$

$$= \prod_{i=1}^{n} (\theta x_i + (1-\theta) y_i) \geq \prod_{i=1}^{n} x_i \cdot y_i^{(1-\theta)}$$

$$\geq \left(\prod_{i=1}^{n} x_i\right)^{\theta} \left(\prod_{i=1}^{n} y_i\right)^{1-\theta}$$

$$\geq 1$$

Hence $x \in S$.

Q.4(c) Show that S is convex iff its intersection with any line is convex

Theorem A: Intersection of 2 convex sets is convex

$\Rightarrow$

using theorem A, if S is convex set then the intersection of S with a line is also convex set.

$\Leftarrow$ Suppose the intersection of S with any line is convex, take any 2 points x_1 & $x_2 \in S$. The intersection of S with the line through x_1 & x_2 is convex.

$\therefore$ convex combination of x_1 & x_2 belong to the intersection, hence they also belong to S.

Show that S is affine iff its intersection with line is affine.

The above statement can be proved similarly as in the case of convex sets, instead of convex sets, we use affine sets.

Q.4(d) Let cnvch(S) be the convex hull of S and let C be the intersection of all convex sets containing S, i.e.

$$C = \bigcap \{ c \mid c \text{ is convex}, c \supseteq S \}$$

To show: cnvch(S) = C.

For this, we need to show that

cnvch(S) $\subseteq$ C & C $\subseteq$ cnvch(S)

1) cnvch(S) $\subseteq$ C

Let x be a convex combination of some points $x_1, x_2, \ldots x_n$ ~~∈~~ $\in S$, or $x \in$ cnvch(S) —(A)

Let c be any convex set containing S [$c \supseteq S$]

$\therefore x_1, x_2, \ldots x_n \in c$. ——(B)

From A & B $\Rightarrow x \in c$ —(C)

From (C), for any convex set c, that contains S, $x \in c$. $\therefore x$ also belongs to intersection of all such convex sets c containing S

$\therefore x \in C$.

$\Rightarrow$ cnvch(S) $\subseteq$ C

2.) C $\subseteq$ cnvch(S)

Since cnvch(S) is a convex set containing

(d) continued

S, we must have $cnvxh(S) = C$ for some C during construction of C.

Hence proved.

From ① & ② $\Rightarrow$ $cnvxh(S) = C$.

The above proof can also be used for affine & ~~~~ conic hulls as well, instead of convex combinations, we use affine & conical combinations for them.

Q.4(e) $S = \{ x \mid \|x - a\|_2 \le \theta \|x - b\|_2 \}$

a & b are fixed points

$0 \le \theta \le 1$.

We assume that $a \ne b$.

$$S = \{ x \mid \|x - a\|_2^2 \le \theta^2 \|x - b\|_2^2 \}$$

$$= \{ x \mid (1-\theta^2) x^T x - 2(a - \theta^2 b)^T x + a^T a - \theta^2 b^T b) \le 0 \}$$

case 1. When $\theta = 1$.

$\Rightarrow$ It is a half space.

case 2 $\theta < 1$.

$\Rightarrow$ it is a ball

$$\{ x \mid (x - x_0)^T (x - x_0) \le R^2 \}$$

case 1 & case 2 $\Rightarrow$ S is convex.