2019213

ML - Assignment 3

Q2

Ans a) This non-negativity constraint on ε i.e $\varepsilon \geq 0$ can be removed. We can do so because even if we remove this contraint the optimal value of ~~of~~ our target will still be same, we ~~a~~ can s this as

~~S~~ Assume solution ~~is~~ with $\varepsilon < 0$ ~~that~~ the constraint

$$y^{(i)}(w^T x^{(i)} + b) \geq 1 - \varepsilon \text{ is true for } \varepsilon = 0$$

Hence we can say this can't be ~~an~~ optimal solution, hence ~~once~~ disproving our assumption, hence can have solution with $\varepsilon < 0$ so ~~x~~ ~~no~~ $\varepsilon \geq 0$ can be removed

Ans b) $$L(w, b, \varepsilon, \alpha) = \left\{ \frac{1}{2} w^T w + \frac{C}{2} \sum_{i=1}^{m} \varepsilon_i^2 - \sum_{i=1}^{m} \alpha_i \left[y^{(i)} (w^T x^{(i)} + b) - 1 + \varepsilon_i \right] \right\} \quad \forall \alpha_i \geq 0$$

Ans c) $$W(\alpha) = \min_{w, b, \varepsilon} L(w, b, \varepsilon, \alpha)$$

$$= \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} (\alpha_i y^{(i)} x^{(i)})^T (\alpha_j y^{(j)} x^{(j)}) + \frac{1}{2} \sum_{i=1}^{m} \frac{\alpha_i}{\varepsilon_i} \varepsilon_i^2 - \sum_{i=1}^{m} \alpha_i \left(y^{(i)} \left(\left(\sum_{j=1}^{m} \alpha_j y^{(j)} x^{(j)} \right)^T x^{(i)} + b \right) - 1 + \varepsilon_i \right)$$

dual $\Rightarrow$ $$\left[= \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y^{(i)} y^{(j)} x^{(i)T} x^{(j)} - \frac{1}{2} \sum_{i=1}^{m} \frac{\alpha_i^2}{C} \right] \quad \forall \alpha_i \geq 0$$

Q3

Ans a) Following hint we set $\alpha_i = 1$ and $b = 0$.

we get

$$|f(x^{(i)}) - y^{(i)}| = \left| \sum_{j=1}^{m} y^{(j)} K(x^{(j)}, x^{(i)}) - y^{(i)} \right|$$

$$= \left| \sum_{j=1}^{m} y^{(j)} e^{-\|x^{(j)} - x^{(i)}\|^2/\tau^2} - y^{(i)} \right|$$

if we get out the $y^{(i)}$ from $\sum_{j=1}^{m}$ we can cancel out $y^{(i)}$.

$$= \left| \sum_{j=1, j\neq i}^{m} y^{(j)} e^{-\|x^{(j)} - x^{(i)}\|^2/\tau^2} \right|$$

$$|y^{(j)}| = 1 \qquad \geq \varepsilon$$

$$\leq \sum_{j=1, j\neq i}^{m} e^{-\varepsilon^2/\tau^2} = (m-1) e^{-\varepsilon^2/\tau^2}$$

So, we need $(m-1) e^{-\varepsilon^2/\tau^2} < 1$

$$\boxed{\tau < \frac{\varepsilon}{\log(m-1)}}$$

Ans b) Yes, the resulting classifier will obtain zero training errors

So if a SVM finds a solution without slack variables then it always returns zero training error.

Showing there is a solution.

$$y^{(i)}(w^T x^{(i)} + b) = y^{(i)} \cdot f(x^{(i)}) > 0$$

as $f(x^{(i)}) \cdot y^{(i)} > 0$, they will have same sign.

Hence we choose large α_i

$y^{(i)} w^T x^{(i)} + b > 1$, hence optimisation problem is possible. Hence we get a solution.