Ans-2 Given: SVM algorithm with $l2$ norm soft such that the optimization function is given as

$$\min_{\omega,b,\epsilon} \frac{1}{2}||\omega||^2 + \frac{C}{2}\sum_{i=1}^{m} \epsilon^2 \quad \text{such that}$$

$$y^{(i)}(\omega^T x^{(i)} + b) \geq 1 - \epsilon_i \quad \forall i \in 1 \dots m$$

a. To find : wheather there should be a non negative constraint on ϵ

There shouldn't be any non negativity constraint in this. This is so because if we consider $\epsilon < 0$ hence the constraint $y(i)(\omega^T x^{(i)} + b) >= 1 - \epsilon$ will change

$$\because \quad \epsilon < 0$$

$$-\epsilon > 0$$

$$1 - \epsilon > 1$$

$\Rightarrow y^{(i)}(\omega^T x^{(i)} + b) > 1$ and hence would be solvable and the objective $f x^n$ would be lesser $\Rightarrow$ it is the optimal soln.

b. Lagrangian of l_2 norm

$$\mathcal{L}(\omega, b, \epsilon, \alpha) = \frac{1}{2}\omega^T\omega + \frac{C}{2}\sum_{i=1}^{m} \epsilon_i^2 - \sum_{i=1}^{m} \alpha_i [y^{(i)}(\omega^T x^{(i)} + b) - 1 + \epsilon]$$

$$\forall \alpha_i \geq 0 \quad \forall i \in [1, m]$$

c. Dual of optimization function

let dual be represented as $\omega(\alpha) = \min_{\omega, b, \epsilon} \mathcal{L}(\omega, b, \epsilon, \alpha)$

$$= \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} (\alpha_i y^i x^i)^T (\alpha_i y^j x^i) + \frac{1}{2} \sum_{i=1}^{m} \frac{\alpha_i}{\epsilon_i} \epsilon_i^2$$

$$- \sum_{i=1}^{m} \alpha_i \left[y^{(i)} \left\{ \left(\sum_{j=1}^{m} \alpha_j y^{(j)} x^{(j)} \right)^T x^i + b \right\} - 1 + \epsilon \right]$$

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

$$- \left(\sum_{i=1}^{m} \alpha_i y^i \right) b + \sum_{i=1}^{m} \alpha_i - \sum_{i=1}^{m} \alpha_i \epsilon_i$$

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

dual formation:

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

such that $\alpha_i \geq 0 \quad i = 1 \ldots m$

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

Ans-3 Given: SVM machine using gaussian kernel.

$$K(x, z) = \exp(-\|x - z\|^2/\tau^2)$$

To show: As long as there are no identical points in training set we can always find a value for the bandwidth parameter τ s.t. SVM achieves zero training error.

let $\alpha_i = 1$ $\forall i$ from 1 to m and $b=0$

for any training point $\{x^{(i)}, y^{(i)}\}$

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

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

$$= \left|\cancel{y^{(i)}} + \sum_{j\neq i} y^{(j)} \exp(\|x^j - x^i\|^2/\tau^2) \cancel{- y^i}\right|$$

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

$$\left|\sum_{j\neq i} y^{j} \exp(\|x^j - x^i\|^2/\tau^2)\right| \leq \sum_{j\neq i} y^j \exp(-\|x^j - x^i\|^2/\tau^2)$$

[triangular inequality]

$$\Rightarrow \sum_{j\neq i} |y^{(j)}| \exp(\|x^j - x^i\|^2/\tau^2)$$

$$\Rightarrow \sum_{j\neq i} \exp(-\|x^{(j)} - x^{(i)}\|^2/\tau^2)$$

$$\leq \sum_{j\neq i} \exp(\epsilon^2/\tau^2) = (m-1)\exp(-\epsilon^2/\tau^2)$$

$$[\because \|x^j - x^i\| \geq \epsilon]$$

now $(m-1)\exp(-\epsilon^2/\tau^2) < 1$

$$\exp(-\epsilon^2/\tau^2) < \frac{1}{m-1}$$

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

Hence for $\tau = \frac{\epsilon}{\log m}$ SVM will achive 0 training error.

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

b. Yes the model will achive zero training error. If there are no slack variables and if the machine is able to find a point which is the soln.

consider $y^{(i)}(w^T x^i + b)$ for some i

$$\text{let } b = 0$$

$$y^{(i)}(w^T x^i) = y^i f(x^i)$$

$$y^i f(x^i) > 0$$

~~The possi~~ for this to happen y^i & $f(x^i)$ should be of same sign

Hence for large α_i optimization problem $y^i(w^T x^i + b) \geq$ will be feasible.