AMIT YADAV

Assignment 3

① To prove : Norm is a convex function on R^D.

Sol : (i) Domain of Norm $= \mathbb{R}^D$

Clearly domain of Norm is a convex set because any linear combination of x_1 and x_2 s.t $x_1, x_2 \in \mathbb{R}^D$ is in $\mathbb{R}^D$.

(ii) $\|\theta x + (1-\theta)y\| \leq \theta\|x\| + (1-\theta)\|y\|$

Sol: LHS

$\|\theta x + (1-\theta)y\| \leq \|\theta x\| + \|(1-\theta)y\|$ (using triangle inequality)

$\leq \theta\|x\| + (1-\theta)\|y\|$ $(\because \|\alpha v\| = \alpha\|v\|)$

RHS

Since both conditions hold for norm function, therefore norm is a convex function.

Scanned by CamScanner

② <u>Langrangion</u> <u>functional</u> :

Given primal:

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

s.t

$$y_i (w^T \phi(x_i) + b) \geq 1 - \xi_i$$

$$\xi_i \geq 0 \quad \forall i \in (1, \ldots m)$$

Introducing langrange multipliers for every constraint:

$$L(w, \xi, b, \alpha, \beta) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i + \sum_{i=1}^{n} \alpha_i (1 - \xi_i - y_i (w^T \phi(x_i) + b)) - \sum_{i=1}^{n} \beta_i \xi_i$$

s.t. $\boxed{\alpha \geq 0, \ \beta \geq 0}$

---

③ $$\frac{\partial L}{\partial w} = w + \sum_{i=1}^{n} \alpha_i (-y_i \phi(x_i)) = w - \sum_{i=1}^{n} \alpha_i y_i \phi(x_i)$$

Setting $\frac{\partial L}{\partial w} = 0$, we get

$$w = \sum_{i=1}^{n} \alpha_i y_i \phi(x_i) \quad \text{——} \ (1)$$

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

Setting $\frac{\partial L}{\partial b} = 0$, we get

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

$$\frac{\partial L}{\partial \varepsilon_i} = c + (-\alpha_i) - \beta_i$$

Setting $\frac{\partial L}{\partial \varepsilon_i} = 0$, we get

$$\beta_i = c - \alpha_i \quad \text{——— (3)}$$

But given constraint $\alpha_i \geq 0$, $\beta_i \geq 0$

we get $\boxed{0 \leq \alpha_i \leq c}$ ——— (4)

---

(4) substituting value of $w = \sum \alpha_i y_i \phi(x_i)$ in $L(w, \varepsilon, b, \alpha, \beta)$

$$L = \frac{1}{2}\left(\sum_{i=1}^{n} \alpha_i y_i \phi(x_i)\right)^T \left(\sum_{i=1}^{n} \alpha_i y_i \phi(x_i)\right) + c\sum_{i=1}^{n} \varepsilon_i$$

$$+ \sum_{i=1}^{n} \alpha_i \left(1 - \varepsilon_i - y_i \left(\left(\sum_{j=1}^{n} \alpha_j y_j \phi(x_j)\right)^T \phi(x_i) + b\right)\right) - \sum_{i=1}^{n} \beta_i \varepsilon_i$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \left(\phi^T(x_i) \cdot \phi(x_j)\right) + \sum_{i=1}^{n} (c - \beta_i) \varepsilon_i$$

$$+ \sum_{i=1}^{n} \alpha_i (1 - \varepsilon_i) - \sum_{i=1}^{n} \alpha_i y_i \left(\sum_{j=1}^{n} \alpha_j y_j \phi^T(x_j)\right) \phi(x_i) - b \sum_{i=1}^{n} \alpha_i y_i$$

Now, using (2) and (3) we get

$$= \sum_{i=1}^{n} \alpha_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$= \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

where $K(x_i, x_j) = \phi^T(x_i) \cdot \phi(x_j)$

$$\therefore \boxed{L(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i, x_j)}$$

We can get $\alpha^* = \underset{\alpha}{\arg\max}\ L(\alpha)$

s.t. $0 \leq \alpha_i \leq C \quad \forall i = 1 \dots n$

$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

[redacted]

(5) Given : C : a convex set.

To show : $\theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C$

s.t. $x_1, x_2, x_3 \in C$

$\theta_1 + \theta_2 + \theta_3 = 1$

Sol : Since $x_1, x_2 \in C$

$$\therefore \frac{\theta_1 x_1}{\theta_1+\theta_2} + \frac{\theta_2 x_2}{\theta_1+\theta_2} \in C \quad \left(\because \theta x_1 + (1-\theta)x_2 \in C \text{ s.t. } x_1, x_2 \in C \right) \quad \text{--- (1)}$$

Now, since $x_3 \in C$, using (1) we get

$$(\theta_1+\theta_2)\left(\frac{\theta_1 x_1}{\theta_1+\theta_2} + \frac{\theta_2}{\theta_1+\theta_2} x_2\right) + (1-\theta_1-\theta_2)x_3 \in C$$

$$\Rightarrow \theta_1 x_1 + \theta_2 x_2 + (1-\theta_1-\theta_2)x_3 \in C$$

$$\Rightarrow \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C \quad (\because \theta_1+\theta_2+\theta_3 = 1)$$

Hence proved.