23-02-2022

Wednesday, February 23, 2022 2:05 PM

Adding d to x_0 will lead to point ~~z~~ z_0 in H_1

d → scalar (convert to vector)

$x_0 + d.\vec{u}$ — unit vector | $w \perp H_2$

$\therefore x_0 + d.\frac{w}{||w||}$ → why? — class lecture

$$\therefore z_0 = x_0 + d.\frac{w}{||w||} \quad \text{— (1)}$$

z_0 is a pt in H_1 $\therefore w z_0 + b = 1$ — (2)

$$\therefore w\left(x_0 + d.\frac{w}{||w||}\right) + b = 1$$

$$w x_0 + d.\frac{w.w}{||w||} + b = 1$$

→ sq of the norm

$$w x_0 + d\frac{||w||^2}{\cancel{||w||}} + b = 1$$

$$w x_0 + d||w|| + b = 1$$

$$\underbrace{w x_0 + b} + d||w|| = 1$$

$$-1 + d||w|| = 1$$

$$\therefore \left\{ d = \frac{2}{\|w\|} \right\}$$

to maximise the margin, we have to choose HP with smallest $\|w\|$

Objective function: minimise $\|w\|$ subject to the cond^n. that all +ve samples are above H_1 & −ve samples are below H_0

min $\|w\|$ s.t. $w x_i + b \geq 1 \ \forall x_i \in$ +ve class
& $w x_i + b \leq -1 \ \forall x_i \in$ −ve "

Min $\|w\|$ s.t. $y_i (w x_i + b) \geq 1$

for $y_i = 1$, $y_i (w x_i + b) \geq 1$
$1 (w x_i + b) \geq 1$

for $y_i = -1$, $-(w x_i + b) \geq 1$
$\therefore w x_i + b \leq 1$

obj. fn is min $\|w\|$ s.t. $y_i (w x + b) \geq 1 \ \forall i$

$\|w\|$ is not a convex func. & so it leads to local minimum

minimising $\|w\|$ is equivalent to minimising $\frac{1}{2} \|w\|^2$

Quadratic obj func.

convex & diff. ...

Quadratic obj func.
↓
convex & differentiable

Obj. fn. is $\underbrace{\min \frac{1}{2}\|w\|^2}_{\text{constrained obj func.}}$ s.t. $y_i(wx+b) \geq 1 \ \forall i$

$\|w\|^2$ = multiplying vector with itself

primal form of obj func →

$$\min L(w,b,\alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha_i \left(y_i (w^T x + b) - 1 \right) \quad (1)$$

$\alpha_i \geq 0 \ \forall i$

set of Lagrange multiplier ?

"Find support vectors closer to HP"
"Choose HP with max. dist. from those support vectors."

→ complex

→ Refer class as min wrt w, b max wrt α

∴ Convert to dual form

If w & b have to be min

$$\frac{\partial L}{\partial w} = 0 \quad \text{and} \quad \frac{\partial L}{\partial b} = 0$$

∴ i) $\frac{\partial L}{\partial w} = \frac{1}{\cancel{2}} \cdot \cancel{2} w - \sum_{i=1}^{N} \alpha_i y_i x_i + 0 = 0$

$$\Rightarrow \quad w = \sum_{i=1}^{N} \alpha_i y_i x_i \quad (2)$$

Optimal w is a linear combo of training exs, scaled by α

Optimal ω is a linear combo of training ex., scaled by α

ii) $$\frac{\partial L}{\partial \beta} = -\sum_{i=1}^{N} \alpha_i y_i = 0$$

$$\Rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0 \quad \text{—} ③$$

All Lagrange multiplier for +ve class should be = all LMs for -ve class

Check class recording for DA | LMs etc.