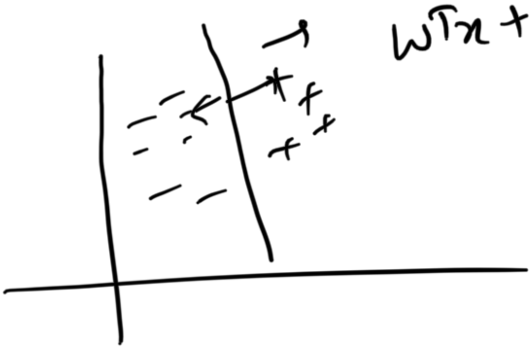
COL774

Machine Learning

Sep 15, 2021

Last Class:- Support Vector Machines (SVM)



$w^Tx + b = 0$

↳ Find the hyperplane which maximizes the min. margin

$\hat{\gamma}^{(i)}$:- functional margin

$\gamma^{(i)}$:- geometric margin

$$\max_{\gamma,} \ \gamma$$

$$\gamma \le \gamma^{(i)} \quad \forall i$$

$$\Downarrow$$

↳ SVM:-

$$\min_{w,b} \ \frac{1}{2} w^T w$$

$$y^{(i)}(w^T x^{(i)} + b) \ge 1$$

Convex Constrained optimization problem

↳

$$\min_{w} \ f(w) \rightarrow \text{convex} \qquad w \in \mathbb{R}^n$$

$$\text{convex} \leftarrow g_i(w) \le 0 \qquad i \in \{1 \ldots m\}$$

$$\text{affine} \leftarrow h_l(w) = 0 \qquad l \in \{1 \ldots p\}$$

Primal Problem

Constrained convex optimization problem.

Theory of Langrangeans:-

$$L(w, \alpha, \beta) = f(w) + \sum_{i=1}^{m} \alpha_i g_i(w) + \sum_{l=1}^{p} \beta_l h_l(w)$$

↳ Langrange multipliers $\alpha_i \geq 0$

w:- feasible:-] :- w satisfies the constraints in the primal problem.

$\max_{\alpha, \beta, \forall i: \alpha_i \geq 0} \mathcal{L}(w, \alpha, \beta) = ?$ $f(w)$

suppose:

if w is not feasible:- $h_i(w) \neq 0$ OR $g_i(w) \geq 0$

$\max_{\alpha, \beta, \forall i: \alpha_i \geq 0} \mathcal{L}(w, \alpha, \beta) = ?$ ∞

$\alpha \geq 0$

→ $f(w)$:- if w feasible
∞ :- o.w.

$$\Rightarrow \left[\min_{w} \max_{\alpha, \beta, \alpha \geq 0} \mathcal{L}(w, \alpha, \beta)\right]$$:- is equivalent to solving the primal problem

$$= \min_{w} \left[\theta_P(w)\right]$$

↳ primal objective.

primal problem.

$\theta_D(\alpha, \beta)$:- Dual objective:-

↳ Dual problem:-

$$\max_{\alpha, \beta, \alpha \geq 0} \left[\min_{w} \mathcal{L}(w, \alpha, \beta)\right]$$:- Dual problem

$$\equiv \max_{\alpha, \beta, \alpha \geq 0} \left[\theta_D(\alpha, \beta)\right]$$

$\min \theta_P(w) \geq \max \theta_D(\alpha, \beta)$] prove this

$\min_w$ Op(w) $\ge$ $\max_{\alpha,\beta;\ \alpha \ge 0}$ Dual(...)

p^* $\ge$ d^*

↳ optimal value of primal — ↳ optimal value of dual problem

What we will be interested in:- when is the case? that

$$[p^* = d^*]$$

Aside: 1 $\min_u \left[\max_v h(u,v)\right] \ge \max_v \min_u h(u,v)$

Let LHS: 1

let u^* be the optimal value of u

$$\left[\max_v h(u,v) = g(u)\right]$$

LHS:- $\min_u g(u)$:- Let u^* be optimal value:-

Consider

$$g(u^*) = \max_v h(u^*,v) \quad \text{— 1}$$

2 — $h(u^*,v) \ge \min_u h(u,v)$] $\cancel{g(u^*) \ge \min_u g(u)}$

3 — $\max_v [h(u^*,v)] \ge \max_v \left[\min_u h(u,v)\right]$

4 — $\min_u \max_v h(u,v) \ge \max_v \min_u h(u,v)$

↳ $\min_{w} \theta_P(w) \geq \max_{\substack{\alpha,\beta \\ \alpha \geq 0}} \theta_D(\alpha,\beta)$

weak duality ⇒ $\theta_P(w) \geq \theta_D(\alpha,\beta)$ $\forall (w,\alpha,\beta)$, $\alpha \geq 0$

under what conditions does above inequality become equality?] Strong duality

↳ ① Primal problem is convex.] AND

② Slater's conditions are satisfied.

↳ $\exists w$: $g_i(w) < 0$ $\forall i$:- [$h_e(w) = 0$ $\forall e$]

↳ feasible

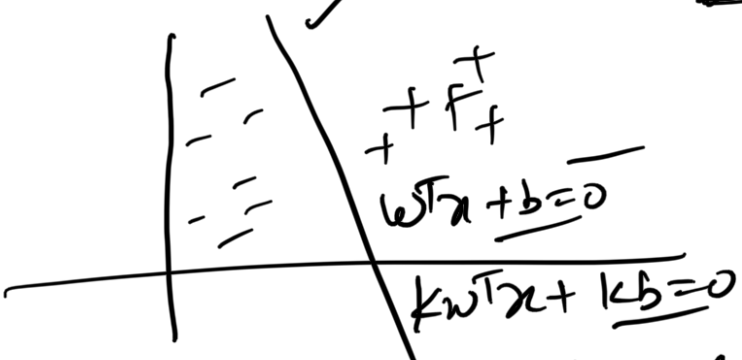
⇒ SVM:- $1 - y^{(i)}(w^T x^{(i)} + b) \leq 0$ } $g_i(w) \leq 0$

$\equiv y^{(i)}(w^T x^{(i)} + b) \geq 1$

? $\exists w, b$, $y^{(i)}(w^T x^{(i)} + b) > 1$ $\forall i$:-

LHS:- +ve

↳ if hyperplane separates data



↳ if strong duality holds:-

① Primal problem is convex

② Slater's conditions are satisfied:-

primal optimal — dual optimal

$\exists\, w^*, \alpha^*, \beta^*$ s.t

$\theta_P(w^*) := p^*$

$\theta_D(\alpha^*, \beta^*) = d^*$

$p^* = d^*$

$\underbrace{L(w^*, \alpha^*, \beta^*)}_{\Updownarrow} = \theta_P(w^*) = \theta_D(\alpha^*, \beta^*)$

KKT Conditions are satisfied

w, α, β