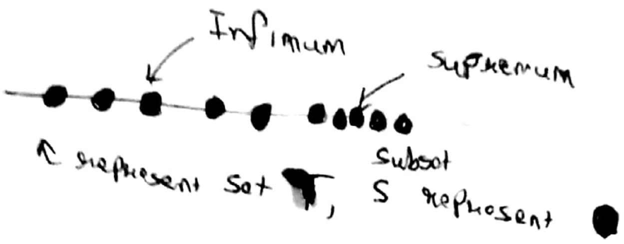
Date: 30.05.2018 (10.1)

Geometric interpretation of Primal-dual problem ?

Infimum / Supremum and Geometric interpretation of lagrange duality ?



Infimum: Infimum of a Subset S of a Partially ordered Set T is the greatest element in T that is less than or equal to all elements of S. Also called greatest lower bound.

Supremum: Supremum of a Subset S of a Paritialy ordered Set T is the least element in T that is greater than or equal to all elements of S. Also called least upper bound.

Scanned by CamScanner

Geometric interpretation of Primal-dual Problem:

Primal Problem:

$$\min_{x} f(x)$$

Subject to:

$g_i(x) \leq 0 \quad \forall i \in \{1, 2, \ldots, m\}$

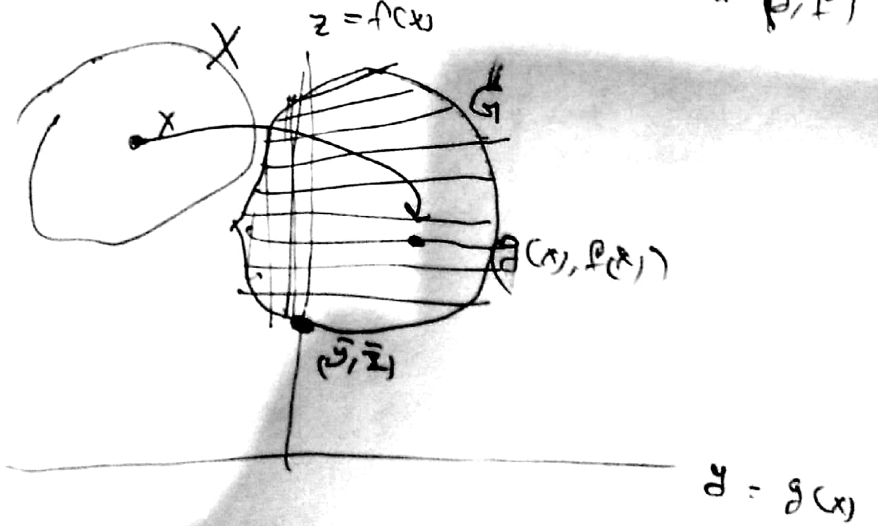
$h_i(x) \leq 0 \quad \forall i \in \{1, 2, \ldots, m\}$

where: $f : \mathbb{R}^n \to \mathbb{R}$

$g : \mathbb{R}^n \to \mathbb{R}$

Define. $G = \{(y, z) \mid y = g(x), z = f(x) \text{ for some } x \in X\}$

$G \in \mathbb{R}^2$, G is the image of X under (g, f) map



Primal Problem says following $g(x) \leq 0$ find the value of $f(x)$ that is minimum. Such Point is $(\bar{y}, \bar{z})$

Lagrangian dual subproblem

$$\theta_D(u) = \operatorname{Inf}\{f(x) + u g(x) \mid x \in X\}$$

$$\max_u \; \theta_D(u)$$

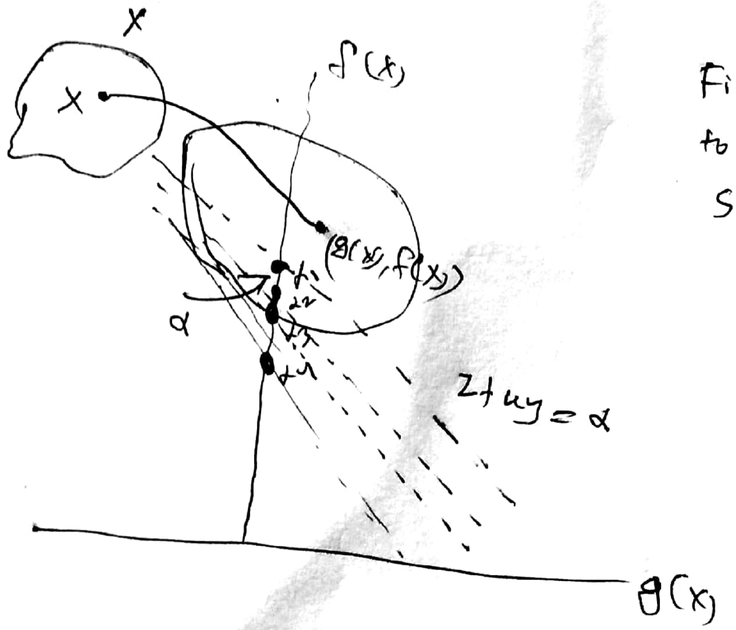
Subject to $u \geq 0$

$\operatorname{Inf}\{f(x) + u g(x) \mid x \in X\}$??

$\operatorname{Inf}\{\cancel{z} + uy = \alpha \mid x \in X\}$

minimize $z + uy = \alpha$ ~~Such that~~ over Points (y, z) in G.

$z + uy = \alpha \rightarrow$ equation of a straight line with slope $-u$ and Intercept α



Find a straight line Parallel to $z + uy = \alpha$ that has still in G.

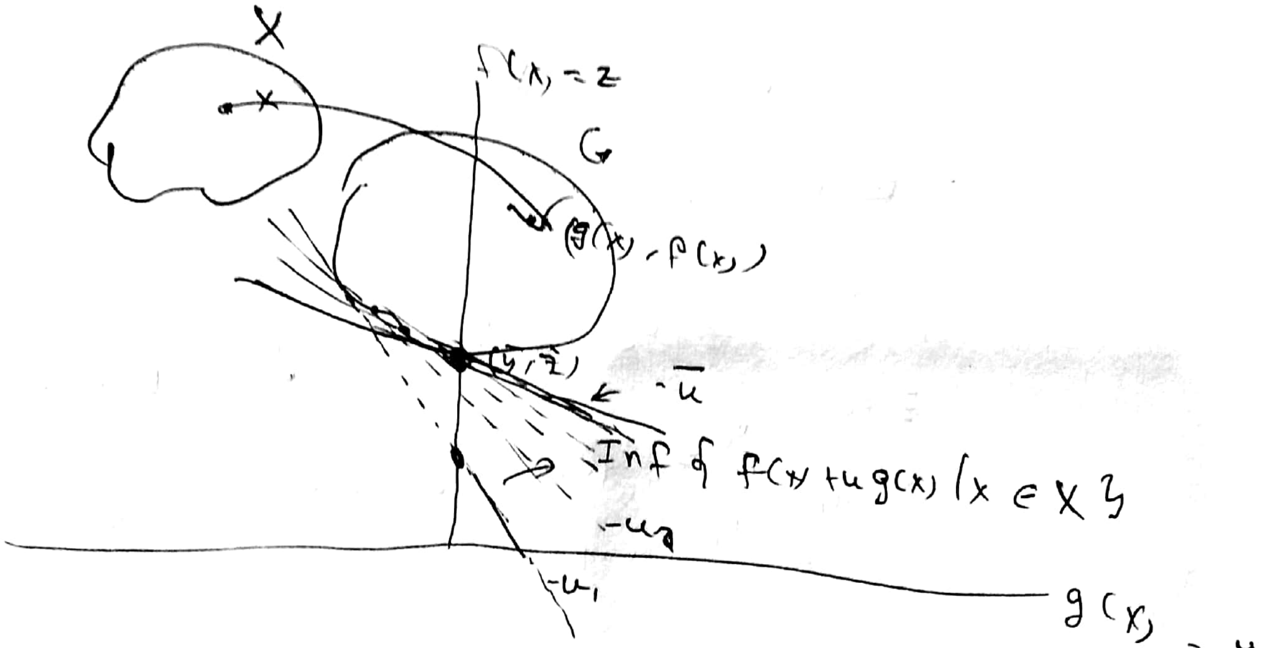
$$\operatorname{Inf}\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\} = \alpha_4$$

Finally, $\max_u \operatorname{Inf}\{f(x) + u g(x) \mid x \in X\}$, to solve dual Problem, find line with ~~-ve~~ slope $-u$ $(u \geq 0)$ s.t

Scanned by CamScanner

last intercept on z axis, $\theta(u)$ is maximal.

The solution to the dual Problem is $\bar{u}$ and optimal ~~[illegible]~~ dual objective occurs at $(\bar{y}, \bar{z})$ $\therefore$ same as Primal Problem.



Scanned by CamScanner

Date: 28-05-2018

Support vector machine primal-dual optimization problem?

$$\min_{\omega} \frac{1}{2}\|\omega\|^2$$

s.t $y^{(i)}(\omega^T x^{(i)}+b) \geq 1 \quad \forall i \in \{1,2,\ldots,m\}$

$g_i(\omega, b)$

Rewrite:

$$\min_{\omega} \frac{1}{2}\|\omega\|^2$$

s.t $-y^{(i)}(\omega^T x^{(i)}+b)+1 \leq 0 \quad ,\forall i \in \{1,2,\ldots,m\}$

Lagrangian: $\mathcal{L}(\omega, b, \alpha) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^{m} \alpha_i \left(y^{(i)}[\omega^T x^{(i)}+b]-1\right)$ — ⓪

Write in the form of dual problem:

$$\theta_D(\alpha) = \min_{\omega,b} \mathcal{L}(\omega, b, \alpha)$$

$\nabla_\omega \mathcal{L}(\omega,b,\alpha) = 0, \quad \|\omega\| - \sum_{i=1}^{m} \alpha_i (y^{(i)} x^{(i)}) = 0$

$$\omega = \sum_{i=1}^{m} \alpha_i y^{(i)} x^{(i)} \quad — ①$$

$\nabla_b \mathcal{L}(\omega,b,\alpha) = 0, \quad -\sum_{i=1}^{m} \alpha_i y^{(i)} = 0, \quad \sum_{i=1}^{m} \alpha_{(i)} y^{(i)} = 0$

Put eqn ① into eqn ⓪

$$\theta_D(\alpha) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i y^{(i)} \alpha_j y^{(j)} \langle x^{(i)}, x^{(j)}\rangle - \sum_{i=1}^{m} \alpha_i \left(y^{(i)}[\omega^T x^{(i)}+b] - 1\right)$$

Scanned by CamScanner

$$\theta_D(\alpha) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y^{(i)} y^{(j)} \langle x^{(i)}, x^{(j)} \rangle - \sum_{i=1}^{m} \alpha_i \left(y^{i}(\omega^T x^{(i)} + b) - 1\right)$$

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

$$- \cancel{\sum_{i=1}^{m} \alpha_i y^{(i)} b} \to 0 + \sum_{i=1}^{m} \alpha_i$$

because $\sum_{i=1}^{m} \alpha_i y^{(i)} = 0$ from $\nabla_b \mathcal{L}(\omega, b, \alpha) = 0$

$$= \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y^{(i)} y^{(j)} \langle x^{(i)}, x^{(j)} \rangle \rightarrow \text{Dot Product}$$

$$\boxed{\theta_D(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y^{(i)} y^{(j)} \langle x^{(i)}, x^{(j)} \rangle}$$

Now dual Problem becomes:

$$\boxed{\begin{aligned} &\max_{\alpha}\ \theta_D(\alpha) \\ &\text{s.t.}\ \ \alpha_i \geq 0 \quad \forall i \in \{1, 2, \ldots, m\} \\ &\qquad \sum_{i=1}^{m} \alpha_i y^{(i)} = 0 \end{aligned}}$$

Try to $\max_{\alpha} \theta_D(\alpha)$. Knowing α^*, we can compute ω, b as.

$$\omega^* = \sum_{i=1}^{m} \alpha_i^* x^{(i)} y^{(i)}$$

$$b = \frac{-\max_{i, y^{(i)}=-1} \omega^{*T} x^{(i)} + \min_{i, y^{(i)}=1} \omega^{*T} x^{(i)}}{2}$$

Scanned by CamScanner

ediction:

Given new ~~train~~ example x_{new}

$$h_{w,b}(x\_new) = g(w^T x\_new + b)$$

$$w^T x\_new + b = \left(\sum_{i=1}^{m} \alpha_i y^{(i)} x^{(i)}\right)^T x\_new + b$$

$$= \sum_{i=1}^{m} \alpha_i y^{(i)} \langle x^{(i)}, x\_new \rangle + b$$

α_i will be All zero's except for Support vectors ($\alpha_i g_i(w) = 0$
$\alpha_i \neq 0$ / $g_i(w) = 0$
functional margin $= 1$)

So, we need to find inner product b/w Support vectors and New x (x_new

By Examining dual optimization Problem, we gained significant insight into the Structure of the Problem. and, we can able to write entire algorithm in terms of only inner products b/w input feature vectors.

Scanned by CamScanner