Lecture 20 (Lagrangian)

1 Formulating the Problem

$$L(w, \alpha, \beta) = f(w) + \sum_{i=1}^{m} \alpha_i g_i(w) + \sum_{l=1}^{k} \beta_l h_l(w), \alpha_i \geq 0$$

Consider a point w which is *feasible*, then

$$\max_{\alpha,\beta}(L(w, \alpha, \beta)) = f(w)$$

Also, consider a point w which is not feasible, then

$$\max_{\alpha,\beta}(L(w, \alpha, \beta)) = \infty$$

Therefore, the *primal* problem can be written as solving:

$$\min_{w}(\max_{\alpha,\beta,\alpha\geq 0}(L(w, \alpha, \beta)))$$

$$\implies \min_{w}[\theta_P(w)]$$

where, $\theta_p(w)$ is the *primal* objective and the entire problem is called the *primal* problem. The constraints have been made simpler and have been absorbed. The *dual* problem is given as:

$$\max_{\alpha,\beta,\alpha\geq 0}[\min_{w} L(w, \alpha, \beta)]$$

$$\implies \max_{\alpha,\beta,\alpha\geq 0}[\theta_D(\alpha, \beta)]$$

The relation between $\theta_P(w)$ and $\theta_D(\alpha, \beta)$ is given as:

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

$$p^* \geq d^*$$

We are interested in finding the condition when $p^* = d^*$ so that we can solve the problem easily.

2 (Sufficient) Conditions for Strong Duality ($p^* = d^*$)

1. Primal problem is convex, and
2. Slaters conditions are satisfied
   $\exists w : g_i(w) < 0 \ \forall i \in \{1, \dots, m\}$ and $h_l(w) = 0 \ \forall l \in \{1, \dots, k\}$

The 2^{nd} condition will be true for SVMs if the data is linearly separable (this is where the concept of support vector comes)