Machine Learning Homework 4

He Li, 2014012684

2017/10/31

Problem. *Given a matrix* $A = (a_{ij})_{n\times m}$*, show that* $\min_{i\le n}\max_{j\le m} a_{ij} \ge \max_{i\le m}\min_{j\le n} a_{ji}$.

Proof. Denote $\min_{i\le n}\max_{j\le m} a_{ij} = a_{pq}$, $\max_{i\le m}\min_{j\le n} a_{ji} = a_{rt}$. Then

$$a_{pq} \ge a_{rq} \ge a_{rt}$$

Therefore,

$$\min_{i\le n}\max_{j\le m} a_{ij} \ge \max_{i\le m}\min_{j\le n} a_{ji}$$

□

Problem. *Show that KKT conditions are necessary and if* f, g_i *are convex and each* h_i *is linear then it's also sufficie for* (X^*, λ^*, μ^*) *to be the optima of primal and dual programmings.*

Proof. First we show that KKT conditions are necessary. Denote (x^*, λ^*, μ^*) the solution of primal and dual problems. Obviously, it satisfies (1), (2) and (3) of KKT conditions. For (4), consider the primal problem, if $g_i(x^*) < 0$, then $\lambda_i(x^*)$ should equal to 0 in order to max the primal function. Therefore, $\lambda_i g_i(x^*) = 0$.

Then we show that with the additional condition that $f(x), g_i(x)$ are convex and $h_i(x)$ is linear, KKT condition is sufficient.

First, a x_0 satisties the above conditions also satisfies the constraints in primal and dual problems. Denote $x1$ the solution of primal problem and x_2 the solution of dual problem, deonte $L(x,\mu,\lambda) = f(x) + \sum_i \mu_i h_i(x) + \sum_i \lambda_i g_i(x)$.

With the additional condition, $L(x,\mu,\lambda)$ is a convex function w.r.t x, then $x_0 = \arg\min_x L(x,\mu,\lambda)$. Therefore,

$$L(x_0,\mu,\lambda) \le L(x_2,\mu,\lambda) \le L(x_1,\mu,\lambda) \tag{1}$$

Note that x_0 satisfies primal constraints and x_2 is the argmin of primal problem. Therefore,

$$L(x_1,\mu,\lambda) \le L(x_0,\mu,\lambda) \tag{2}$$

With (1) and (2) we have our conclusion that

$$L(x_0,\mu,\lambda) = L(x_2,\mu,\lambda) = L(x_1,\mu,\lambda)$$

□