KKT Conditions and Lagrangian Duality

2022年5月10日 11:46

Recall the setting:

Let $f, g_1, \cdots, g_{m_1}, h_1, h_2, \cdots, h_{m_2} : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable functions. Consider

$$
(Q) \qquad \begin{aligned} \inf \quad & f(x) \\ \text{s.t.} \quad & \left.\begin{aligned} & g_i(x) \le 0, \quad i = 1, \cdots, m_1, \quad (v_i) \\ & h_j(x) = 0, \quad j = 1, \cdots, m_2, \quad (w_j) \\ & x \in \mathbb{R}^n. \end{aligned}\right\} S : \text{feasible region} \end{aligned}
$$

Q: When do the KKT conditions also sufficient for optimality?

A: Consider convexity of (Q).

Theorem (KKT Sufficient optimality conditions)

Suppose that in (Q), $f, g_1, \cdots, g_{m_1}$ are convex; $h_1, \cdots, h_{m_2}$ are affine. Suppose that $(\bar{x}; \bar{v}, \bar{w})$ satisfies

(a) $\quad g_i(\bar{x}) \le 0 \quad \forall i \;;\; h_j(\bar{x}) = 0 \quad \forall j \qquad$ (primal feasibility)

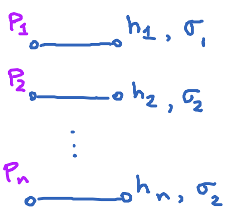
(b) $\quad \nabla f(\bar{x}) + \sum_{i=1}^{m_1} \bar{v}_i \nabla g_i(\bar{x}) + \sum_{j=1}^{m_2} \bar{w}_j \nabla h_j(\bar{x}) = 0, \qquad$ (dual feasibility)

$\qquad \bar{v} \geq 0,$

(c) $\quad \bar{v}_i \, g_i(\bar{x}) = 0 \quad \forall i \qquad$ (complementarity)

Then, $\bar{x}$ is a global minimum of (Q).

Example: (Power Allocation)



$h_i \stackrel{>0}{:}$ channel gain

$\sigma_i \stackrel{>0}{:}$ Gaussian noise power

P_i : allocated power

Goal: To allocate power to each channel s.t. the information rate is maximized.

$$\max \sum_{i=1}^{n} \ln\left(1+\frac{h_i p_i}{\sigma_i}\right) \longleftrightarrow \min \ -\sum_{i=1}^{n} \ln\left(1+\frac{h_i p_i}{\sigma_i}\right)$$

$$\text{s.t.} \quad \sum_{i=1}^{n} p_i \le \underline{P}. \leftarrow \text{given} \qquad (v_0)$$

$$p_i \ge 0 \quad \forall i. \qquad (v_i) \longleftrightarrow -p_i \le 0$$

Exercise: $p \mapsto \sum_{i=1}^{n} \ln\left(1+\frac{h_i p_i}{\sigma_i}\right)$ is concave

Remark: Since this is a linearly constrained convex optimization problem, KKT conditions are both necessary and sufficient.

$$\frac{-h_i/\sigma_i}{1+\frac{h_i p_i}{\sigma_i}} + v_0 - v_i = 0 \qquad (*)$$

$$v_0\left(\sum_{i=1}^{n} p_i - P\right) = 0, \quad v_i p_i = 0, \quad \begin{matrix} v_0 \ge 0, \\ v_i \ge 0 \end{matrix} \quad \forall i. \qquad (**)$$

1°: From $(*)$: $v_0 > 0$. $\overset{(**)}{\Longrightarrow} \sum_{i=1}^{n} p_i = P$.

2°: Solving p_i from $(*)$: $p_i = \frac{1}{v_0 - v_i} - \frac{\sigma_i}{h_i}$. — $(***)$

From $(**)$: If $p_i > 0$, then $v_i = 0$. $\Rightarrow p_i = \frac{1}{v_0} - \frac{\sigma_i}{h_i} > 0$.

If $p_i = 0$, then from $(***)$, $\frac{1}{v_0} - \frac{\sigma_i}{h_i} \le 0$.

Hence, we can write

$$(\Delta) \ — \qquad p_i = \left(\frac{1}{v_0} - \frac{\sigma_i}{h_i}\right)_+, \qquad (x)_+ = \max\{x, 0\}.$$

waterfilling solution

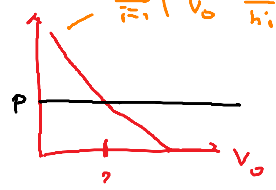
3°: From 1°:

single-variable nonlinear equation

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \left(\frac{1}{v_0} - \frac{\sigma_i}{h_i}\right)_+ = P$$

unknown

$v_0 \mapsto \sum_{i=1}^{n} \left(\frac{1}{v_0} - \frac{\sigma_i}{h_i}\right)_+$



v_0 can be found by a bisection method

Then, we can recover the optimal $\{P_i\}$ from (Δ)

Lagrangian Duality

Consider

$$(Q) \quad v_q^* = \inf f(x) \quad \text{s.t. } G(x) \le 0, \ (v) \quad H(x) = 0, \ (w) \quad x \in X \subseteq \mathbb{R}^n$$

$G(x) = (g_1(x), \dots, g_{m_1}(x)) \in \mathbb{R}^{m_1}$

$H(x) = (h_1(x), \dots, h_{m_2}(x)) \in \mathbb{R}^{m_2}$

X : arbitrary, non-empty set.

Q. How to construct the dual of (Q)?

Idea: Use a penalty function approach.

Specifically, define

$$L(x; v, w) = f(x) + v^T G(x) + w^T H(x)$$

to be the Lagrangian function associated with (Q). Now, for a fixed $x \in X$,

$$\sup_{\substack{v \ge 0 \\ w}} \{ \underbrace{f(x) + v^T G(x) + w^T H(x)}_{L(x;v,w)} \} = \begin{cases} f(x) & \text{if } G(x) \le 0, H(x) = 0 \\ +\infty & \text{otherwise} \end{cases}$$

Hence,

$$(Q) \iff \inf_{x \in X} \sup_{\substack{v \ge 0 \\ w}} L(x; v, w)$$

Then, we can define the dual of (Q) as

$$(D): \quad v_d^* = \sup_{\substack{v \ge 0 \\ w}} \underbrace{\inf_{x \in X} L(x; v, w)}_{\theta(v,w)} \quad \text{(interchange inf and sup)}$$

Theorem (Weak Duality)

Let $\bar{x}$ be feasible for (Q) and $(\bar{v}, \bar{w})$ be feasible for (D)

Then, $\theta(\bar{v}, \bar{w}) \leq f(\bar{x})$.

Example: Consider

$$v_p^* = \min \; -x$$

$$\text{s.t.} \quad x \leq 1 \qquad (v)$$

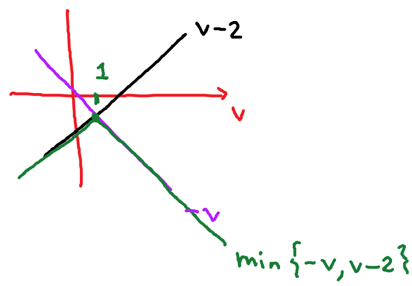
$$x \in X = \{0, 2\} \subseteq \mathbb{R}$$

Clearly, $v_p^* = 0$, attained by $x^* = 0$.

The dual is

$$v_d^* = \sup_{v \geq 0} \underbrace{\inf_{x \in \{0,2\}} \{ -x + v(x-1) \}}_{\theta(v)}$$

Observe that for any $v \geq 0$,



$$\theta(v) = \min\{-v, v-2\}$$

Hence, $v_d^* = \sup_{v \geq 0} \min\{-v, v-2\} = -1$,

attained by $v^* = 1$