Getting the Primal Solution

Theorem u, v given, $u \geq 0$.

Let $\bar{x}$ be an optimal solution to $\min_{x \in X} L(x, u, v)$

Then $\bar{x}$ is an optimal solution to

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq g_i(\bar{x}) \qquad i \in I := \{i : u_i > 0\} \\ & h_j(x) = h_j(\bar{x}) \qquad j = 1, \dots, p \\ & x \in X \end{aligned}$$

"Proof" Clearly $\bar{x}$ feasible.

Also, trying to improve on f over $\bar{x}$, worsens $g_i(x)$ for some $i \in I$:

$\forall x \in X$, $f(x) + u^T g(x) + v^T h(x)$

$$\geq f(\bar{x}) + u^T g(\bar{x}) + v^T h(\bar{x}) \quad \text{by defn of } \bar{x}.$$

Also, $u^T g(x) \leq u^T g(\bar{x})$, $v^T h(x) = v^T h(\bar{x})$, if x feas. for problem in theorem.

$\therefore f(x) \geq f(\bar{x})$.

//

Corollary.

Under assumptions of theorem, suppose $g(\bar{x}) \le 0$, $h(\bar{x}) = 0$, $u^T g(\bar{x}) =$

Then $\bar{x}$ is optimal for the primal problem, and (u, v) is opt for dual.

Proof. ~~$\bar{x}$ is optimal for~~ (Follows from weak duality: $\bar{x}$ is feasible and $f(\bar{x}) = L(\bar{x}, u, v) = \theta(u,v)$.)

$$\cancel{\begin{aligned} \min\ & f(x) \\ & g_i(x) \le g_i(\bar{x}) = 0 \quad i \in I \\ & h(x) = h \end{aligned}}$$

$u^T g(\bar{x}) = 0 \quad \therefore \quad u_i > 0 \Rightarrow g_i(\bar{x}) = 0$, since $u \ge 0$, $g(\bar{x}) \le 0$

$\therefore$ $\bar{x}$ optimal for

$$\begin{aligned} \min\ & f(x) \\ & g_i(x) \le g_i(\bar{x}) = 0 \quad \forall i \in I \\ & h_j(x) = h_j(\bar{x}) = 0 \\ & x \in X \end{aligned}$$

$\therefore$ $\bar{x}$ optimal for

$$\begin{aligned} \min\ & f(x) \\ & g(x) \le 0 \\ & h(x) = 0 \\ & x \in X. \end{aligned}$$

Also, $f(\bar{x}) = f(\bar{x}) + \bar{u}^T g(\bar{x}) + \bar{v}^T h(\bar{x}) \le \theta(u, v)$ $\therefore (u, v)$ solves dual //.

<u>Theorem</u> (Absence of duality gap $\Rightarrow$ optimal soln satisfies comp slackness).

$\bar{x}$ primal optimal, $\bar{u}, \bar{v}$ dual optimal, $f(\bar{x}) = \theta(\bar{u}, \bar{v})$

Then $\bar{u}^T g(\bar{x}) = 0$, $\bar{x} \in X(\bar{u}, \bar{v})$

<u>Proof</u>

$$f(\bar{x}) + \bar{u}^T g(\bar{x}) + \bar{v}^T h(\bar{x}) \cancel{\geq \inf \{ f(\bar{x}) + \bar{u}^T g(\bar{x})}$$
$$\geq \inf \{ f(x) + \bar{u}^T g(x) + \bar{v}^T h(x) : x \in X \}$$
$$= \theta(\bar{u}, \bar{v}) = f(\bar{x})$$

Now $h(\bar{x}) = 0$, $\bar{u} \geq 0$, $g(\bar{x}) \leq 0$

$\therefore$ must have $\bar{u}^T g(\bar{x}) = 0$.

Also, $\cancel{f(\bar{x}) + \bar{u}^T g(\bar{x}) + \bar{u}}$

have equality above, so $\bar{x} \in X(\bar{u}, \bar{v})$. //

(So $\exists$ optimal solution to primal problem in $X(\bar{u}, \bar{v})$).

Interested in directions along which θ increases.

Defn A vector d is called an ascent direction of θ at w if there exists a $\delta > 0$ such that

$$\theta(w+\lambda d) > \theta(w) \qquad \forall \lambda \in (0, \delta)$$

STEEPEST ASCENT DEFN.

~~If θ concave, d is an ascent direction iff~~ $\phi : \mathbb{R} \to \mathbb{R}$

Let $\phi_d(\lambda) = \theta(w+\lambda d)$. ~~Then d is an ascent direction~~

If θ concave, d is an ascent direction iff $\phi_d'(0) > 0$.

Theorem $X \subseteq \mathbb{R}^n$, $X \neq \phi$ compact, $\phi_{d_u,d_v}(\lambda) = \theta(\bar{u}+\lambda d_u, \bar{v}+\lambda d_v)$

Then $\phi_{d_u,d_v}'(0) \geq d_u^T g(\bar{x}) + d_v^T h(\bar{x})$ for some $\bar{x} \in X(\bar{u}, \bar{v})$

(One of these points is vital, no points not in $X(\bar{u},\bar{v})$ matter infinitesimally)

Corollary ~~$X \in \mathbb{R}, X$~~

$$\phi_{d_u,d_v}'(0) = \inf \left\{ d_u^T \xi_u + d_v^T \xi_v : (\xi_u, \xi_v) \in \underbrace{\partial\theta(\bar{u},\bar{v})}_{\text{Set of subgradients of } \theta \text{ at } (\bar{u},\bar{v})} \right\}$$

(minimizing linear function over a convex set, where extreme points are subgradients corresponding to $x \in X(\bar{u},\bar{v})$.)

Defn A vector $\bar{d}$ is called a direction of steepest ascent of θ at w if
$$\phi'_{\bar{d}}(0) = \underset{\|d\| \leq 1}{\text{maximum}} \ \phi'_d(0).$$

Theorem $X \subseteq \mathbb{R}^n$, $X \neq \emptyset$, X compact. f, g, h continuous.
Let $(\bar{\xi}_u, \bar{\xi}_v) \in \partial\theta(u,v)$ with smallest norm.
Then the direction of steepest ascent $\bar{d}_u, \bar{d}_v$ of θ at $\bar{u}, \bar{v}$ is
$$\bar{d} = \begin{cases} 0 & \text{if } \bar{\xi}_{\bar{u}}, \bar{\xi}_{\bar{v}} = 0 \\ \dfrac{\bar{\xi}_{\bar{u}}, \bar{\xi}_{\bar{v}}}{\|\bar{\xi}_{\bar{u}}, \bar{\xi}_{\bar{v}}\|} & \text{if o/w.} \end{cases}$$

($\bar{\xi}_{\bar{u}}, \bar{\xi}_{\bar{v}} = 0$ case is clear: already at maximum.
O/w: make same inner product with all subgradients, or nearest face, so move all points same (infinitesimally)).

(NB: if $0 \notin \partial\theta(\bar{u},\bar{v})$, then can separate 0 from $\partial\theta(\bar{u},\bar{v})$, so $\exists$ ascent direction).

Gradient method for solving dual problem

(Will assume subproblems have unique solutions).

0: Initialization: Choose vector u^1, v^1 with $u^1 \geq 0$. Let $k=1$.

1. Solve the subproblem:

$$\min_{x \in X} \; f(x) + u^{k^T} g(x) + v^{k^T} h(x)$$

(= find $\theta(u, v)$)

Let x^k be the unique optimal solution.

2. Define

$$\hat{g}_i(x^k) = \begin{cases} g_i(x^k) & \text{if } u^k_i > 0 \\ \text{maximum}\{0, g_i(x^k)\} & \text{if } u^k_i = 0. \end{cases}$$

(So dirn is feasible).

3. If $(\hat{g}(x^k), h(x^k)) = 0$, STOP: u_k, v_k is optimal solution

4. Solve the problem

$$\max \; \theta\left[(u^k, v^k) + \lambda(\hat{g}(x^k), h(x^k))\right]$$
$$\text{s.t.} \quad u^k + \lambda \hat{g}(x^k) \geq 0$$
$$\lambda \geq 0.$$

Let λ_k be an optimal solution

5. $u^{k+1} \leftarrow u^k + \lambda_k \hat{g}(x^k)$

$v^{k+1} \leftarrow v^k + \lambda_k h(x^k)$

What if problem in 4 has no solution, ie λ is unbdd?

~~If~~ Two cases: i) obj fn is unbdd: then primal is infeasible

ii) obj fn bounded: then just take some large λ_k.

~~Example~~ ~~min~~ $\cancel{x_1^2 + x_2^2}$

When solution to problem in part 1 is not unique,

algo is more complicated because of step 2.

Two questions regarding $(\hat{g}(x^k), h(x^k))$:

1) Why is u^k, v^k optimal if $(\hat{g}(x^k), h(x^k)) = 0$?

(Know u^k, v^k optimal if $(g(x^k), h(x^k)) = 0$.)

KKT conditions for $\max \theta(u,v)$

$$\text{s.t. } u \geq 0 :$$

~~are $\nabla_u \theta(\bar{u},\bar{v}) - b = 0$, $\nabla_v \theta(\bar{u},\bar{v}) = 0$~~

~~$b^T u = 0$ for some vector $b \geq 0$~~

~~Take $b = -\hat{g}(x^k)$:~~

~~$\nabla_u \theta(\bar{u},\bar{v}) = g(x^k)$~~

Equiv to

$$\min -\theta(u,v)$$
$$\text{s.t. } -u \leq 0$$

KKT conditions:

$$-\nabla_u \theta(\bar{u},\bar{v}) - b = 0$$
$$-\nabla_v \theta(\bar{u},\bar{v}) = 0$$
$$b^T u = 0, \quad b \geq 0 \quad (b \text{ variable}).$$

Now
$$\nabla_u \theta(\bar{u},\bar{v}) = g(x^k)$$
$$\nabla_v \theta(\bar{u},\bar{v}) = h(x^k)$$

Take $b = -g(x^k)$, so clearly satisfy ∇ cond, and $b \geq 0$

Also, $u_i > 0 \Rightarrow g(x^k) = \hat{g}(x^k) = 0$, so b does satisfy KKT.

2) If $(\hat{g}(x^k), h(x^k)) \neq 0$, why is it an ~~descent~~ ascent direction?

$$
\begin{aligned}
&(\hat{g}(x^k), h(x^k))^T \nabla \theta(\bar{u}, \bar{v}) \\
&\quad = (\hat{g}(x^k), h(x^k))^T (g(x^k), h(x^k)) \\
&\quad = \| \hat{g}(x^k), h(x^k) \|_2^2 > 0.
\end{aligned}
$$

Note $\hat{g}(x^k), h(x^k)$ feasible by construction.