Duality & Reductions

Monday, January 27, 2020 10:35 PM

<u>ORACLES</u> for a convex set K

MEM(x) : YES if $x \in K$

NO o.w.

SEP(x) : YES if $x \in K$

NO o.w, and c : $c^T y < c^T x \;\; \forall y \in K$.

VAL(c) : OUTPUTS $\max_{x \in K} c^T x$

or "K is EMPTY"

OPT(c) : $\underset{\in K}{x}$ s.t. $c^T x \le c^T y \quad \forall\, x \in K$

or "K is EMPTY".

Cutting Plane Method: $OPT_K \longrightarrow SEP_K$.

SEP is stronger than MEM

SEP is stronger than MEM

OPT ———————— VAL.

For different problems, different oracles can be more convenient/efficient.

E.g. $K = \{x : Ax \geq b\}$ $\qquad SEP_K$ is easy – check all constraints

$K =$ Conv Hull $\{a_1, \ldots, a_m\}$ $\qquad OPT_K$ is easy – $\sup_i c^T a_i$

Q. Are these fundamentally equivalent?

ORACLES for convex functions.

$EVAL_f(x)$: $\quad f(x)$.

$GRAD_f(x)$: $f(x), g$ s.t. $\forall y \quad f(y) \geq f(x) + g^T(y-x)$.

(g is a subgradiant of f at x).

Recall $\quad \delta_K(x) = \begin{cases} 0 & x \in K \\ \infty & x \notin K \end{cases}$ $\qquad$ convex.

$MEM_K \equiv \delta_K$.

A useful (and important) concept.

A useful (and important) concept.

Dual of a convex function $f : \mathbb{R}^n \to \mathbb{R}$.

$$f^*(\theta) = \sup_{x \in \mathbb{R}^n} \theta^T x - f(x) \qquad \forall \theta \in \mathbb{R}^n.$$

Note: f^* is max of affine functions (one per x)

so f^* is convex.

- $f^*(0) = -\inf_x f(x)$.

- $\delta_K^*(c) = \sup_x c^T x - \delta_K(x)$

$$= \sup_{x \in K} c^T x$$

$$\boxed{\mathrm{EVAL}_{\delta_K^*} \equiv \mathrm{VAL}_K.}$$

Lemma $\nabla f^*(\theta) = \arg\max_x \theta^T x - f(x)$

Pf. $x_\theta = \arg\max_x \theta^T x - f(x)$.

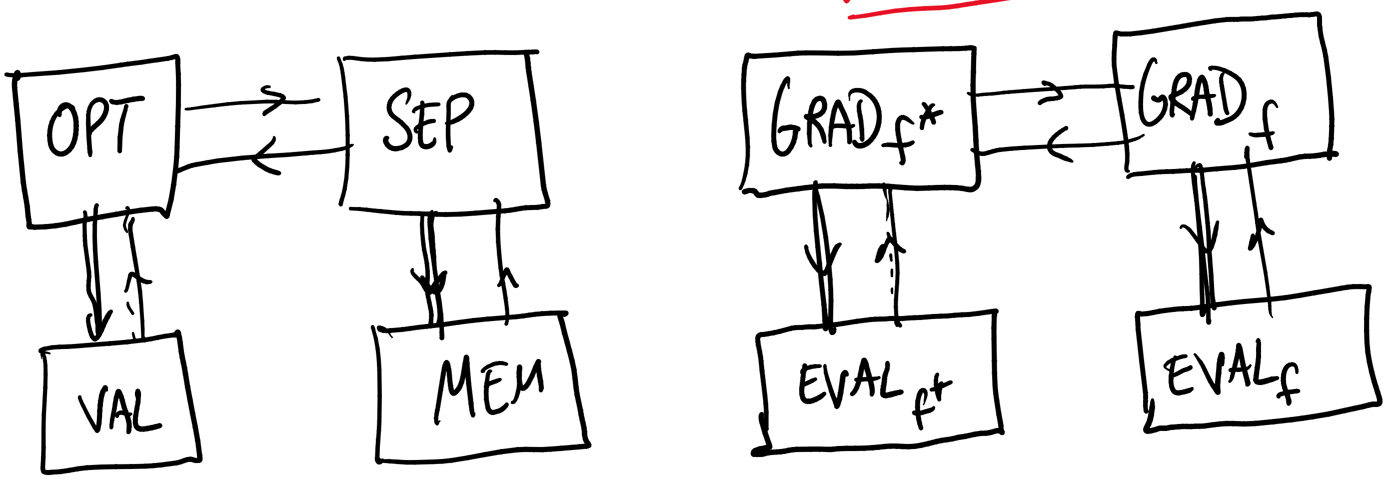
$$f^*(\theta) = \theta^T x_\theta - f(x_\theta)$$

$\forall \eta \quad f^*(\eta) \geq \eta^T x_\theta - f(x_\theta)$

$\Rightarrow \quad f^*(\eta) - f^*(\theta) \geq x_\theta^T (\eta - \theta)$

$\Rightarrow \quad x_\theta \in$ subgrad.(f^*).

$$\boxed{\mathrm{GRAD}_{\delta^*_K} \equiv \mathrm{OPT}_K.}$$



Thm. convex f, Epi(f) closed, then $f^{**} = f$.

Pf. $\mathrm{Epi}(f) = \{(x,t) : f(x) \leq t\}$ is a convex set.

So it is an intersection of halfspaces H.

We can assume of the form $(\theta, b) : \theta^T x \geq b \quad \forall x \in \mathrm{Epi}(f)$.

(Why?: $\theta^T x + \alpha t \leq b$

but if $(x,t) \in \mathrm{Epi} f$, then $\forall t' \geq t$, $(x, t') \in \mathrm{Epi} f$.

So take $\bar{t} = \operatorname{argmax} \theta^T x + \alpha t \leq b$

So take $\bar{t} = \underset{\substack{t \\ (x,t)\in \mathrm{Epi}(f)}}{\arg\max}\ \theta^T x + \alpha t \le b$

$\Leftrightarrow \theta^T x \le b - \alpha \bar{t}$.)

$f(x) \geq \theta^T x - b \quad \forall (\theta, b) \in \mathcal{H}$

fix θ. $\quad b \geq \theta^T x - f(x) \quad \forall x$

$b \geq \sup_x \theta^T x - f(x) = f^*(\theta)$

$\Rightarrow f(x) = \sup_{(\theta,b)} \theta^T x - b = \sup_\theta \theta^T x - f^*(\theta) = f^{**}(x)$ $\boxtimes$

Examples.

$f(x) = \frac{1}{p} \sum x_i^p$

$f^*(x) = \frac{1}{q} \sum x_i^q$

$\frac{1}{p} + \frac{1}{q} = 1$

$f(x) = ax - b$

$f^*(\theta) = \begin{cases} 0 & \theta = a \\ \infty \end{cases}$

$\nabla f \longleftrightarrow \nabla f^*$ $\qquad h = h^{**}$

$\nabla f \longleftrightarrow \nabla f^*$ $\qquad\qquad h = h^{**}$

$$\min_x \; g(x) + h(Ax) = \min_x \max_\theta \; g(x) + \theta^T A x - h^*(\theta)$$

$$= \max_\theta \min_x \; g(x) + (A^T\theta)^T x - h^*(\theta)$$

$$= \max_\theta - \max_x \; \left((-A^T\theta)^T x - g(x)\right) - h^*(\theta)$$

$$= - \min_\theta \; g^*(-A^T\theta) + h^*(\theta) .$$

Sion's minimax Theorem. $X \subset \mathbb{R}^n$ compact, convex set.

$Y \subset \mathbb{R}^n$ convex. $f : X \times Y \longrightarrow \mathbb{R} \cup \{+\infty\}$ s.t.

$f(x, \cdot)$ is upper semi-continuous and quasi-concave on Y

$f(\cdot, y)$ is lower ——— quasi convex on X.

Then

$$\min_{x \in X} \sup_{y \in Y} f(x,y) = \sup_{y \in Y} \min_{x \in X} f(x, y)$$

<u>Example 1.</u> unit capacity flow problem $G = (V, E)$

$$\max \; c^T f$$
$$Af = d \quad -1 \leq f \leq 1$$

f: flow

$A : \mathbb{R}^{V \times E}$

vertex-edge adjacency.

d: demands on vertices.

if s-t maxflow then
$$d_s = d \quad d_t = -d \quad d_i = 0 \quad i \neq s, t.$$

$$\max_{Af = d,\, -1 \leq f \leq 1} c^T f = \max_{-1 \leq f \leq 1} \min_{\phi} c^T f - \phi^T (Af - d)$$

$$= \min_{\phi} \max_{-1 \leq f \leq 1} \phi^T d + (c - A^T \phi)^T f$$

$$= \min_{\phi} \phi^T d + \sum_{e \in E} |c - A^T \phi|$$

$$c = 0, \quad d = \{ \underset{\uparrow \atop s}{F}, 0 \ldots 0, \underset{\uparrow \atop t}{-F} \}$$

$\phi_j - \phi_i$

cut

$S \; v \in V : \phi(v) \geq t \}$

cut

$\{v \in V : \phi(v) \geq t\}$

$\phi_1 \leq \phi_i \quad \cdots \quad \phi_j \leq \phi_n$

<u>Exercise</u>: show this is min s-t cut

<u>Example 2.</u> SDP.

<u>Primal</u>: $\max\limits_{X \succeq 0} \quad C \bullet X \qquad A_i \bullet X = b_i \quad i = 1, 2, \dots m$

<u>Dual</u>: $\min\limits_{y} \quad b^T y \qquad \sum\limits_{i=1}^{m} y_i A_i \succeq C$

X is $n \times n$ symmetric.

Primal takes $O(n^2(Z + n^4))$ $\quad Z$: total # nnz in A_i.

Dual takes $O(m(Z + n^\omega + m^2))$

better when $m < n^2$, often the case.

But how to recover primal solution from dual?

But how to recover primal solution from ... .
(we only solve to some error ε).

$$\min_{\sum_i y_i A_i \succcurlyeq C} b^T y = \min_{v^T(\sum y_i A_i - C)v \geq 0 \ \forall v.} b^T y$$

When running the cutting plane method on DUAL, we get planes of the form $v^T(\sum_i y_i A_i - C)v \geq 0$

Let S be the set of all such v. At the end,

$$\min_{\sum y_i A_i \succcurlyeq C} b^T y \leq \min_{\substack{v^T(\sum_i y_i A_i - C)v \geq 0 \\ v \in S}} b^T y + \varepsilon.$$

Now consider RHS.

$$\min_{\substack{v^T(\sum_i y_i A_i - C)v \geq 0 \\ v \in S}} b^T y = \min_{y} \max_{\substack{\lambda_v \geq 0 \\ v \in S}} b^T y - \sum_{v \in S} \lambda_v v^T(\sum_i y_i A_i - C)v$$

$$= \max_{\lambda \geq 0} \min_{y} C \bullet \sum_{v \in S} \lambda_v v v^T + b^T y - \sum_i y_i \left(A_i \bullet \sum_{v \in S} \lambda_v v v^T\right)$$

$$= \max_{\lambda_v \geq 0} \min_{y} C \cdots$$

$$= \max_{X = \sum_{v \in S} \lambda_v v v^T,\ \lambda_v \geq 0} \ \min_{y} \ C \bullet X + \sum_i y_i (b_i - A_i \bullet X)$$

$$= \max_{\substack{X = \sum \lambda_v v v^T,\ \lambda_v \geq 0 \\ A_i \bullet X = b_i}} C \bullet X$$

(else the OPT is $-\infty$)

this is

$$\max \sum_{v \in S} \lambda_v (v^T C v)$$

$$\sum_v \lambda_v (v^T A_i v) = b_i, \quad \lambda_v \geq 0$$

an LP !