Duality & Reductions

Monday, January 27, 2020 10:35 PM

ylw

ORACLES 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) : $x \in K$ s.t. $c^T x \leq c^T y \quad \forall x \in K$

or "K is EMPTY".

Cutting Plane Method: $OPT_K \rightarrow SEP_K$.

SEP is stronger than MEM

SEP is stronger than MEM

OPT ———————— VAL.

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

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

$K =$ Conv Hull $\{a_1, \ldots, a_m\}$ $\quad$ 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 subgradient of f at x).

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

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

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

$$= \sup_{x \in K} c^T x \qquad \boxed{\mathrm{EVAL}_{\delta_K^*} \equiv \mathrm{VAL}_K.}$$

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

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

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

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

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

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

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

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

$$\begin{tikzcd}
\boxed{\mathrm{OPT}} \arrow[r, shift left] \arrow[d, shift right] & \boxed{\mathrm{SEP}} \arrow[l, shift left] \arrow[d, shift right] & \boxed{\mathrm{GRAD}_{f^*}} \arrow[r, shift left] \arrow[d, shift right] & \boxed{\mathrm{GRAD}_f} \arrow[l, shift left] \arrow[d, shift right] \\
\boxed{\mathrm{VAL}} \arrow[u, shift right] & \boxed{\mathrm{MEM}} \arrow[u, shift right] & \boxed{\mathrm{EVAL}_{f^*}} \arrow[u, shift right] & \boxed{\mathrm{EVAL}_f} \arrow[u, shift right]
\end{tikzcd}$$

Thm. Convex f, $\mathrm{Epi}(f)$ 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 $\bar{t} = \underset{\substack{t \\ (x,t) \in \mathrm{Epi}(f)}}{\operatorname{argmax}} \theta^T x + \alpha t \leq b$

So

$$t$$
$$(x,t) \in \mathrm{Epi}(f)$$

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

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

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.

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

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

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

$$f(x) = ax - b$$

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

$$\nabla f \longleftrightarrow \nabla f^*$$

$$h = h^{**}$$

$$\ldots \theta^T x - h^*(\theta)$$

v .

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

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

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

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

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

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

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

$f(\cdot, y)$ is lower semi-continuous and 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</u>. SDP.

$C \cdot X \quad A_i \bullet X = b_i \quad i = 1, 2, \ldots m$

Primal: $\max_{X \succeq 0} \quad C \bullet X \qquad A_i \bullet X = b_i \quad i=1,2,\dots m$

Dual: $\min_{y} \quad b^T y \qquad \sum_{i=1}^{m} y_i A_i \succeq C$

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?
(we only solve to some error ε).

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

When running the cutting plane method on DUAL,

we get planes of the form $v^T(\sum y_i A_i - C)v \geq 0$

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,

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

Now consider RHS.

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

$$= \max_{\lambda_v \geq 0} \min_y C \bullet \sum \lambda_v v v^T + b^T y - \sum_i y_i \Big(A_i \bullet \sum_{v \in S} \lambda_v v v^T\Big)$$

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

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

(else the OPT is $-\infty$)

$X = \sum \lambda_v v v^T$, $\lambda_v \geq 0$

$A_i \cdot X = b_i$

---

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

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

an LP !
