Lecture 17 (Fast) polytope intersection

Outline

(1) Problem definition

(2) why does it make sense?

(3) relaxed formulation and solving the dual

(4) error analysis

Problem formulation

Given two polytopes K_1, K_2 where we know how to solve fast

$$\min c^T x \quad \text{s.t } x \in K_1$$

and

$$\min c^T x \quad \text{s.t } x \in K_2 \qquad \forall c \in \mathbb{R}^n$$

we want to solve

$$\min c^T x \quad \text{s.t } x \in K_1 \cap K_2$$

(2) why does it make sense?

Problem

Input: a graph $G=(V,E)$, a weight function $w: E \to \mathbb{R}_+$, and a partition of the edges in different color sets:

$$E = C_1 \cup C_2 \cup \dots \cup C_k, \quad C_i \cap C_j = \phi \quad \forall i \neq j$$

Output: a minimum weight colorful spanning tree.

each edge of the tree has to be of a different color.

The problem can be described as the minimization of a linear function over the intersection of two polytopes

$\Longrightarrow$

spanning tree

$$\min \sum_{e\in E} w_e x_e$$

s.t. $\sum_{e\in E} x_e = n-1$

$\sum_{e\in E(S)} x_e \le |S|-1 \quad \forall S \subseteq V$

$0 \le x_e \le 1 \quad \forall e\in E$

$\Longrightarrow$ greedy works (no need to solve an LP)

($E(S)$ is the set of edges with both endpoints at S)

colorful graph

$$\min \sum_{e\in E} w_e x_e$$

st $\sum_{e\in C_i} x_e \le 1 \quad \forall i\in[k]$

\# of colors

$0 \le x_e \le 1 \quad \forall e\in E$

$\Longleftarrow$ greedy works

minimum weight colorful spanning tree

$\min \sum_{e\in E} w_e x_e$ s.t.

$$\left.\begin{aligned} &\sum_{e\in E} x_e = n-1 \\ &\sum_{e\in E(S)} x_e \le |S|-1 \quad \forall S\subseteq V \\ &0\le x_e\le 1 \quad \forall e\in E \end{aligned}\right\} K_1$$

linear function

$$\left.\begin{aligned} &\sum_{e\in C_i} x_e \le 1 \quad \forall i\in[k] \\ &0\le x_e\le 1 \quad \forall e\in E \end{aligned}\right\} K_2$$

since both polytopes can be described by a matroid $\Longrightarrow$ the intersection has integral solutions

The naive approach to solve the colorful spanning tree problem would be to use directly the faster cutting plane method. This will lead to a slow algorithm, as the separation oracle for the spanning tree problem alone, required solving a min-cut max-flow problem. Consequently each iteration is very costly. Moreover we do not use at all the fact that minimizing a linear function over each one of the polytopes separately is easy and fast.

(3) relaxed formulation and solving the dual.

Problem definition

$$\min \langle c, x\rangle \quad \text{s.t } x \in K_1 \cap K_2$$

if

$\min \langle c, x\rangle$ s.t $x \in K_1$ is easy and

$\min \langle c, x\rangle$ s.t $x \in K_2$ is easy.

we will use

$\min \langle c, x\rangle$ s.t $x \in K_i$

as an optimization subroutine → optimization oracle

Roadmap

We will ultimately use a cutting plane method to solve the problem. We will overcome the costly separation oracle computation by transforming the problem such that:

① separation oracle becomes trivial

② subgradient oracle is computed using the optimization oracle

② Claim

an optimization oracle for $\max c^T x$ s.t. $x \in K$

is a subgradient oracle for $f(c) = \max_{x \in K} c^T x$

Proof

let $x^* = \arg\max_{x \in K} c^T x$

$$f(d) = \max_{x \in K} d^T x \geq d^T x^* =$$

$$= c^T x^* + (x^*)^T (d-c) =$$

$$= f(c) + \underbrace{(x^*)^T}_{\text{subgradient definition.}} (d-c)$$

$\max c^T x$

$x \in K_1 \cap K_2$

assuming that

$\max_{x \in K_1} \|x\|_2 \leq M$

$\max_{x \in K_2} \|x\|_2 \leq M$

$$\max_{x \in K_1, y \in K_2} \frac{1}{2} c^T x + \frac{1}{2} c^T y - \frac{\lambda}{2} \|x-y\|_2^2$$

forcing $x \approx y$

If λ is big enough

we will use instead

$$\max_{x \in K_1, y \in K_2} f_\lambda(x,y) = \frac{1}{2}c^T x + \frac{1}{2}c^T y - \frac{\lambda}{2}\|x-y\|_2^2 \underbrace{- \frac{1}{2\lambda}\|x\|_2^2 - \frac{1}{2\lambda}\|y\|_2^2}$$

these terms will help recover an almost optimum solution.

good proxy lemma

① for $\lambda \geq 1$ there is a unique minimizer to the problem $\max_{x \in K_1, y \in K_2} f_\lambda(x,y)$, let that be (x_λ, y_λ)

and ② $\max_{x \in K_1 \cap K_2} c^T x \leq f(x_\lambda, y_\lambda) + \frac{\ell^2}{\lambda}$

③ $\|x_\lambda - y_\lambda\|_2^2 \leq \frac{6\ell^2}{\lambda}$

proof

① unique maximizer (x_λ, y_λ) because of $\frac{1}{\lambda}$-strong concavity.

② let $x^* \in \arg\max_{x \in K_1 \cap K_2} c^T x$ then

$$f_\lambda(x_\lambda, y_\lambda) \geq f_\lambda(x^*, x^*) = c^T x^* - \frac{\|x^*\|_2^2}{\lambda} \geq c^T x^* - \frac{\ell^2}{\lambda} \geq -2\ell^2$$

③ $f_\lambda(x_\lambda, y_\lambda) \leq \frac{1}{2}\|c\|_2 \cdot \|x_\lambda\|_2 + \frac{1}{2}\|c\|_2\|y_\lambda\|_2 - \frac{\lambda}{2}\|x_\lambda - y_\lambda\|_2^2 \leq$

$$\leq \frac{\mu^2}{2} + \frac{\mu^2}{2} - \frac{\lambda}{2}\|x_\lambda - y_\lambda\|_2^2 \Rightarrow$$

$$\Rightarrow \|x_\lambda - y_\lambda\|_2^2 \leq \frac{2}{\lambda}\left(\mu^2 - f_\lambda(x_\lambda, y_\lambda)\right) \overset{②}{\leq}$$

$$\leq \frac{2}{\lambda}(\mu^2 + 2\mu^2) = \frac{6\mu^2}{\lambda}$$

Dual transformation

$$\max_{x\in K_1,\, y\in K_2} f_\lambda(x,y) = \frac{1}{2}c^T x + \frac{1}{2}c^T y - \frac{\lambda}{2}\|x-y\|_2^2 - \frac{1}{2\lambda}\|x\|_2^2 - \frac{1}{2\lambda}\|y\|_2^2$$

$$\frac{1}{2}\|x\|^2 = \max_{\theta : \|\theta\| \leq \mu} \left\{ \theta^T x - \frac{1}{2}\|\theta\|_2^2 \right\} =$$

(since $\|x\| \leq \mu$)

$$= - \min_{\theta : \|\theta\| \mu} \left\{ \frac{1}{2}\|\theta\|_2^2 - \theta^T x \right\}$$

$$f_\lambda(x,y) = \frac{1}{2}c^T x + \frac{1}{2}c^T y$$
$$+ \lambda \min_{\theta_1 : \|\theta_1\|_2 \leq \mu} \left\{ \frac{1}{2}\|\theta_1\|^2 - \theta_1^T (x-y) \right\}$$
$$+ \frac{1}{\lambda} \min_{\theta_2 : \|\theta_2\|_2 \leq \mu} \left\{ \frac{1}{2}\|\theta_2\|^2 - \theta_2^T x \right\}$$
$$+ \frac{1}{\lambda} \min_{\theta_3 : \|\theta_3\|_2 \leq \mu} \left\{ \frac{1}{2}\|\theta_3\|^2 - \theta_3^T x \right\}$$

$$\Rightarrow \max_{x\in K_1, y\in K_2} f_\lambda(x,y) =$$

$$= \max_{\substack{x\in K_1\\ y\in K_2}} \min_{\substack{\theta_1,\theta_2,\theta_3\\ \|\theta_i\|\le M\\ \forall i=1,2,3}} \left(\frac{c}{2} - \lambda\theta_1 - \frac{1}{\lambda}\theta_2\right)^T x$$

$$\left(\frac{c}{2} - \lambda\theta_1 - \frac{1}{\lambda}\theta_3\right)^T y$$

$$+ \frac{\lambda}{2}\|\theta_1\|_2^2$$

$$+ \frac{1}{2\lambda}\|\theta_2\|_2^2$$

$$+ \frac{1}{2\lambda}\|\theta_3\|_2^2$$

Observation

For x, y fixed
$$\left|\begin{array}{l}\theta_1 = x - y\\ \theta_2 = x\\ \theta_3 = y\end{array}\right.$$

θ_2, θ_3 will help us recover x_λ and y_λ that's why we needed the terms $-\frac{1}{2\lambda}\|x\|_2^2$ and $-\frac{1}{2\lambda}\|y\|_2^2$

$$\max_{\substack{x\in K_1\\ y\in K_2}} \ \min_{\substack{\theta_1,\theta_2,\theta_3\\ \|\theta_i\|\le M\\ \forall i=1,2,3}} \left(\frac{c}{2}-\lambda\theta_1-\frac{1}{\lambda}\theta_2\right)^T x + \left(\frac{c}{2}-\lambda\theta_1-\frac{1}{\lambda}\theta_3\right)^T y + \frac{\lambda}{2}\|\theta_1\|_2^2 + \frac{1}{2\lambda}\|\theta_2\|_2^2 + \frac{1}{2\lambda}\|\theta_3\|_2^2$$

Sion's theorem

$\Longrightarrow$

K_1, K_2 closed

$\theta_1,\theta_2,\theta_3$ s.t. $\|\theta_i\|_2\le M$ and convex

convex in $\theta_1,\theta_2,\theta_3$

concave in x,y

$$\Rightarrow \min_{\substack{\theta_1,\theta_2,\theta_3\\ \|\theta_i\|\le M\\ \forall i=1,2,3}} \ \max_{\substack{x\in K_1\\ y\in K_2}} \left(\frac{c}{2}-\lambda\theta_1-\frac{1}{\lambda}\theta_2\right)^T x + \left(\frac{c}{2}-\lambda\theta_1-\frac{1}{\lambda}\theta_3\right)^T y + \frac{\lambda}{2}\|\theta_1\|_2^2 + \frac{1}{2\lambda}\|\theta_2\|_2^2 + \frac{1}{2\lambda}\|\theta_3\|_2^2 =$$

$$= \min_{\substack{\theta_1,\theta_2,\theta_3\\ \|\theta_i\|\le M\\ \forall i=1,2,3}} \left\{ \max_{x\in K_1}\left(\frac{c}{2}-\lambda\theta_1-\frac{1}{\lambda}\theta_2\right)^T x + \max_{y\in K_2}\left(\frac{c}{2}-\lambda\theta_1-\frac{1}{\lambda}\theta_3\right)^T y + \lambda/2\|\theta_1\|_2^2 + \frac{1}{2\lambda}\|\theta_2\|_2^2 + \frac{1}{2\lambda}\|\theta_3\|_2^2 \right\}$$

$$= \min_{\substack{\theta_1,\theta_2,\theta_3 \\ \|\theta_i\| \le M \\ \forall i=1,2,3}} h_\lambda(\theta_1,\theta_2,\theta_3)$$

separation oracle of h_λ

just check if $\|\theta_1\| \le M$, $\|\theta_2\| \le M$
$\|\theta_3\| \le M$

optimization oracle of h_λ

$$x^* + y^* + \lambda\theta_1 + \frac{1}{\lambda}\theta_2 + \frac{1}{\lambda}\theta_3$$

$$x^* = \arg\max_{x \in K_1} \left(\frac{c}{2} - \lambda\theta_1 - \frac{1}{\lambda}\theta_2\right)^T x \Rightarrow$$ from the optimization oracle

$$y^* = \arg\max_{y \in K_2} \left(\frac{c}{2} - \lambda\theta_1 - \frac{1}{\lambda}\theta_3\right)^T y \Rightarrow$$ from the optimization oracle

$\Rightarrow$ we can use a cutting plane method to find $\theta_1,\theta_2,\theta_2$ s.t

$$h_\lambda(\theta_1,\theta_2,\theta_3) - \underbrace{\min_{\substack{\theta_1,\theta_2,\theta_3 \\ \|\theta_i\| \le M \\ \forall i=1,2,3}} h_\lambda(\theta_1,\theta_2,\theta_3)}_{\theta_1^*,\theta_2^*,\theta_3^* \leadsto \arg\min \ldots} \le \varepsilon$$

Error analysis

we know that $x_\lambda = \theta_2^*$ and we

$y_\lambda = \theta_3^*$

want to prove that $\theta_2 \simeq \theta_2^*$, $\theta_3 \simeq \theta_3^*$

$$h_\lambda(\theta_1,\theta_2,\theta_3) \leq h_\lambda(\theta_1^*,\theta_2^*,\theta_3^*) + \varepsilon$$

$$\xRightarrow[h_\lambda \ \frac{1}{\lambda}\text{-strongly convex}]{} \|\theta_1-\theta_1^*\|_2^2 + \|\theta_2-\theta_2^*\|_2^2 + \|\theta_3-\theta_3^*\|_2^2 \leq \frac{2\varepsilon}{1/\lambda} = 2\lambda\cdot\varepsilon$$

$$\xRightarrow[\theta_3^* = y_\lambda]{\theta_2^* = x_\lambda} \|\theta_2 - x_\lambda\|_2^2 + \|\theta_3 - y_\lambda\|_2^2 \leq 2\lambda\cdot\varepsilon$$

$a^2+b^2 \geq \frac{1}{2}(a+b)^2$

$$\Rightarrow \begin{cases} \|x_\lambda - y_\lambda\|_2 \geq \|\theta_2-\theta_3\| - 2\sqrt{2\lambda\varepsilon} \\ \|x_\lambda\|_2 \geq \|\theta_2\|_2 - \sqrt{2\lambda\varepsilon} \\ \|y_\lambda\|_2 \geq \|\theta_3\|_2 - \sqrt{2\lambda\varepsilon} \end{cases}$$

$$\Longrightarrow$$

Now we are ready to bound

$$f_\lambda(\theta_2,\theta_3) - f_\lambda(x_\lambda, y_\lambda) =$$

$$= \frac{1}{2}\langle c, \theta_2 - x_\lambda\rangle + \frac{1}{2}\langle c, \theta_3 - y_\lambda\rangle$$

$$-\frac{\lambda}{2}\left(\|\theta_2-\theta_3\|_2^2 - \|x_\lambda - y_\lambda\|_2^2\right)$$

$$-\frac{1}{2\lambda}\left(\|\theta_2\|_2^2 - \|x_\lambda\|_2^2\right)$$

$$-\frac{1}{2\lambda}\left(\|\theta_3\|_2^2 - \|y_\lambda\|_2^2\right) \geq$$

$$\geq -\frac{1}{2}u\sqrt{2\lambda\varepsilon}$$

$$-\frac{1}{2}u\sqrt{2\lambda\varepsilon}$$

$$-2\lambda\sqrt{2\lambda\varepsilon}\,\|x_\lambda - y_\lambda\|_2^2 - 4\lambda^2\varepsilon$$

$$-\frac{1}{\lambda}\|x_\lambda\|_2\sqrt{2\lambda\varepsilon} - \varepsilon$$

$$-\frac{1}{\lambda}\|y_\lambda\|_2\sqrt{2\lambda\varepsilon} - \varepsilon \geq$$

$\|x_\lambda\|_2 \leq u$

$\|y_\lambda\|_2 \leq u$

$\|x_\lambda - y_\lambda\|_2 \leq \sqrt{\frac{6u^2}{\lambda}}$

$$\underset{\lambda\geq 2}{\geq} -20u\lambda\sqrt{\varepsilon} - 10\lambda^2\varepsilon \implies$$

$$f_\lambda(\theta_2,\theta_3) - f_\lambda(x_\lambda,y_\lambda) \geq -2\,0\,\mu\lambda\sqrt{\varepsilon} - 10\lambda^2\varepsilon$$

$$\max_{x\in K_1\cap K_2} c^T x \leq f(x_\lambda,y_\lambda) + \frac{\mu^2}{\lambda}$$

using as solution $\xi = \frac{\theta_2+\theta_3}{2}$ we get

$$\max_{x\in K_1\cap K_2} c^T x \leq f_\lambda(x_\lambda,y_\lambda) + \frac{\mu^2}{\lambda} \leq$$

$$\leq f_\lambda(\theta_2,\theta_3) + \frac{\mu^2}{\lambda} + 20\mu\lambda\sqrt{\varepsilon} + 10\lambda^2\varepsilon$$

$$\leq \langle c,\xi\rangle + \frac{20\,\mu^2}{\lambda} + 20\lambda^3\varepsilon$$

and $\|\xi - x_\lambda\|_2 + \|\xi - y_\lambda\|_2 \leq 4\sqrt{2\lambda\varepsilon} + \sqrt{\frac{6\mu^2}{\lambda}}$

(using $\|\theta_2 - x_\lambda\|_2^2 + \|\theta_3 - y_\lambda\|_2^2 \leq 2\cdot\lambda\cdot\varepsilon$)

All together

$\lambda = \frac{410\mu^2}{\delta^2} \quad \varepsilon = \frac{\delta^3}{10^3\mu^6}$ then

$$\max_{x \in K_1 \cap K_2} c^T x \le c^T \xi + \delta$$

and $\|\xi - x_\lambda\|_2 + \|\xi - y_\lambda\|_2 \le \delta$

optimization oracles complexity

$\Rightarrow$ complexity =

$$O\left(n\left(OO(K_1) + OO(K_2)\right)\log n\mu/\delta + n^3 \operatorname{polylog}\left(\frac{n\mu}{\delta}\right)\right)$$

Lee, Sidford, Wong