Lecture 7

Tuesday, March 23, 2021 10:24 AM

→ We assume that the polytope P is bounded, full-dimensional & non-empty.

Initialization: Goal: Find $\eta_0 > 0$ & $x_0 \in \mathrm{int}(P)$ s.t.

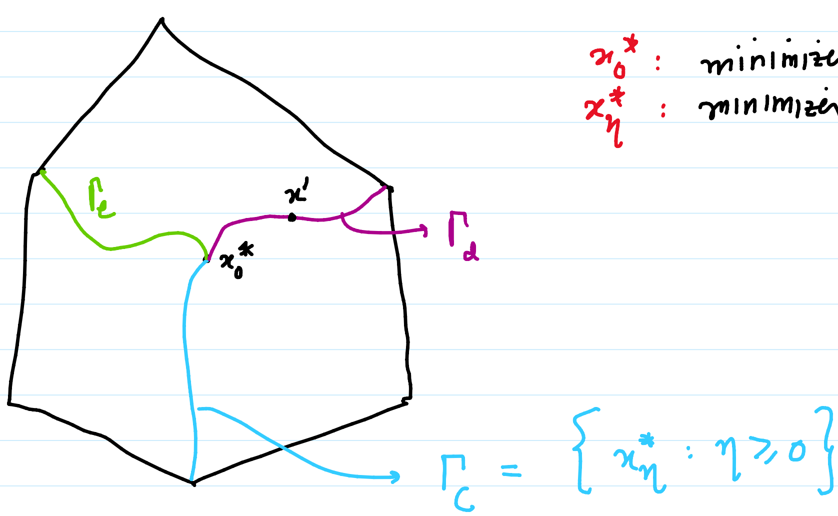
$$\| N_{\eta_0}(x_0) \|_{x_0} \le 1/6$$

Recall: $N_{\eta_0}(x_0) := -H_F(x_0)^{-1} \nabla f_{\eta_0}(x_0) = -H_F(x_0)^{-1}(\eta_0 c + \nabla F(x_0))$

η_0 will be $\geq 2^{-\tilde{O}(nL)}$ ($\tilde{O}(\cdot)$ hides logarithmic factors in various parameters).

First let us assume that we have access to an interior point $x' \in P$. Infact, we will assume something stronger. We will assume that we have an $x' \in P$ s.t.

$$\langle a_i, x \rangle \le b_i - 2^{-\tilde{O}(nL)} \quad \forall i \in [m]$$



x_0^* : minimizer of F

x_η^* : minimizer of $f_\eta = \eta \langle c, x \rangle + F(x)$

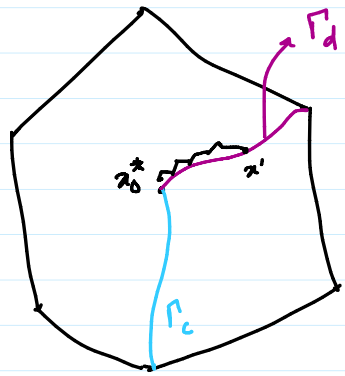
Observation: Each $x' \in \mathrm{int}(P)$ lies on some central path!

Pf: $d := -\nabla F(x')$. Then x' lies on Γ_d.

$$f(x) = \langle d, x \rangle + F(x)$$

Then $\nabla f(x') = d + \nabla F(x') = 0$. Hence x' is the minimizer of f.

Then $\nabla f(x') = d + \nabla F(x') = 0$. Hence x' is the minimizer of f.



Strategy: Start from x' & traverse the central path Γ_d in reverse. Once we get close to x_0^*, that will close to the central path Γ_c.

$$\left(s_i(x) = b_i - \langle a_i, x\rangle \right)$$

Lemma (Following the central path Γ_d in reverse): Suppose x' is a point s.t. $s_i(x') \geq \beta \max_{x \in P} s_i(x) \quad \forall\, i \in [m]$.

$$N'_\eta(x) := -H_F(x)^{-1}\left(\eta \underbrace{d}_{-\nabla F(x')} + \nabla F(x)\right)$$

Then whenever $\eta \leq \frac{\beta}{24\sqrt{m}}$ & $\|N'_\eta(x)\|_x \leq 1/24$, it holds that $\|H_F(x)^{-1}\nabla F(x)\|_x \leq 1/12$.

Proof: Use triangle inequality for the local norm

$$\|H_F(x)^{-1}\nabla F(x)\|_x \leq \|N'_\eta(x)\|_x + \eta\|H_F(x)^{-1} d\|_x$$

$$\leq \frac{1}{24} + \eta\|H_F(x)^{-1}\nabla F(x')\|_x$$

$$\|H_F(x)^{-1}\nabla F(x')\|_x^2 = \left\langle \nabla F(x'), H_F(x)^{-1}\nabla F(x')\right\rangle$$

Let A be an $m \times n$ matrix whose rows are a_i's.

$$A = \begin{pmatrix} - a_1 - \\ - a_2 - \\ \vdots \\ - a_m - \end{pmatrix}$$

Let S_x denote the $m \times m$ diagonal matrix that contains the slack values for x.

$$S_x = \begin{pmatrix} s_1(x) & & 0 \\ & \ddots & \\ 0 & & s_m(x) \end{pmatrix}$$

Then $H_F(x) = A^T S_x^{-2} A$, $\quad \nabla F(x) = A^T S_x^{-1} \mathbb{1}$

$\nabla F(x') = A^T S_{x'}^{-1} \mathbb{1}$

$\mathbb{1}$: all 1's vector $(m \times 1)$.

Then $\langle \nabla F(x'), H_F(x)^{-1} \nabla F(x') \rangle$

$$= \mathbb{1}^T S_{x'}^{-1} A \left(A^T S_x^{-2} A\right)^{-1} A^T S_{x'} \mathbb{1}$$

$$= \mathbb{1}^T S_{x'}^{-1} S_x \, S_x^{-1} A \left(A^T S_x^{-2} A\right)^{-1} A^T S_x^{-1} \, S_x S_{x'}^{-1} \mathbb{1}$$

$$= v^T \Pi v$$

where $v = S_x S_{x'}^{-1} \mathbb{1}$, $\quad \Pi = S_x^{-1} A \left(A^T S_x^{-2} A\right)^{-1} A^T S_x^{-1}$

Let $B = S_x^{-1} A$, Then $\Pi = \underbrace{B \left(B^T B\right)^{-1} B^T}_{\text{projection matrices}}$

$$\Pi^2 = B \left(B^T B\right)^{-1} \cancel{B^T B} \cancel{\left(B^T B\right)^{-1}} B^T$$

$$= B \left(B^T B\right)^{-1} B^T = \Pi$$

since Π is a projection matrix

$$\Rightarrow \quad v^T \Pi v = \|\Pi v\|_2^2 \leq \|v\|_2^2$$

$$\|v\|^2 = \sum^m s_i(x)^2 \leq \underline{m}$$

$$\|v\|_2^2 = \sum_{i=1}^{m} \frac{s_i(x)^2}{s_i(x')^2} \leq \frac{m}{\beta^2}$$

$$\Rightarrow \quad \| H_f(x)^{-1} \nabla F(x') \|_x \leq \frac{\sqrt{m}}{\beta}$$

$$\Rightarrow \quad \| H_f(x)^{-1} \nabla f(x) \|_x \leq \frac{1}{24} + \eta \frac{\sqrt{m}}{\beta}$$

$$\leq \frac{1}{24} + \frac{1}{24} = \frac{1}{12}$$

Lemma (Switching the central path): Suppose $x \in \operatorname{int}(P)$ s.t. $\| H_f(x)^{-1} \nabla F(x) \|_x \leq 1/12$. & η_0 is s.t.

$$\eta_0 \leq \frac{1}{12} \cdot \frac{1}{\langle c, x - x^* \rangle}. \Rightarrow \| N_{\eta_0}(x) \|_x \leq 1/6$$

Proof:

$$\| N_{\eta_0}(x) \|_x = \| H_f(x)^{-1} (\eta_0 c + \nabla f(x)) \|_x$$

$$\leq \eta_0 \| H_f(x)^{-1} c \|_x + \| H_f(x)^{-1} \nabla f(x) \|_x$$

$$\leq \eta_0 \| H_f(x)^{-1} c \|_x + \frac{1}{12}$$

Let $c_x := H_f(x)^{-1} c$. Then by the Dikin's Ellipsoid property, $x - \frac{c_x}{\|c_x\|_x} \in P$

$$\Rightarrow \quad \left\langle c, x - \frac{c_x}{\|c_x\|_x} \right\rangle \geq \langle c, x^* \rangle$$

$\|c_x\|_x$

$$\Rightarrow \langle c, x - x^* \rangle \geq \left\langle c, \frac{c_x}{\|c_x\|_x} \right\rangle = \frac{\|c_x\|_x^2}{\|c_x\|_x} = \|c_x\|_x$$

$$\leq \eta_0 \langle c, x - x^* \rangle + \frac{1}{12}$$

$$\leq \frac{1}{12} + \frac{1}{12} = 1/6 .$$

Lemma (Finding a starting point given an interior point):

There is an algorithm that given x' s.t. $s_i(x') \geq \beta \max_{x \in P} s_i(x)$ $\forall i$, outputs a point $x_0 \in int(P)$ & $\eta_0 > 0$ s.t.

$$\|N_{\eta_0}(x_0)\|_{x_0} \leq 1/6 .$$

And also $\eta_0 \geq \dfrac{1}{\|c\|_2 \, diam(P)} \geq 2^{-\tilde{O}(nL)}$.

Running time: $poly(n, m, L, \log(\frac{1}{\beta}))$

Proof: 1. Start by running the IPM in reverse for the central path Γ_d, $d := -\nabla F(x')$.

Initially $\eta = 1$ & $\|N_1(x')\|_{x'} = 0 \leq 1/6$.

2. Decrease η till it becomes less than

$$\eta_0 := \min \left\{ \frac{\beta}{24\sqrt{m}}, \frac{1}{\|c\|_2 \, diam(P)} \right\}$$

multiplicative decrease by a factor $(1 - \frac{1}{\quad})$

$\Downarrow$ multiplicative decrease by a factor $\left(1 - \frac{1}{20\sqrt{m}}\right)$ at each step.

At this point, we get that $\| N'_{\eta_0}(x) \|_x \le 1/6$.

3. Do a couple of Newton iterations to get y s.t.
$$\| N'_{\eta_0}(y) \|_y \le 1/24 .$$

4. From the first lemma, we get that
$$\| H_F(y)^{-1} \nabla F(y) \|_y \le 1/12$$

5. Since $\langle c, y - x^* \rangle \le \|c\|_2 \, \|y - x^*\|_2$
$$\le \|c\|_2 \, \mathrm{diam}(P)$$

$$\eta_0 \le \frac{1}{\|c\|_2 \, \mathrm{diam}(P)} \le \frac{1}{\langle c, y - x^* \rangle} .$$

Hence we can apply the second lemma to get
$$\| N_{\eta_0}(y) \|_y \le 1/6 .$$

Exercise: $\|c\|_2, \mathrm{diam}(P) \le 2^{\tilde{O}(nL)}$.

Now how to find an interior point $x^0 \in P$?

Solution: Look at a different polytope $P' \subseteq \mathbb{R}^{n+1}$

$$P' = \left\{ x \in \mathbb{R}^n, t \in \mathbb{R} \ \text{ s.t. } \ \langle a_i, x \rangle \le b_i + t \ \ \forall i \in [m] \atop -C \le t \le C \right\}$$

Take $C = 2 + \sum_i |b_i|$. Then the point

$x' = 0$, $t' = 1 + \sum_i |b_i|$ is a point in $\mathrm{int}(P')$

& slack of each constraint ≥ 1.

<u>use this to solve the following linear program</u>

$$\min_{(x,t) \in P'} t$$

suppose (x^*, t^*) is the optimal sol^n.

Then $\langle a_i, x^* \rangle \leq b_i + t^* \quad \forall i$

<u>Exercise:</u> $t^* \leq -2^{-\tilde{O}(nL)}$.

$$\boxed{\text{Solving the above program to precision/error } 2^{-\tilde{O}(nL)} \text{ will give us a point } x' \in P \text{ s.t. } \langle a_i, x' \rangle \leq b - 2^{-\tilde{O}(nL)} \ \forall i}$$

<u>Exercise:</u> If P is bounded, full-dimensional & non-empty, then so is P'.

<u>Alternate:</u>

$$\boxed{\min_{x \in P} \langle c, x \rangle}$$

$\downarrow$

reduce to the following feasibility problem

$$\boxed{\text{find a point in } P \cap \{x : \langle c, x \rangle \leq \gamma\}}$$

& do a binary search on γ.

Self-concordant barrier functions

Definition (Self-concordant barrier function):

Let $K \in \mathbb{R}^n$ be a convex set & let $F : \text{int}(K) \to \mathbb{R}$ be a thrice differentiable function. We say that F is a self-concordant barrier function with parameter ν if it satisfies the following properties:

1. (Barrier): $F(x) \to \infty$ as x approaches the boundary of K.
2. (Convexity): F is strictly convex.
3. (Complexity parameter): $\| N_F(x) \|_x^2 = \langle \nabla F(x), H_F(x)^{-1} \nabla F(x) \rangle \le \nu \quad \forall x \in \text{int}(K)$.
4. (Self-concordance): $\forall x \in \text{int}(K)$ & $\forall h$

$$\left| \frac{d^3}{dt^3} f(x+th) \Big|_{t=0} \right| \le 2 \left(\frac{d^2}{dt^2} f(x+th) \Big|_{t=0} \right)^{3/2}$$

The factor of $3/2$ in the exponent ensures scale-invariance

i.e. if $F(x)$ is self-concordant, then so is $F(\lambda x)$ for any $\lambda \in \mathbb{R}$.

The logarithmic barrier function satisfies the above properties with $\nu = m$.

Exercise: A self-concordant function F satisfies the NL condition

Exercise: A self-concordant function F satisfies the NL condition with $\delta_0 = 1/8$.

Theorem: Let F be a self-concordant barrier function on a convex set $K \in \mathbb{R}^n$ with complexity parameter ν. Then there is a IPM that after $O\left(\sqrt{\nu}\,\log(\nu/\varepsilon\eta_0)\right)$ iterations outputs $\hat{x} \in K$ s.t. $\langle c, \hat{x}\rangle - \langle c, x^*\rangle \leq \varepsilon$. Each iteration requires solving a linear system of the kind

$$H_F(x)\, y = z \quad (y \text{ are the unknowns}).$$

Search for efficient barrier functions:

$P = \{x \in \mathbb{R}^n : \langle a_i, x\rangle \leq b_i \ \forall\ i \in [m]\}$.

The logarithmic barrier has complexity parameter m.

Can we make it $O(1)$?

Lower bound: Any self-concordant barrier function for $K = [0,1]^n$ has complexity parameter at least n.

Every convex set $K \in \mathbb{R}^n$ has a self-concordant barrier function with complexity parameter $O(n)$.

$\downarrow$

Constructions non-explicit

$\downarrow$

[illegible] function [illegible]

Lee-Sidford constructed an explicit barrier function for polytopes which has complexity parameter $O(n)$.

[reduce linear programs to solving $\tilde{O}(\sqrt{n})$ linear systems]