Lecture 13

Last time

- Affine invariance
- A new guarantee for Newton's.

Today

- Back to IPM
- A "complete" IPM
- Answering our questions

Back to IPM

Recall the η problem

$$\min_{x \in \mathbb{R}^d} f_\eta(x) \quad \text{with } f_\eta(x) = \eta c^T x + B(x)$$

and our informal template:

IPM (Informal)

- Pick $x_0 \in \operatorname{int} P$ sufficiently close to $x^*(0)$ and pick $\eta_0 > 0$ small.
- Loop $k = 0, 1, \ldots, T$:
  - Find an approximate minimizer x_{k+1} of f_{η_k} using x_k for initialization
  - Increase $\eta_{k+1} = q\,\eta_k$ with $q > 1$.

▷ Return x_T

and we had several questions

Q_0: what B function to use?

Q_1: How to find x_0? What is sufficiently close?

Q_2: what method to use to find x_{k+1}?

Q_3: How to pick q?

Q_4: How to show that the method finds an approximate solution to the original LP in poly time?

← Intuition
Description
New questions

Answering questions

Today we answer all these questions except Q_1 (Exercise).

Our ultimate goal is to find

$$c^T x - c^T x^* \le \varepsilon \text{ with } x \in P.$$

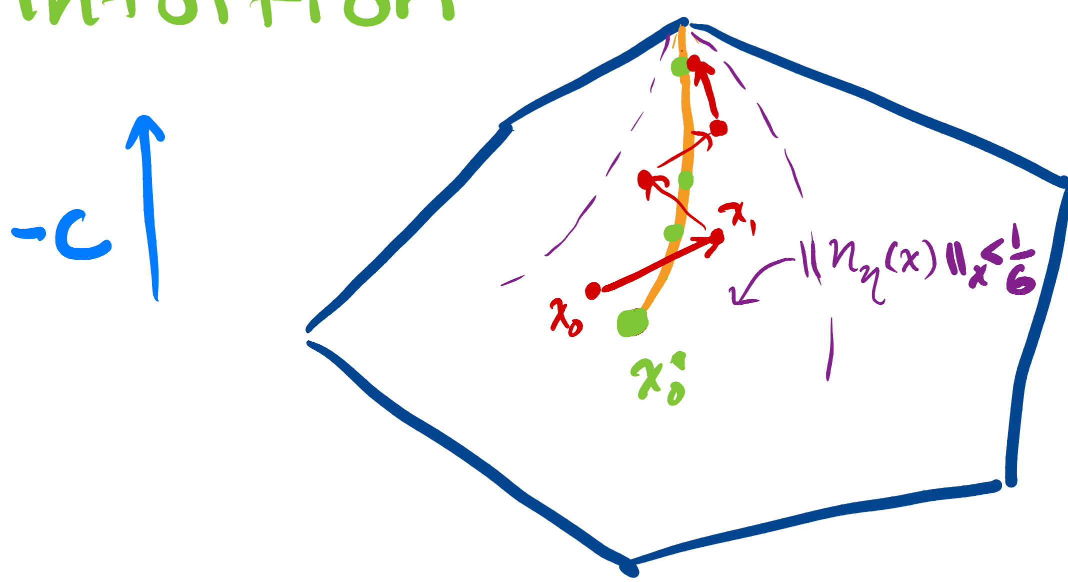
We will like to follow the

central path because we know $x^*_\eta \to x^*$. However, we don't need to solve for $x^*_{\eta_k}$ exactly, until the very end once we know

$$c^T x^*_{\eta_T} - c^T x^* \lessapprox \varepsilon.$$

So we could loosely track $x^*_{\eta_k}$ and then at the very end solve more accurately for $x^*_{\eta_T}$

Intuition



Given what we proved last class it is natural to expect that

$\|n_\eta(x)\|_x \leq \frac{1}{6}$ with

$$n_\eta(x) = -\left[\nabla^2 f_\eta(x)\right]^{-1} \nabla f(x).$$

is a natural condition to expect from all our iterates. This motivates

IPM (Full)

- Pick $x_0 \in \operatorname{int} P$ s.t. $\|n_{\eta_0}(x_0)\| \leq \frac{1}{6}$
- Loop $k = 0, 1, \dots, T-1$: so that $\eta_T > \frac{m}{2\varepsilon}$
  - $x_{k+1} \leftarrow x_k + n_{\eta_k}(x_k)$
  - $\eta_{k+1} \leftarrow \eta_k\left(1 + \frac{1}{20\sqrt{m}}\right)$
  - Increase $\eta_{k+1} = q\,\eta_k$ with $q > 1$.
- Run 2 steps of Newton for f_{η_T} starting from x_T. $\Big\}\ \hat{x}$

A_0: $B(x) = -\log \sum_{i=1}^{m} \log(b_i - a_i^T x)$.

The reason for this is that B is SC.

Lemma: The log barrier function

$$B(x) = -\sum_{i=1}^{m} \log(\underbrace{b_i - a_i^T x}_{r_i(x)})$$

is self concordant.

Proof: Recall that

$$H(x) = \sum_{i=1}^{m} \frac{a_i a_i^T}{s_i(x)^2}$$

Let $\delta = \|y - x\|_x < 1$, then

$$\delta^2 = (y-x)^T H(x)(y-x) = \sum_{i=1}^{m} \left(\frac{a_i^T(y-x)}{s_i(x)}\right)^2.$$

Hence, each individual term

$$\left(\frac{s_i(y) - s_i(x)}{s_i(x)}\right)^2 = \left(\frac{a_i^T(y-x)}{s_i(x)}\right)^2 \leq \delta^2.$$

Therefore,

$$(1-\delta)|s_i(x)| \leq |s_i(y)| \leq (1+\delta)|s_i(x)|$$

which implies

$$\frac{(1+\delta)^{-2}}{s_i(x)^2} \le \frac{1}{s_i(y)^2} \le \frac{(1-\delta)^{-2}}{s_i(x)^2}.$$

Thus

$$\frac{(1+\delta)^{-2}}{s_i(x)^2} a_i a_i^T \preceq \frac{a_i a_i^T}{s_i(y)^2} \preceq \frac{(1-\delta)^{-2}}{s_i(x)^2} a_i a_i^T.$$

Summing over all i and using the fact that

$$(1-3\delta) \le (1+\delta)^{-2} \le (1-\delta)^{-2} \le 1+3\delta$$

for all $\delta \in [0,1]$, yields the result. $\square$

This means that the theory we developed in Lecture 12 applies. Let

$$n_\eta(x) = -[\nabla^2 f_\eta(x)]^{-1} \nabla f(x).$$

Lemma +: For a given x, let $x_+ = x + n_\eta(x)$. Then if $\|n_\eta(x)\|_x \le \frac{1}{6}$ we have

$$\|n_\eta(x_+)\|_{x_+} \le 3\|n_\eta(x)\|_x^2 \le \frac{1}{12}$$

+

A_1: We do not answer how to get x_0. But we notice it will suffice for it to satisfy

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

A_2: We will run Newton's method for how many iterations are necessary to ensure $\|n_{\eta_{k+1}}(x_{k+1})\|_{x_{k+1}} \le 1/6$.

At the last iteration we might run it for longer to ensure

$$|c^T\hat{x} - c^T x^*| < \varepsilon.$$

Lemma η: For every $x \in \operatorname{int} P$ and $\eta', \eta > 0$, we have

$$\|n_{\eta'}(x)\|_x \le \frac{\eta'}{\eta}\|n_\eta(x)\|_x + \sqrt{m}\left|\frac{\eta'}{\eta} - 1\right|. \dashv$$

Before proving this result,

notice that, it tells us now to increment η, since

$$\| n_{\eta_{k+1}}(x_{k+1}) \|_{x_{k+1}} \leq q \, \| n_{\eta_k}(x_{k+1}) \|_{x_{k+1}}$$

$$\sqrt{m}\, |q - 1|$$

$$\leq \frac{q}{12} + \sqrt{m}\, |q + 1| .$$

A_3: We set $q = \left(1 + \frac{1}{20\sqrt{m}}\right)$

$$\| n_{\eta_{k+1}}(x_{k+1}) \|_{x_{k+1}} \leq \frac{1}{12} + \frac{1}{240\sqrt{m}} + \frac{1}{20}$$

One step suffices

$$\leq \frac{1}{6} .$$

Proof of Lemma η: We use $H(x) = \nabla^2 B(x)$ and $g(x) = \nabla B(x)$. Then

$$n_{\eta'}(x) = -H(x)^{-1} (\eta' c + g(x))$$

$$= -\frac{\eta'}{\eta} \underbrace{H(x)^{-1} (\eta c + g(x))}_{n_\eta(x)}$$

$$+ \left(1 - \frac{\eta'}{\eta}\right) H(x)^{-1} g(x)$$

Taking the $\|\cdot\|_x$ and applying triangle inequality yields

$$\|n_{\eta'}(x)\|_x \leq \frac{\eta'}{\eta'} \|n_\eta(x)\|_x + \left|\frac{\eta'}{\eta} - 1\right| \|H(x)^{-1} g(x)\|_x .$$

It suffices to bound $\|\overbrace{H(x)^{-1}g(x)}^{z}\|_x \leq \sqrt{m}$ for any $x \in \operatorname{int} P$.
Applying Cauchy-Schwarz:

$$\begin{aligned}
\|z\|_x^2 &= g(x)^T H(x)^{-1} g(x) \\
&= z^T g(x) \\
&= \sum^{m} 1 \cdot \frac{(z^T a_i)^2}{(b_i - a_i^T x)^2} \\
&\leq \sqrt{m} \sqrt{\sum_{i=1}^{m} \frac{(z^T a_i)^2}{(b_i - a_i^T x)^2}}
\end{aligned}$$

$$= \sqrt{m} \sqrt{z^T \underbrace{\left(\sum \frac{a_i a_i^T}{(b_i - a_i^T x)^2} \right)}_{H(x)} z}$$

$$= \sqrt{m} \, \| z_x \| .$$

We conclude $\|z\|_x \le \sqrt{m}$. $\square$

Finally we need to decide how many steps to run for the last execution of Newton so that

$$c^T x_{T+1} - c^T x^* \le \epsilon .$$

Proposition (♡) Suppose $x \in \operatorname{int} P$ and $\eta > 0$ s.t. $\|n_\eta(x)\|_x < 1/24$, then

$$c^T x - c^T x^* \le \frac{m}{\eta} \left(1 - 4 \|n_\eta(x)\|_x \right)^{-1} .$$

$\dashv$

Notice that this proposition implies that if

$$\| n_{\eta}(x) \|_{x} \leq \frac{1}{24}$$

$$\frac{m}{2\varepsilon} \leq \eta$$

Then, we obtain

$$c^{T}x - c^{T}x^{*} \leq \varepsilon.$$

Recall that $\| n_{\eta_T}(x_T) \|_{x_T} < \frac{1}{6}$.

Therefore after 2 iterations of Newton

$$\begin{aligned} \| n_{\eta_T}(\hat{x}) \|_{\hat{x}} &\leq 3 \| n_{\eta_T}(x_T^{+}) \|_{x_T^{+}}^{2} \\ &\leq 27 \| n_{\eta_T}(x_T) \|_{x_T}^{4} \\ &\leq \frac{1}{48} \end{aligned}$$

Thus,

$$c^T \tilde{x} - c^T x^* \le \frac{12}{11} \frac{m}{\eta_T} < \varepsilon.$$

In turn, if we establish Proposition (♡), we will have an answer for Q4.

Theorem A4: IPM (Full) using $T = O\left(\sqrt{m} \log \frac{m}{\varepsilon \eta_0}\right)$ iterations, outputs a point $\hat{x} \in \operatorname{int} P$ satisfying

$$c^T \hat{x} - c^T x^* \le \varepsilon.$$

Further, each step involves solving a linear system, which can be executed in polynomial time. $\dashv$