Ellipsoid Algorithm & Related Topics

Theorem 1: Given an LP $\max\{c^Tx \mid Ax \le b, x \ge 0\}$. there is a polynomial (in $\langle A\rangle, \langle b\rangle, \langle c\rangle$) time algorithm that produces an optimal bfs for the LP.

Note: $\langle A\rangle$ is the length of the bit representation of the matrix A, etc. We do not know a strongly polynomial time algorithm for linear programming, that is an important open question.

Theorem 2: Given ~~an~~ (finite) LP $\max\{c^Tx \mid Ax \le b\}$ (wlog.) with n variables, suppose we are given a ~~stng~~ "strong separation" oracle for $K = \{x \mid Ax \le b\}$, ~~and~~ Then in polynomial time (poly in n, $\max_i\{\langle a_i\rangle, \langle b_i\rangle\}$) we can exactly optimize over the LP. (find an optimal bfs)

[The strong separation oracle: given $\hat{x} \in \mathbb{R}^n$, either say $\hat{x} \in K$ or output a hyperplane $a^Tx \le b$ st $K \subseteq \{x \mid a^Tx \le b\}$ and $a^T\hat{x} > b$.]
such a hyperplane must exist [Not hard to prove!] [intuitive!].

So basically if you can, given a proposed point $\hat{x} \in \mathbb{R}^n$, correctly tell me whether $\hat{x} \in K$ or not (and if not, tell me a violated constraint) then we can optimize over K.

$$\boxed{\text{separation} \Rightarrow \text{optimization}}$$

Cite: the Grötschel-Lovasz-Schrijver book contains all of the details about this and related theorems. It also shows that the separation and optimization problems are "essentially" equivalent. And also equivalent to just testing membership (under some technical conditions).

In the rest of the lecture, give some sense of how to optimize an arbitrary convex function over K to get error $+\varepsilon$ in time $O(\log 1/\varepsilon)$.

This convergence means that if we are looking for a solution whose bit complexity is polynomial (ie the numbers are only exponential) then we can find it in poly time (I am waving my hands a bit here, but this can be done. see GLS 88).

The Center-of-Gravity approach.

$f: K \to [-B, B]$. convex

Min (f, K) $\quad K_0 \leftarrow K$

- Compute the center of gravity of K_t (say c_t)
- Find the gradient (or subgradient) of f at c_t

Set $K_{t+1} \leftarrow K_t \cap \{x \mid \langle \nabla f(c_t), x - c_t \rangle \leq 0\}$

the optimum point must lie here.

Repeat for T times
return argmin $f(c_t) =: \hat{x}_T$

Yudin Nemirovski '76
N. Shor 70s

- Ellipsoid (Khachiyan) 79
- Interior point (Karmarkar).
  - many variants
  - most used in practice
  - may talk about it later.
- Center of gravity approach
  - cute and simple
    - but requires some heavy machinery to implement
- Geometric approaches
  - no dependence on numbers but time $\exp(\sqrt{n})$ right now.

Theorem: $f(\hat{x}_T) - f(x^*) \leq 4B(1-1/e)^{T/n}$

- where x^* is the minimizer of f in K
- B is the max value f can take (and $-B$ the least)
- n is the dimension of the space.

Pf: [Fact: [Grünbaum] Any hyperplane thru the center of gravity cuts the body into $(1-1/e, 1/e)$-balanced parts.]

Hence the volume falls exponentially. $\mathrm{vol}(K_t) \leq \mathrm{vol}(K) \cdot (1-1/e)^t$.

Now take $K_\varepsilon = \{(1-\varepsilon)x^* + \varepsilon x \mid x \in K\}$.

- value of f on any point in K_ε is at most $(1-\varepsilon) f(x^*) + \varepsilon(B)$.

$$\leq f(x^*) + \epsilon(B - f(x^*))$$
$$\leq f(x^*) + 2\epsilon B.$$

- volume of $K_\varepsilon = \epsilon^n \cdot \mathrm{vol}(K)$

Now if $(1-1/e)^{t/n} \leq \epsilon/2$ then some portion of K_ϵ got lopped off by some c_t that c_t had value $f(c_t) \leq f(x_{\text{lopped off}})$ for any $x_{\text{lopped off}}$ in K_ϵ.

$\Rightarrow$ we get: $f(\hat{x}_T) - f(x^*) \le 2B\epsilon + 4B(1-\frac{1}{e})^{T/n}$.

<u>Fly in the ointment</u>: How to compute center of gravity?

This algo was proposed by Levin [1965] & Walker [1965], but no idea how to implement it

[Bertsimas & Vempala 2005] used random walks in polytopes to sample enough points and show that you can estimate the C-of-G pretty well, enough to give a polytime algorithm for Constrained Convex Minimization

---

<u>Ellipsoid: A different</u> approach with a similar idea

<u>Basic Idea:</u> now just want to solve feasibility:- given some description of K, and guarantee that $K \subseteq Ball(0, R)$ for $R > 0$ and $\{Ball(c, r) \subseteq K$ for some c or $K = \emptyset$ (both r, R are given)

does $\exists$ any point in K?

[i.e. is K of volume 0 or volume $r^n . B(1)$?]

- If we can solve feasibility, can solve other problems too.
- <u>Say K is given by a separation oracle</u>

Start off: ~~some~~ the ball $B(0,R)$.

Is the center of $B(0,R)$ in K?

If yes, we are done.

If no, look at the separating hyperplane $a^T x \le b$ s.t. $K \subseteq \{x : a^T x \le b\}$ and center is not in.

Find a ball that contains $\underbrace{B(0,R) \cap \{a^T x \le b\}}$, this set must contain K.

☹ May be that the smallest ball containing this set is $B(0,R)$ itself.

☺ Use ellipsoids.

<u>Fact</u>: can show that the volume of the smallest-volume-ellipsoid containing the half-ball is smaller by an $e^{-1/2n}$ factor.

$\Rightarrow$ volume decreases exponentially. Must stop after $O(n^2 \log(R/r))$ iterations. because $vol(init)/vol(final) \le (R/r)^{n}$.

<u>For convex function minimization:</u> given f, K. and also R, r.

Each time:

if center $c_t \notin K$ then find a separating hyperplane. $K \subseteq \{a^T x \le b\}$

if $c_t \in K$ then find the gradient $\nabla f(c_t)$ st $opt \in \{x : \langle \nabla f(c_t), x - c_t \rangle \le 0\}$

Recurse on the correct side. After T steps output ~~smallest of~~ $x_T \leftarrow \underset{c_t \in K}{\operatorname{argmin}} f(c_t)$.

Similar analysis shows that:

$$f(x_T) - f(x^*) \le \frac{2BR}{r} \exp\left(-\frac{T}{2n^2}\right).$$

Let's do the simplest case just to get some familiarity with the process.

Suppose current ball is $\mathcal{E}_0 = B(0,R)$. and we want to find the smallest ellipsoid containing the right half-ball $\mathcal{E}_0 \cap \{x \mid x_1 \ge 0\}$.

<u>First</u>: how to represent ellipsoids?

e.g. $\frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1$ $\qquad \begin{pmatrix} x \\ y \end{pmatrix}^T \begin{pmatrix} 1/a^2 & 0 \\ 0 & 1/b^2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \le 1.$

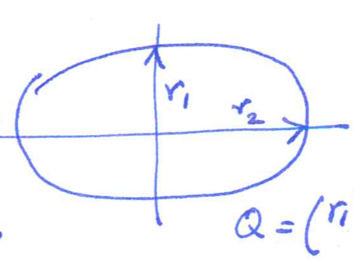
An ellipsoid is a linear transformation of a ball $B(0,1) = \{x : x^T x \le 1\}$.

$$\begin{aligned} L\,\text{Ball}(0,1) &= \{Lx : x \in B(0,1)\}. \\ &= \{y : L^{-1} y \in B(0,1)\} \\ &= \{y : (L^{-1}y)^T (L^{-1}y) \le 1\} \\ &= \{y : y^T (LL^T)^{-1} y \le 1\} = \{y : y^T Q^{-1} y \le 1\} \end{aligned}$$

(LL^T) is a psd matrix $\qquad$ Q psd matrix.

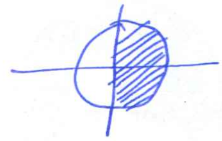
the standard ball is $\{y : y^T I^{-1} y \le 1\}$.

$$\boxed{\mathcal{E}(c, Q) = \{x : (x - c)^T Q^{-1} (x - c) \le 1\}}$$



$\text{vol}(L(A)) = \text{vol}(A) \cdot |\det(L)| = \sqrt{\det(Q)} \cdot \text{vol}(A).$ $\qquad Q = \begin{pmatrix} r_1^2 & \\ & r_2^2 \end{pmatrix}$

$\mathcal{E}_0 = B(0,1)$



and suppose separating hyperplane is $a_0 = (-1, 0, 0 \dots 0)$. $\Rightarrow$ want

want $\mathcal{E}_0 \cap \{x : x_1 \geq 0\}$. to be contained within $\mathcal{E}_1$ (of smallest volume.)

lemma: $c_1 = (\frac{1}{n+1}, 0, 0 \dots 0)$

and $Q_1 = \frac{n^2}{n^2-1}\begin{pmatrix} 1-2/n+1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}$ then $\mathcal{E}(c_1, Q_1)$ is the min volume ellipsoid containing the half-ball.

Also: $\frac{\mathrm{vol}(\mathcal{E}_1)}{\mathrm{vol}(\mathcal{E}_0)} \leq e^{-1/2(n+1)}$.

Pf: we'll not prove the min volume part. just that. half-ball $\subseteq \mathcal{E}(c_1, Q_1)$. and the volume ratio.

Fact: $Q_1^{-1} = \frac{n^2-1}{n^2}\begin{pmatrix} \frac{n+1}{n-1} & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}$ just do it!

Say $x \in$ Half ball $= (x_1 \geq 0, \tilde{x})$. with $\|x\|^2 \leq 1$

$$(x-c_1)^T Q^{-1} (x-c_1) = \frac{n^2-1}{n^2}\left(x_1 - \frac{1}{n+1}, \tilde{x}\right)^T \begin{pmatrix} \frac{n+1}{n-1} & & \\ & 1 & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} x_1 - \frac{1}{n+1} \\ \tilde{x} \end{pmatrix}$$

$$= \frac{n^2-1}{n^2}\left[\left(x_1 - \frac{1}{n+1}\right)^2 \cdot \frac{n+1}{n-1} + \|\tilde{x}\|^2\right]$$

$$\leq \frac{1}{n^2}\left[(x_1(n+1)-1)^2 + (n^2-1)(1-x_1^2)\right]$$

$$= \frac{1}{n^2}\left[x_1^2(n+1)^2 - 2x_1(n+1) + 1 + n^2 - 1 - n^2 x_1^2 + x_1^2\right]$$

$$= \frac{1}{n^2}\left[2x_1^2 - 2x_1\right] + 1 \leq 1. \quad \checkmark \quad \Rightarrow x \in \mathcal{E}_1$$

And: $\frac{\mathrm{vol}(\mathcal{E}_1)}{\mathrm{vol}(\mathcal{E}_0)} = \sqrt{\det(Q)} = \sqrt{\left(\frac{n^2}{n^2-1}\right)^n \left(\frac{n-1}{n+1}\right)} = \sqrt{\left(\frac{n^2}{n^2-1}\right)^{n-1} \cdot \left(\frac{n}{n+1}\right)^2}$

$$\leq \exp\left(\frac{n-1}{2} \cdot \frac{1}{n^2-1} - \frac{1}{n+1}\right) = \exp\left(\frac{-1}{2(n+1)}\right).$$

The general case: It's all to be subjected to an affine translation L

i.e. if $\mathcal{E}_k = L(B(0, c_k))$ then $\mathcal{E}_{k+1} = L($of the appropriate transformed half ball$)$.

Since the volumes scale the same, $\dfrac{vol(\mathcal{E}_{k+1})}{vol(\mathcal{E}_k)} = \dfrac{|\det L| \cdot vol(\mathcal{E}')}{|\det L| \cdot vol(B(0,1))} \le e^{-\frac{1}{2(n+1)}}$.

Suppose $\mathcal{E}_k = \mathcal{E}(c_k, Q_k)$ and the new ellipsoid has to contain

$\mathcal{E}_k \cap \{a_k^T x \cancel{\geq} \leq a_k^T c_k\}$. i.e hyperplane $a_k^T(x - c_k) \cancel{\geq} \leq 0$.

then $\mathcal{E}_{k+1} = (c_{k+1}, Q_{k+1})$ where $c_{k+1} \leftarrow c_k - \frac{1}{n+1} h_k$

and $Q_{k+1} = \dfrac{n^2}{n^2-1}\left(Q_k - \dfrac{2}{n+1} h_k h_k^T\right)$

where $h_k = \left(\sqrt{a_k^T Q_k a_k}\right)^{-1} Q_k a_k$

N.b. assume exact arithmetic — the real trick is to get all the numerical issues correct and under control. cf. Khachiyan, and Lovasz-Gacs. etc.
