Mathematical Programming II	**Lecture 17**
ORIE 6310 Spring 2014	**March 27, 2014**
Scribe: Jian Wu	

Theorem 1 *Let $C \subset \mathcal{R}^n$ be a convex body (compact with nonempty interior). Then there exist B and y with*

$$E(n^{-2}B, y) \subset C \subset E(B, y), \tag{1}$$

where $E(n^{-2}B, y)$ is $E(B, y)$ shrunk by a factor of n around its center.

Proposition 1 *For any convex body $C \subset \mathcal{R}^n$, there is an ellipsoid $E = E(B, y)$ of minimum volume containing C.*

Proof: We can formulate this as

$$\min_{B,y} \det B \tag{2}$$

$$(x - y)^T B^{-1}(x - y) \leq 1, \text{ for all } x \in C, \tag{3}$$

$$B \in \mathcal{R}^{n*n} \text{ symmetric, positive definite}, y \in \mathcal{R}^n. \tag{4}$$

Using $D = B^{-1}$ as the matrix variable, we get

$$\min_{D,y} -\det D \tag{5}$$

$$(x - y)^T D(x - y) \leq 1, \text{ for all } x \in C, \tag{6}$$

$$D \in \mathcal{R}^{n*n} \text{ symmetric, positive definite}, y \in \mathcal{R}^n. \tag{7}$$

Assume, without loss of generality, that $B(0, r) \subset C \subset B(0, R)$. Then $D = R^{-2}I$, $y = 0$ is feasible, so we can add the constraint $\det D \geq R^{-2n} > 0$ and relax the positive definite condition to positive semidefinite. Then the feasible region is closed and nonempty, and the objective function is continuous.

From $B(0, r) \subset C$, we find that for all $a \in \mathcal{R}^n$ with $||a|| = 1$, the width of the ellipsoid in direction a is at least that of the ball: $2\sqrt{a^T D^{-1} a} \geq 2r$. So all eigenvalues of D^{-1} are at least r^2, and so all eigenvalues of D are at most r^{-2}. $D = Q\Lambda Q^T$ is therefore bounded. Hence all eigenvalues of D are at least $\frac{R^{-2n}}{r^{-2(n-1)}} =: \epsilon$. Then

$$\epsilon y^T y \leq y^T D y \leq 1, \tag{8}$$

since $x = 0 \in C$. So $||y|| \leq \epsilon^{-1/2}$ for all feasible D, y. Hence (D, y) is bounded for all feasible (D, y), so the optimization problem has an optimal solution.

Proof: Proof of the theorem.

Claim: We can choose B, y such that $E(B, y)$ is the minimum-volume ellipsoid containing C (which exists by the proposition).

Suppose not, so $E(n^{-2}B, y) \not\subset C$. So there exists $x \in E(n^{-2}B, y)$ but $x \notin C$. Choose a with $a^T Ba = 1$ and $a^T x > \beta \geq \max\{a^T z : z \in C\}$. Note $a^T x \leq \max\{a^T z : z \in E(n^{-2}B, y)\} = a^T y + \frac{1}{n}$.

So $z \in C$ implies $a^T z \leq \beta = a^T y - \alpha$ for $\alpha > -\frac{1}{n}$. Hence by Proposition 3 from last time, C can be enclosed in an ellipsoid of smaller volume than $E(B, y)$, a contradiction.

Corollary 1 *Given a convex body $C \subset \mathcal{R}^n$, the minimum-volume ellipsoid containing it, E^+, has*

$$\text{vol}(E^+) \leq n^n \text{vol}(C),$$

and the maximum-volume ellipsoid contained in it, E^-, has

$$\text{vol}(E^-) \geq n^{-n} \text{vol}(C).$$

We have a lower bound of $\Omega(n \ln \frac{2}{\delta\epsilon})$ on the complexity of oracle-algorithms for problem (f, G). We have the upper bound $2.2n \ln \frac{2}{\delta\epsilon}$ from MCG, which is hard to implement. We also have the upper bound $2n(n+1) \ln \frac{2\sqrt{n}}{\delta\epsilon}$ from the ellipsoid method, which is easy to implement. Can we get the better complexity of MCG with an algorithm that is easier to implement?

We now briefly discuss the Method of Inscribed Ellipsoids (MIE) (Tarasov, Khachiyan, Erlikh, 1988). It's based on a different notion of center and a different notion of size. It generates a sequence $\{(H_k, z_k)\}$ of localizers with each H_k a polyhedron, $H_0 = C$.

We compute the maximum-volume ellipsoid contained in H_k, say E_k with center x_k. We call the oracle at x_k, and then update as in MCG. We also measure the size of H_k by $\text{vol}'(H_k) := \text{vol}(E_k)$.

Key fact: for any half-space with x_k on its boundary, the resulting H_{k+1} has vol' at most 0.844 times that of H_k.

Note also E_0 is $B(0, 1)$, so $\text{vol}'(H_0) \leq 2^n$ and if $\text{vol}(G) \geq \delta^n$, then $\text{vol}'(G) \geq n^{-n}\delta^n$ by the corollary, and similarly, $\text{vol}(G(\epsilon)) \geq (\delta\epsilon)^n$, so $\text{vol}'(G(\epsilon)) \geq n^{-n}(\delta\epsilon)^n$. Hence by the same arguments as for MCG, we can get

Theorem 2 *If the MIE takes more than $cn \ln \frac{2n}{\delta}$ steps for some constant c and $z_k = *$, then $G = \emptyset$. If it produces $z_k \in G$, then it gets $\epsilon(z_k, f, G) \leq \epsilon$ in $cn \ln \frac{2n}{\delta\epsilon}$ steps.*

In fact, given H_k defined by $\mathcal{O}(n \ln n)$ constraints we can approximate x_k in $\mathcal{O}(n^{3.5+\epsilon})$ arithmetic operations using an interior-point method (Khachiyan-Todd, Anstreicher).

There are three conceptual concerns about the ellipsoid method for LP.

(1) Does it exploit sparsity? Maybe each a_k will be sparse, but $B_k a_k$ not so sparse, and then $B_{k+1} = \delta(B_k - \sigma \frac{B_k a_k a_k^T B_k}{a_k^t B_k a_k})$ becomes increasingly dense (compare with the simplex method and interior-point methods).

(2) How do we know E_k contains the feasible region after many iterations? We would like a certificate.

(3) If the half-space generated misses the current ellipsoid completely, the problem is infeasible, but again we want a certificate.