Mathematical Programming II	**Lecture 15**
ORIE 6310 Spring 2014	**March 18, 2014**
Scribe: Raphael Louca	

Last time, we introduced the Method of Centers of Gravity (MCG) and the Ellipsoid Method of Yudin and Nemirovski (1976) and Shor (1977), which is a simple modification of MCG.

We represent an ellipsoid by

$$E(B, y) := \{x \in \mathbb{R}^n : (x - y)^\top B^{-1}(x - y) \leq 1\},$$

where $y \in \mathbb{R}^n$, and $B = B^\top \in \mathbb{R}^{n\times n}$ is symmetric positive definite.

Observe that $C = [-1, 1]^n \subseteq E(nI, 0) = \left\{x \in \mathbb{R}^n : x^\top \frac{1}{n} Ix \leq 1\right\} = \{x \in \mathbb{R}^n : \|x\|_2 \leq \sqrt{n}\}$ (Figure 1). Hence, if $E_0 = E(nI, 0)$, then $(E_0, *)$ is a localizer and $\text{vol}(E_0) \leq (2\sqrt{n})^n$.

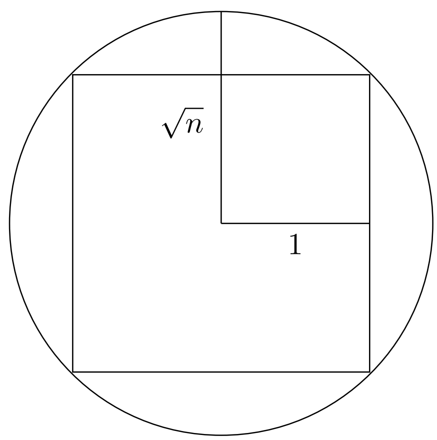

Figure 1: $E(nI, 0) = \left\{x \in \mathbb{R}^n : x^\top \frac{1}{n} Ix \leq 1\right\} = \{x \in \mathbb{R}^n : \|x\|_2 \leq \sqrt{n}\}$.

The Ellipsoid Method Algorithm:

- Initialize with $E_0 = E(nI, 0),\ z = *$.
- At iteration k, we are given a localizer (E_k, z_k), where $E_k = E(B_k, x_k)$ and x_k is the center of gravity of E_k. Call the oracle at x_k.
- If $x_k \notin G \cap \text{int}(C)$, the oracle returns a separating hyperplane $G \subseteq \{x : v_k^\top x \leq v_k^\top x_k\}$. Set $z_{k+1} := z_k,\ a_k := v_k$.
- If $x_k \in G \ \cap \ \text{int}(C)$, the oracle returns $f(x_k),\ g(x_k) \in \partial f(x_k)$. Set $z_{k+1} := \text{argmin}\{f(x_k), f(z_k)\}$, and $a_k := g(x_k)$.
- Set E_{k+1} to be the minimum volume ellipsoid containing $E_k^{1/2} := \{x \in E_k : a_k^\top x \leq a_k^\top x_k\}$ (Figure 2).

- Stop if $\mathrm{vol}(E_{k+1}) < \delta^n$ and $z_{k+1} = *$ (then $G = \emptyset$) or if $z_{k+1} \in G$ and $\mathrm{vol}(E_{k+1}) < (\varepsilon\delta)^n$ (then $\epsilon(z_{k+1}, f, G) < \epsilon$).

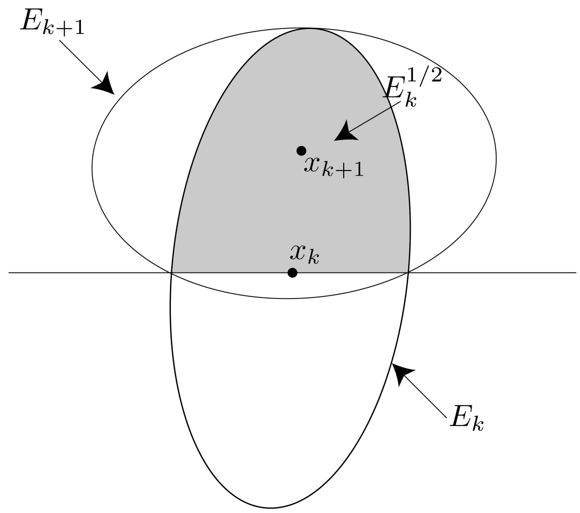

Figure 2: Ellipsoid Method

Proposition 1. For every k, (E_k, z_k) is a localizer of (f, G).

Proof. The proof is exactly analogous to the proof of the Method of Centers of Gravity from last lecture (i.e., Proposition 2). □

Proposition 2. $E_{k+1} = E(B_{k+1}, x_{k+1})$, where

$$x_{k+1} = x_k - \tau \frac{B_k a_k}{\sqrt{a_k^\top B_k a_k}}, \qquad B_{k+1} = \delta \left(B_k - \sigma \frac{B_k a_k a_k^\top B_k}{a_k^\top B_k a_k} \right),$$

with $\tau = \frac{1}{n+1}$, $\delta = \frac{n^2}{n^2-1}$, and $\sigma = \frac{2}{n+1}$.

Note that we are always assuming that $n > 1$. In fact, we will prove a stronger proposition for which Proposition 2 is a special case ($\alpha = 0$).

Proposition 3. Let $E = E(B, y)$ and $-\frac{1}{n} \leq \alpha < 1$. Then the minimum volume ellipsoid containing $E_\alpha = \{x \in E : a^\top x \leq a^\top y - \alpha\sqrt{a^\top B a}\}$ is $E_+ := E(B_+, y_+)$, where

$$B_+ = \delta \left(B - \sigma \frac{B a a^\top B}{a^\top B a} \right) \quad \text{and} \quad y_+ = y - \tau \frac{Ba}{\sqrt{a^\top B a}}.$$

Here, $\tau = \frac{1+n\alpha}{n+1}$, $\delta = \frac{(1-\alpha^2)n^2}{n^2-1}$, and $\sigma = \frac{2(1+n\alpha)}{(n+1)(1+\alpha)}$. Moreover,

$$\frac{\mathrm{vol}(E_+)}{\mathrm{vol}(E)} = \frac{n}{n+1} \left(\frac{n^2}{n^2-1} \right)^{\frac{n-1}{2}} (1-\alpha)(1-\alpha^2)^{\frac{n-1}{2}}.$$

This is at most $\exp\left(-\frac{1}{2(n+1)}\right)$ if $\alpha \geq 0$.

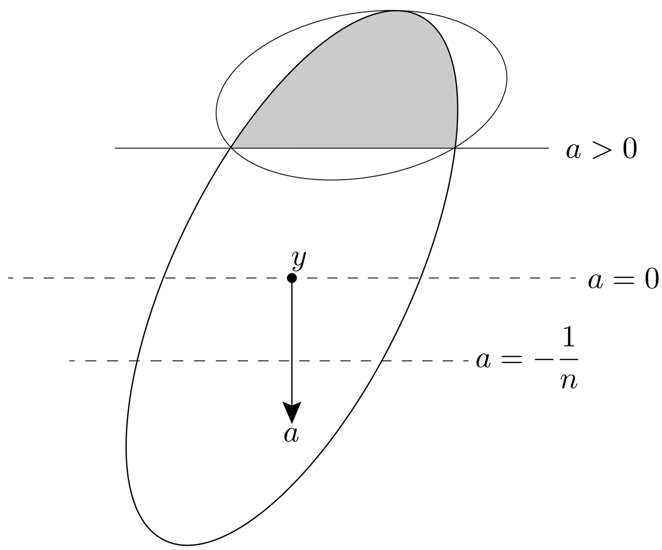

Figure 3: Minimum Volume Ellipsoid. If $\alpha \le -\frac{1}{n}$, $E_+ = E$.

We'll prove Proposition 3 next time using the following Lemmata.

Lemma 1. Let $B = B^\top \in \mathbb{R}^{n\times n}$ be symmetric positive definite. Then:

(a) B has a unique symmetric positive definite square root, denoted by $B^{1/2}$, satisfying $B^{1/2}B^{1/2} = B$.

(b) $E(B, y) = \{y + B^{1/2}w : \|w\|_2 \le 1\}$.

(c) $\text{vol}(E(B, y)) = \sqrt{\det(B)}\ \text{vol(unit ball)}$.

(d) For any $0 \ne a \in \mathbb{R}^n$, $a^\top x$ is maximized/minimized over $E(B, y)$ at $y \pm \frac{Ba}{\sqrt{a^\top Ba}}$ with optimal values $a^\top y \pm \sqrt{a^\top Ba}$.

Proof.

(a) Write $B = QDQ^\top$, with Q orthogonal (the columns of Q are the eigenvectors of B) and D diagonal (the diagonal entries are the eigenvalues of B). Then, set $D^{1/2}$ to be the diagonal matrix containing the positive square roots of the eigenvalues and set $B^{1/2} = QD^{1/2}Q^\top$. Check that $B = B^{1/2}B^{1/2}$. We will not prove uniqueness here.

(b) Write $B^{-1/2} = (B^{1/2})^{-1} = (B^{-1})^{1/2}$. If $x = y + B^{1/2}w$, $\|w\| \le 1$, then

$$(x - y)^\top B^{-1}(x - y) = w^\top B^{1/2}(B^{1/2}B^{1/2})^{-1}B^{1/2}w = w^\top w \le 1.$$

Conversely, if $(x - y)^\top B^{-1}(x - y) = (x - y)^\top B^{-1/2}B^{-1/2}(x - y) \le 1$. Define $w := B^{-1/2}(x - y)$ and note that

$$\|w\| = \|B^{-1/2}(x - y)\| \le 1.$$

(c) $\mathrm{vol}(E(B,y)) = \mathrm{vol}(\{y + B^{1/2}w : \|w\| \le 1\} = \det B^{1/2}\, \mathrm{vol}(\text{unit ball}) = \sqrt{\det(B)}\, \mathrm{vol}(\text{unit ball})$.

(d) Note that

$$\begin{aligned}\max\{a^\top x : x \in E(B,y)\} &= \max\{a^\top (y + B^{1/2}w) : \|w\| \le 1\} \\ &= a^\top y + \max\{a^\top B^{1/2} w : \|w\| \le 1\} \\ &\overset{(i)}{=} a^\top y + \|B^{1/2}a\|,\end{aligned}$$

where (i) follows by setting $w = \frac{B^{1/2}a}{\|B^{1/2}a\|} = \frac{B^{1/2}a}{\sqrt{a^\top Ba}}$, whence the result for maximizing. Similarly for minimizing. This gives both the optimal value and the optimizing point.

□

Lemma 2. (Sherman - Morrison - Woodbury) Let $A \in \mathbb{R}^{n\times n}$ be invertible and $U,\ V \in \mathbb{R}^{n\times k}$, where $k < n$. Then

(a) $\det(A + UV^\top) = \det(A)\det(I_k + V^\top A^{-1}U)$.

(b) $A + UV^\top$ is invertible if and only if $I_k + V^\top A^{-1}U$ is invertible.

(c) If this holds,

$$(A + UV^\top)^{-1} = A^{-1} + \underbrace{A^{-1}U(I_k + V^\top A^{-1}U)^{-1}V^\top A^{-1}}_{\text{rank}\le k}.$$

Proof.

(a) Consider

$$\begin{aligned}\begin{bmatrix} A & U \\ -V^\top & I_k \end{bmatrix} &= \begin{bmatrix} A + UV^\top & U \\ 0 & I_k \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -V^\top & I_k \end{bmatrix} \\ &= \begin{bmatrix} I_n & 0 \\ -V^\top A^{-1} & I_k \end{bmatrix} \begin{bmatrix} A & U \\ 0 & I_k + V^\top A^{-1}U \end{bmatrix}.\end{aligned}$$

By taking determinants of the two RHS, we conclude that $\det(A + UV^\top) = \det(A)\det(I_k + V^\top A^{-1}U)$.

(b) Follows trivially from (a).

(c) Define $B := I_k + V^\top A^{-1}U$. Then from the equations above,

$$\begin{bmatrix} A + UV^\top & U \\ 0 & I_k \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -V^\top A^{-1} & I_k \end{bmatrix} \begin{bmatrix} A & U \\ 0 & B \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -V^\top & I_k \end{bmatrix}^{-1}.$$

By inverting both sides of the equation, we have,

$$\begin{bmatrix} (A + UV^\top)^{-1} & -(A + UV^\top)^{-1}U \\ 0 & I_k \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -V^\top & I_k \end{bmatrix} \begin{bmatrix} A^{-1} & -A^{-1}UB^{-1} \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} I_n & 0 \\ V^\top A^{-1} & I_k \end{bmatrix}.$$

The top left hand block gives the desired result.

□

Next time, we will prove Proposition 3. We will first show that E_+ contains E_α and then sketch the proof of its minimality. Define $\overline{a} := \frac{a}{\sqrt{a^\top B a}}$ and note that $\overline{a}^\top B \overline{a} = 1$. Therefore, by Lemma 1, $-1 \leq \overline{a}^\top (x-y) \leq 1$, for all $x \in E$. Then,

$$
\begin{aligned}
&a^\top x \leq a^\top y - \alpha\sqrt{a^\top B a} \text{ is equivalent to} \\
&\overline{a}^\top x \leq \overline{a}^\top y - \alpha.
\end{aligned}
$$

Therefore, $x \in E_\alpha$ if and only if $(x-y)^\top B^{-1}(x-y) \leq 1$ and $-1 \leq \overline{a}^\top (x-y) \leq -\alpha$ or equivalently, $(\overline{a}^\top (x-y) + \alpha)(\overline{a}^\top (x-y) + 1) \leq 0$. Thus we have two quadratics.