Another Interpretation/Implementation of the Ellipsoid Method

This section is based on the paper "The ellipsoid method generates dual variables" (found on the course webpage) by Burrell and Todd (1985).

Suppose we want to find $x \in P := \{x \in \mathbb{R}^n : A^\top x \leq b, -e \leq x \leq e\}$. We have three issues to deal with regarding the implementation of the ellipsoid method for linear programming:

(i) If A is sparse, can we get a sparse implementation?

(ii) Can we "certify" at every iteration that $P \subseteq E_k$? That is, we need to check that round-off errors did not affect too much our implementation.

(iii) If the half-space generated completely misses E_k, can we find a certificate of infeasibility?

In the following, we will address the above three problems.
Solution to (i): At every iteration,

$$B_k = \delta \left(B_k - \sigma \frac{B_k a_k a_k^\top B_k}{a_k^\top B_k a_k} \right),$$

where a_k is a column of $\bar{A} := (A, I)$ (note that we do not need to consider columns of $-I$ since the above term is quadratic in a_k). So,

$$B_{k+1}^{-1} = \delta^{-1} \left(B_k^{-1} + \frac{\sigma}{1-\sigma} \frac{a_k a_k^\top}{a_k^\top B_k a_k} \right).$$

Note: $B_0^{-1} = nI$. So at every iteration, $B_k^{-1} = \bar{A} D_k \bar{A}^\top$ for some diagonal positive semidefinite D_k. This is exactly the form of the matrix appearing in interior-point methods, and is often sparse. But if we store and update B_k^{-1}, we need to solve a system to get $B_k a_k$ ($B_k a_k$ is the solution of $B_k^{-1} z = a_k$) !!
But we can update a Cholesky factorization of $B_k^{-1} = \mu_k L_k L_k^\top$ in $\mathcal{O}(n^2)$ work only. Here, μ_k is a scalar, and L_k is a lower triangular matrix, so that the system can be solved efficiently.

Solution to (ii): Note that $P = \{x \in \mathbb{R}^n : \ell \leq \bar{A}^\top x \leq u\}$ (we can lower bound $\bar{A}^\top x$ since x is bounded).
So $\bar{A}^\top x - \ell \geq 0$, and $\bar{A}^\top x - u \leq 0$. Thus $(\bar{A}^\top x - \ell)^\top (\bar{A}^\top x - u) \leq 0$. More generally, for every positive semidefinite diagonal D, $x \in P \Rightarrow (\bar{A}^\top x - \ell)^\top D (\bar{A}^\top x - u) \leq 0$, or

$(x-y)^\top \bar{A} D \bar{A}^\top (x-y) \le \beta$, with $y := (\bar{A} D \bar{A}^\top)^{-1} \bar{A} D \frac{\ell+u}{2}$, and $\beta := y^\top (\bar{A} D \bar{A}^\top) y - \ell^\top D u$ (as long as $\bar{A} D \bar{A}^\top$ is positive definite).

Note that y solves the weighted least-squares problem:

$$D^{1/2} \bar{A}^\top y \approx D^{1/2} \frac{\ell+u}{2}.$$

In fact, the initial ellipsoid E_0 is of this form:

$$E_0 = \hat{E}(D, \ell) := \{x \in \mathbb{R}^n : (\bar{A}^\top x - \ell)^\top D (\bar{A}^\top x - u) \le 0\},$$

for $\ell = \begin{bmatrix} \vdots \\ -e \end{bmatrix}$, and $D = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$.

At iteration k, we have $E_k = \hat{E}(D_k, \ell_k)$: quit if $x_k \in P$, otherwise generate a violated inequality $a^\top x \le \upsilon$. Note: a lower bound on $a^\top x$ comes from E_k (see lecture 15).

E_k is certified by D_k and ℓ_k. So can we get a lower bound directly certified by P?

So improve the lower bound if necessary, and then do a two-sided cut (see Figure 1) to get $E_{k+1} = \hat{E}(D_{k+1}, \ell_{k+1})$. Also, the lower bounds ℓ_k are always certified by duality.

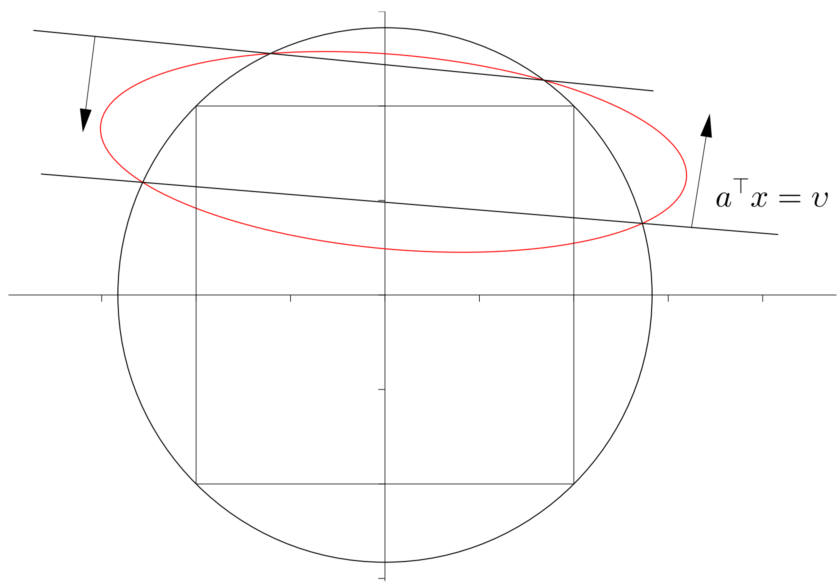

Figure 1: Illustration of the two-sided cut: the top line corresponds to the improved lower bound, the line below corresponds to the violated constraint, and the red ellipse is then E_{k+1}.

Solution to (iii): If at any iteration, we generate $\lambda \ge \upsilon$ (λ from duality always cuts the current ellipsoid), then we can generate a certificate of infeasibility. See Figure 2.

Equivalence of Separation and Optimization

This topic is thoroughly discussed in Grotschel, Lovasz, and Schrijver's book: Geometric Algorithms and Combinatorial Optimization, Springer 1988.

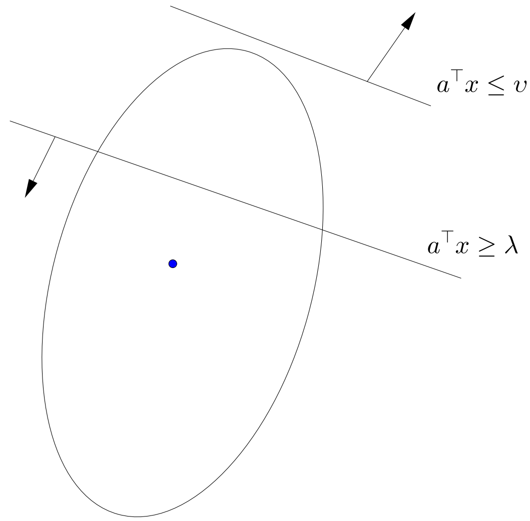

Figure 2: The generated half-space completely misses E_k.

Suppose we have a convex body $G \subseteq \mathbb{R}^n$ with $B(0,r) \subseteq G \subseteq B(0,R)$, with $R/r \geq 2$. We call such a body *well-rounded*.
Consider:

Strong separation problem: Given $x \in \mathbb{R}^n$, either state $x \in G$ or produce $v \in \mathbb{R}^n$ with $\max\{v^\top z : z \in G\} \leq 1$ but $v^\top x > 1$.

Weak optimization problem: Given $\epsilon > 0$ and $c \in \mathbb{R}^n$, $\|c\| = 1$, find $x \in G$ with $c^\top x \geq \max\{c^\top z : z \in G\} - \epsilon$.

Theorem 1 *If we can solve the strong separation problem in time polynomial in n, $\ln \frac{1}{r}$, and $\ln R$, then we can solve the weak optimization problem in time polynomial in n, $\ln \frac{1}{r}$, $\ln R$, and $\ln \frac{1}{\epsilon}$.*

Proof: Apply the (original) ellipsoid method, starting with $E_0 = B(0,R)$. At each iteration, we have the center x_k and in place of calling the oracle, we solve the strong separation problem with $x = x_k$. If $x_k \in G$, set $a_k = -c$, and if not, set $a_k = v$. Within $2n(n+1)\ln\left(\frac{R}{\epsilon r}\right)$ iterations, the volume of the ellipsoid is decreased below that of a ball of radius ϵr. But then $\varepsilon(z_k, -c^\top \cdot, G) \leq \epsilon$.
Since each iteration is polynomial in n, $\ln \frac{1}{r}$, and $\ln R$, we get the desired conclusion. $\square$

Next time, we'll relate the following two problems:
Strong optimization problem: Given $c \in \mathbb{R}^n$, $\|c\| = 1$, find $x \in G$ with $c^\top x = \max\{c^\top z : z \in G\}$.

Weak separation problem: Given $x \in \mathbb{R}^n$, $\frac{1}{2} \geq \eta > 0$, either state $x \in G$ or produce $v \in \mathbb{R}^n$ with

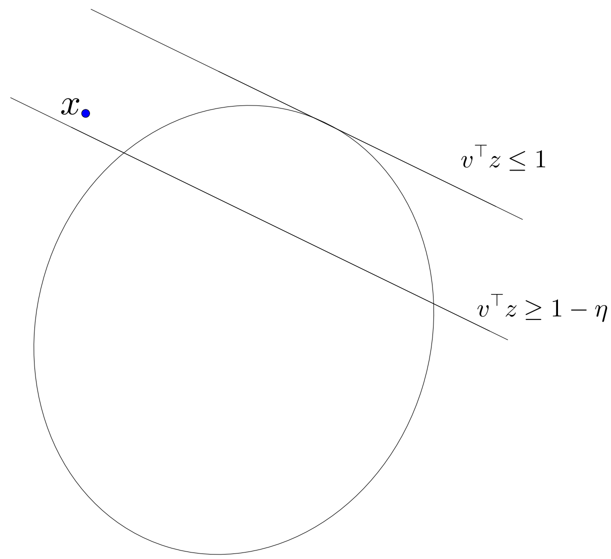

Figure 3: Illustration of the weak separation problem.

$\max\{v^\top z : z \in G\} \leq 1$ and $v^\top x > 1 - \eta$.

Define the polar of a convex set:
Given convex G, define its polar G^* by $G^* = \{y \in \mathbb{R}^n : x^\top y \leq 1 \ \ \forall x \in G\}$.
Examples:

(1) If G is a cube (which is the L_∞ ball), then G^* is the cross-polytope, i.e., the L_1 ball.

(2) If G is the L_2 ball, then $G^* = G$. More generally,

(3) If G is the ellipsoid $\{x \in \mathbb{R}^n : x^T B^{-1} x \leq 1\}$, then G^* is the ellipsoid $\{x \in \mathbb{R}^n : x^T B x \leq 1\}$.