

---

## Lecture 6: Cutting-plane Methods

---

### 6.1 Optimization, Oracles, and a One-Dimensional Example

We study the convex optimization problem

$$\min_{x \in X} f(x),$$

where  $X \subseteq \mathbb{R}^n$  is a convex body and  $f : X \rightarrow \mathbb{R}$  is convex. For the main development below, we assume that the optimal value

$$f^* := \min_{x \in X} f(x)$$

is known, and that we have access to a first-order oracle for  $f$ , meaning that at each query point  $x \in X$  it returns both the value  $f(x)$  and some subgradient  $g \in \partial f(x)$ . Later, when we recast the problem as a cutting-plane search problem, we will also discuss separation oracles for convex target sets. The distinction matters: the center-of-mass method needs a far more explicit description of the current search region than a separation oracle alone can provide, while the ellipsoid method can be implemented from separation cuts together with an explicit outer ellipsoid. Since  $X$  is compact and  $f$  is real-valued convex, the quantity

$$\Delta := \sup_{x \in X} f(x) - \inf_{x \in X} f(x)$$

is finite.

**Definition 6.1** ( $\delta$ -optimal point). For  $\delta \geq 0$ , a point  $x \in X$  is called  $\delta$ -optimal for the problem

$$\min_{y \in X} f(y)$$

if

$$f(x) \leq \inf_{y \in X} f(y) + \delta.$$

We will formulate the guarantees in the following offline form: after  $T$  iterations, one of the first  $T$  query points is guaranteed to be  $\delta$ -optimal. In particular, this also bounds the number of first-order oracle calls by  $T$ .

The basic idea already appears in one dimension. Let  $X = [a, b] \subseteq \mathbb{R}$ . If we query the midpoint  $c = (a + b)/2$ , then the sign of a subgradient at  $c$  tells us on which side a minimizer must lie. Thus optimization has been turned into a search problem: at each step we maintain an interval known to contain at least one minimizer, and we cut that interval in half.



Figure 1: In one dimension, querying the center of mass means querying the midpoint. The separation rule returns one of two half-intervals, and the localization length is cut exactly in half.

**Lemma 6.1** (One-dimensional subgradient cut). *Let  $f : [a, b] \rightarrow \mathbb{R}$  be convex, let  $c = (a + b)/2$ , and let  $g \in \partial f(c)$ .*

1. *If  $g > 0$ , then  $f$  has a minimizer in  $[a, c]$ .*
2. *If  $g < 0$ , then  $f$  has a minimizer in  $[c, b]$ .*
3. *If  $g = 0$ , then  $c$  is a minimizer of  $f$  on  $[a, b]$ .*

*Proof of Lemma 6.1.* For any  $y \in [a, b]$ , the subgradient inequality gives

$$f(y) \geq f(c) + g(y - c).$$

If  $g > 0$  and  $y \in [c, b]$ , then  $f(y) \geq f(c)$ , so

$$\min_{y \in [c, b]} f(y) \geq f(c) \geq \min_{y \in [a, c]} f(y).$$

Since  $f$  attains a minimum on the compact interval  $[a, c]$ , there is a minimizer in  $[a, c]$ . The case  $g < 0$  is symmetric. If  $g = 0$ , then  $f(y) \geq f(c)$  for every  $y \in [a, b]$ , so  $c$  is a minimizer.  $\square$

**Theorem 6.2** (One-dimensional cutting-plane complexity). *Let  $f : [a_0, b_0] \rightarrow \mathbb{R}$  be convex. Start from  $I_0 = [a_0, b_0]$ , and recursively define  $I_{t+1}$  by querying the midpoint  $c_t$  of  $I_t$  and updating according to Lemma 6.1. Then each interval  $I_t$  contains a minimizer of  $f$ , and*

$$|I_T| = 2^{-T} |I_0|.$$

*In particular, if  $x^*$  is any minimizer of  $f$ , then for every  $x \in I_T$ ,*

$$|x - x^*| \leq 2^{-T} |I_0|.$$

*Proof of Theorem 6.2.* By Lemma 6.1, each update keeps at least one minimizer in the surviving interval and halves its length. Thus, by induction,

$$|I_T| = 2^{-T} |I_0|.$$

The final claim follows because both  $x$  and  $x^*$  belong to the same interval  $I_T$ .  $\square$

## 6.2 Feasibility, Separation Oracles, and Outer Approximations

The one-dimensional example suggests a more general search problem.

**Definition 6.2** (Convex feasibility problem). Let  $S_0 \subseteq \mathbb{R}^n$  be a convex body, and let  $K \subseteq S_0$  be an unknown convex target set. The feasibility problem is to find a point  $x \in K$ .

**Definition 6.3** (Separation oracle for a target set). A separation oracle for  $K \subseteq S_0$  is a procedure that, on input  $x \in S_0$ , returns one of the following:

1. a certificate that  $x \in K$ ; or
2. an open halfspace  $H \subseteq \mathbb{R}^n$  such that

$$K \subseteq H \quad \text{and} \quad x \notin H.$$

Any halfspace  $H$  returned in the second case will be called a valid cut for  $K$ .

---

### Algorithm 1 Generic cutting-plane method

---

**Require:** An initial convex body  $S_0 \subseteq \mathbb{R}^n$  and a separation oracle for an unknown target set  $K \subseteq S_0$ .

```

1: for  $t = 0, 1, 2, \dots$  do
2:   Choose a query point  $x_t \in S_t$  and query the separation oracle at  $x_t$ , for example one of:
3:   Center-of-mass method:  $x_t$  is the center of mass of  $S_t$ .
4:   Ellipsoid method: if  $S_t$  is an ellipsoid,  $x_t$  is its center.
5:   if the separation oracle certifies that  $x_t \in K$  then
6:     Output  $x_t$  and stop.
7:   else
8:     Receive a valid cut  $H_t$  for  $K$ .
9:     Construct a new outer approximation  $S_{t+1}$  satisfying  $S_t \cap H_t \subseteq S_{t+1}$ , for example one of:
10:    Center-of-mass method:  $S_{t+1} := S_t \cap H_t$ .
11:    Ellipsoid method: if  $S_t$  is an ellipsoid, take  $S_{t+1}$  as the ellipsoid defined in Theorem 6.9.
12:   end if
13: end for

```

---

**Proposition 6.3** (Volume threshold implies success). *Consider the generic cutting-plane method of this section. Assume that the returned halfspaces are all valid cuts for a target set  $K \subseteq S_0$ . If, after  $T$  iterations, the algorithm has not yet queried a point of  $K$ , then*

$$K \subseteq S_T.$$

Consequently, if

$$\text{vol}(S_T) < \text{vol}(K),$$

then one of the first  $T$  query points must already lie in  $K$ .

*Proof of Proposition 6.3.* We prove by induction that  $K \subseteq S_t$  as long as no queried point has yet been certified to lie in  $K$ . This is true at  $t = 0$  because  $K \subseteq S_0$ . If  $K \subseteq S_t$  and the separation oracle returns a valid cut  $H_t$ , then

$$K \subseteq S_t \cap H_t \subseteq S_{t+1}.$$

Thus  $K \subseteq S_T$ . If additionally  $\text{vol}(S_T) < \text{vol}(K)$ , this is impossible.  $\square$

### 6.3 Sublevel Sets and Optimization Reduction

**Definition 6.4** (delta-approximate optimal set). For  $\delta \geq 0$ , define

$$K_\delta := \{x \in X : f(x) \leq f^* + \delta\}.$$

Thus finding a point in  $K_\delta$  is exactly the same as finding a  $\delta$ -optimal point. Equivalently,  $x \in K_\delta$  if and only if  $x$  is  $\delta$ -optimal.

**Proposition 6.4** (A separation oracle for  $K_\delta$ ). Fix  $\delta \geq 0$ . The separation oracle for  $X$  together with a first-order oracle for  $f$  implement a separation oracle for  $K_\delta$ :

1. If  $x_t \notin X$ , apply the separation oracle for  $X$ . The returned halfspace  $H_t$  satisfies

$$K_\delta \subseteq X \subseteq H_t.$$

2. If  $x_t \in X$  and  $f(x_t) \leq f^* + \delta$ , then  $x_t \in K_\delta$ .
3. If  $x_t \in X$  and  $f(x_t) > f^* + \delta$ , take  $g_t \in \partial f(x_t)$  and return the halfspace

$$H_t := \{x \in \mathbb{R}^n : \langle g_t, x - x_t \rangle < 0\}.$$

Then

$$K_\delta \subseteq H_t.$$

*Proof of Proposition 6.4.* The first two claims are immediate from the definitions of  $X$  and  $K_\delta$ . For the third, let  $x \in K_\delta$ . Then

$$f(x) \leq f^* + \delta < f(x_t).$$

The subgradient inequality at  $x_t$  gives

$$f(x) \geq f(x_t) + \langle g_t, x - x_t \rangle.$$

Hence

$$\langle g_t, x - x_t \rangle \leq f(x) - f(x_t) < 0,$$

so  $x \in H_t$ .  $\square$

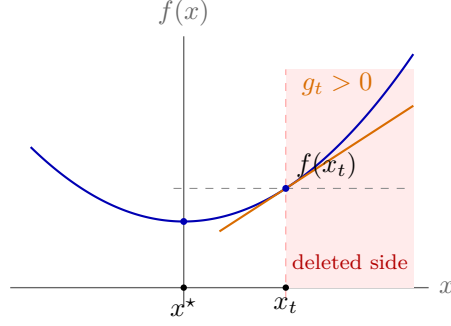


Figure 2: A one-dimensional picture of the subgradient cut when the subgradient at  $x_t$  is positive. The cut deletes the side  $x > x_t$ , because for every such point the subgradient inequality gives  $f(x) \geq f(x_t) + g_t(x - x_t) > f(x_t)$ . The kept side  $x \leq x_t$  is not certified to be good: it still contains points to the left of the minimizer whose objective value is larger than  $f(x_t)$ . The case  $g_t < 0$  is symmetric.

To turn this into a quantitative statement, we need a lower bound on the volume of a  $\delta$ -approximate optimal set. For this purpose alone, introduce the homothetic witness set

$$W_\alpha := (1 - \alpha) \arg \min_{x \in X} f(x) + \alpha X, \quad \alpha \in (0, 1].$$

**Proposition 6.5** (A witness-set lower bound for a delta-approximate optimal set). *For  $\alpha \in (0, 1]$ ,*

$$W_\alpha \subseteq K_{\alpha\Delta} \quad \text{and} \quad \text{vol}(K_{\alpha\Delta}) \geq \alpha^n \text{vol}(X).$$

*Proof of Proposition 6.5.* Let  $y \in W_\alpha$ . Then

$$y = (1 - \alpha)z + \alpha x$$

for some  $z \in \arg \min_{u \in X} f(u)$  and some  $x \in X$ . By convexity,

$$f(y) \leq (1 - \alpha)f(z) + \alpha f(x) = (1 - \alpha)f^* + \alpha f(x) \leq f^* + \alpha\Delta$$

so  $y \in K_{\alpha\Delta}$ . Thus  $W_\alpha \subseteq K_{\alpha\Delta}$ .

Now choose any minimizer  $x^* \in \arg \min_{x \in X} f(x)$ . The affine map

$$T_\alpha(x) := (1 - \alpha)x^* + \alpha x$$

has linear part  $\alpha I$ , so its Jacobian determinant equals  $\alpha^n$ . Since

$$T_\alpha(X) = (1 - \alpha)x^* + \alpha X \subseteq W_\alpha,$$

we have

$$\text{vol}(W_\alpha) \geq \text{vol}(T_\alpha(X)) = \alpha^n \text{vol}(X).$$

Because  $W_\alpha \subseteq K_{\alpha\Delta}$ , it follows that

$$\text{vol}(K_{\alpha\Delta}) \geq \text{vol}(W_\alpha) \geq \alpha^n \text{vol}(X).$$

□

## 6.4 Center of Mass in High Dimension

The midpoint is the center of mass of an interval. Its high-dimensional analogue is therefore the center of mass of the current search region.

**Definition 6.5** (Center of mass). Let  $S \subseteq \mathbb{R}^n$  be measurable, bounded, and of positive volume. Its center of mass is

$$\text{center}(S) := \frac{1}{\text{vol}(S)} \int_S x \, dx.$$

**Theorem 6.6** (Grünbaum's theorem). Let  $S \subseteq \mathbb{R}^n$  be measurable, bounded, convex, and of positive volume. Let  $c = \text{center}(S)$ , and let  $H$  be a closed halfspace whose boundary contains  $c$ . Then

$$\text{vol}(S \cap H) \geq \left( \frac{n}{n+1} \right)^n \text{vol}(S).$$

The sequence  $(n/(n+1))^n$  is strictly decreasing in  $n$  and tends to  $1/e$ . Consequently,

$$\text{vol}(S \cap H) \geq \frac{1}{e} \text{vol}(S) \quad \forall n \geq 1,$$

and this dimension-free lower bound is what we will use in the complexity analysis below. This is the classical Grünbaum inequality; see [Grü60] for the original source and [MSZ18, Theorem 1] for a modern generalization from which this statement follows as a special case.

**Theorem 6.7** (Center-of-mass search complexity). Consider the generic cutting-plane method with exact updates

$$S_{t+1} = S_t \cap H_t,$$

and choose

$$x_t = \text{center}(S_t)$$

at every iteration. Assume that each valid cut  $H_t$  has boundary through  $x_t$ . Then

$$\text{vol}(S_T) \leq (1 - e^{-1})^T \text{vol}(S_0).$$

Consequently, if  $K \subseteq S_0$  is any convex target set satisfying

$$\text{vol}(K) \geq \eta,$$

then

$$\lceil 3 \log(\text{vol}(S_0)/\eta) \rceil + 1$$

separation-oracle queries suffice to find a point in  $K$ .

*Proof idea of Theorem 6.7.* At each step, the cut passes through  $x_t = \text{center}(S_t)$ . By Theorem 6.6, the kept half  $S_t \cap H_t$  contains at most a  $(1 - e^{-1})$ -fraction of the current volume. Iterating gives

$$\text{vol}(S_T) \leq (1 - e^{-1})^T \text{vol}(S_0).$$

If the algorithm had not yet queried a point in  $K$ , then all cuts would still be valid for  $K$ , so

Proposition 6.3 would force

$$\text{vol}(S_T) \geq \text{vol}(K) \geq \eta.$$

Thus once  $\text{vol}(S_T) < \eta$ , some earlier query must already have belonged to  $K$ . Since  $\log((1 - e^{-1})^{-1}) > 1/3$ , it is enough to take

$$T > 3 \log \frac{\text{vol}(S_0)}{\eta}.$$

**Corollary 6.8** (Center-of-mass optimization complexity). *Assume  $0 < \delta \leq \Delta$ . Run the center-of-mass method on the optimization problem  $\min_{x \in X} f(x)$ , using the separation oracle for  $K_\delta$  from Proposition 6.4. Then after*

$$\lfloor 3n \log(\Delta/\delta) \rfloor + 1$$

*iterations, and hence at most that many first-order oracle calls, one of the queried points is  $\delta$ -optimal.*

*Proof of Corollary 6.8.* Choose

$$\alpha := \frac{\delta}{\Delta}.$$

By Proposition 6.5,

$$\text{vol}(K_\delta) \geq \alpha^n \text{vol}(X).$$

Applying Theorem 6.7 with  $K = K_\delta$ ,  $S_0 = X$ , and  $\eta = \alpha^n \text{vol}(X)$  yields

$$\text{vol}(S_T) \leq (1 - e^{-1})^T \text{vol}(X).$$

Thus  $\text{vol}(S_T) < \alpha^n \text{vol}(X)$  as soon as

$$(1 - e^{-1})^T < \alpha^n,$$

or equivalently,

$$T > \frac{n \log(1/\alpha)}{\log((1 - e^{-1})^{-1})} = \frac{n \log(\Delta/\delta)}{\log((1 - e^{-1})^{-1})}.$$

Since

$$\log((1 - e^{-1})^{-1}) > \frac{1}{3},$$

it is enough to take

$$T > 3n \log(\Delta/\delta).$$

Therefore the stated integer bound on  $T$  suffices. Any point in  $K_\delta$  is  $\delta$ -optimal by definition of  $K_\delta$ .  $\square$

*Remark 6.1.* The center-of-mass method is geometrically natural but computationally demanding. At every step it requires a sufficiently explicit representation of the current search region to compute its center of mass. The ellipsoid method replaces the true search region by a tractable outer approximation whose center and update are explicit.

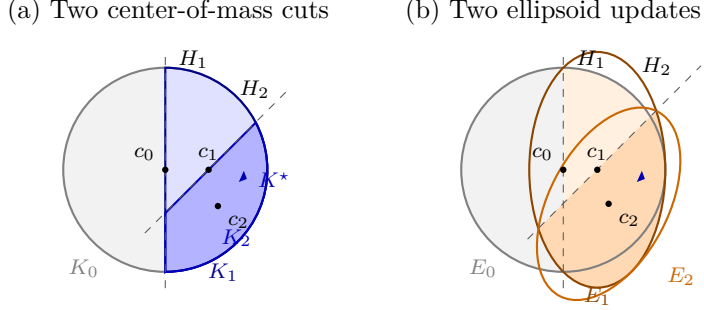


Figure 3: Both panels start from the same unit disk and keep the side containing the same small target set. In (a), the exact localization sets  $K_0 \supset K_1 \supset K_2$  are overlaid in one coordinate system: after two center-of-mass cuts, the shape is visibly no longer ellipsoidal. In (b), the ellipsoid method uses the same two cut directions, but after each cut it replaces the kept region by an explicit covering ellipsoid, so the overlaid iterates  $E_0, E_1, E_2$  remain ellipsoids throughout.

## 6.5 Ellipsoids as Tractable Outer Approximations

The center-of-mass method keeps the exact localization set  $K_t$ , but after only a few cuts that set becomes too irregular to represent explicitly. The ellipsoid method trades exactness for tractability: instead of keeping  $K_t$  itself, it keeps an explicit outer ellipsoid  $E_t \supseteq K_t$ . The price is that the one-step shrink is no longer the dimension-free constant of [Theorem 6.7](#); instead, the update below guarantees only the weaker factor  $e^{-1/(4n)}$ , which is precisely where the extra factor of  $n$  in the final complexity comes from.

**Definition 6.6** (Ellipsoid). Let  $Q \in \mathbb{R}^{n \times n}$  be symmetric positive definite and  $c \in \mathbb{R}^n$ . The ellipsoid with center  $c$  and shape matrix  $Q$  is

$$E(c, Q) := \left\{ x \in \mathbb{R}^n : (x - c)^\top Q^{-1} (x - c) \leq 1 \right\}.$$

**Theorem 6.9** (A covering ellipsoid for a central cut). Assume  $n \geq 2$ . Let  $E(c, Q)$  be an ellipsoid in  $\mathbb{R}^n$ . Let  $a \in \mathbb{R}^n \setminus \{0\}$ , define the central halfspace

$$H := \{x \in \mathbb{R}^n : \langle a, x \rangle \leq \langle a, c \rangle\},$$

Define

$$c^+ = c - \frac{1}{n+1} \frac{Qa}{\sqrt{a^\top Q a}}, \quad Q^+ = \frac{n^2}{n^2-1} \left( Q - \frac{2}{n+1} \frac{Qaa^\top Q}{a^\top Q a} \right).$$

Then

$$E(c, Q) \cap H \subseteq E(c^+, Q^+) \quad \text{and} \quad \text{vol}(E(c^+, Q^+)) \leq e^{-1/(4n)} \text{vol}(E(c, Q)).$$

*Proof idea.* The statement is affine-invariant. If we write

$$x = c + Q^{1/2}z,$$

then  $E(c, Q)$  is sent to the unit ball  $B = \{z : \|z\|_2 \leq 1\}$ , the cut becomes

$$\{z : \langle \nu, z \rangle \leq 0\}, \quad \nu := \frac{Q^{1/2}a}{\sqrt{a^\top Q a}},$$

and both volumes are multiplied by the same Jacobian factor  $\det(Q^{1/2})$ . Thus it suffices to analyze the normalized case  $c = 0$ ,  $Q = I$ , and after a rotation we may further assume  $\nu = e_1$ . In that case the covering ellipsoid is centered at  $-e_1/(n+1)$ , with semiaxis lengths  $n/(n+1)$  in the  $e_1$ -direction and  $n/\sqrt{n^2-1}$  in the orthogonal directions. Therefore its volume ratio relative to the unit ball is

$$\frac{n}{n+1} \left( \frac{n}{\sqrt{n^2-1}} \right)^{n-1} = \left( \frac{n^2}{n^2-1} \right)^{n/2} \sqrt{\frac{n-1}{n+1}},$$

which is the factor computed in the proof below.

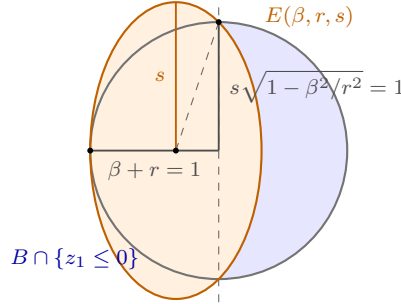


Figure 4: The normalized two-dimensional case, drawn using the notation  $E(\beta, r, s)$ . The left contact point gives the relation  $\beta + r = 1$ , while the vertical section at  $z_1 = 0$  has half-length  $s\sqrt{1 - \beta^2/r^2}$ . To cover the half-disk, this vertical half-length must be at least 1, and the minimum-area solution occurs when it is exactly 1.

To see why these parameters are the optimal ones, work in the normalized case and use the rotational symmetry around the  $e_1$ -axis. The minimum-volume covering ellipsoid must therefore have the form

$$E(\beta, r, s) := \left\{ z \in \mathbb{R}^n : \frac{(z_1 + \beta)^2}{r^2} + \frac{\|z_\perp\|_2^2}{s^2} \leq 1 \right\},$$

where  $z = (z_1, z_\perp)$ ,  $\beta \geq 0$ ,  $r > 0$ , and  $s > 0$ . The leftmost point of  $E(\beta, r, s)$  is  $-(\beta + r)e_1$ , so covering the point  $-e_1$  forces

$$\beta + r \geq 1.$$

Also, the section of  $E(\beta, r, s)$  by the hyperplane  $z_1 = 0$  is the Euclidean ball

$$\left\{ z_\perp \in \mathbb{R}^{n-1} : \|z_\perp\|_2 \leq s\sqrt{1 - \beta^2/r^2} \right\},$$

so covering the equatorial disk  $\{z_1 = 0, \|z_\perp\|_2 \leq 1\}$  forces

$$s\sqrt{1 - \beta^2/r^2} \geq 1.$$

At the minimum-volume solution both inequalities must be tight, since otherwise one could decrease  $r$  or  $s$  and obtain a smaller covering ellipsoid. Therefore

$$\beta + r = 1, \quad s\sqrt{1 - \frac{\beta^2}{r^2}} = 1,$$

so

$$r = 1 - \beta, \quad s^2 = \frac{(1 - \beta)^2}{1 - 2\beta}, \quad 0 \leq \beta < \frac{1}{2}.$$

Therefore it remains to minimize the exact volume ratio

$$\frac{\text{vol}(E(\beta, r, s))}{\text{vol}(B)} = r s^{n-1} = (1 - \beta) \left( \frac{1 - \beta}{\sqrt{1 - 2\beta}} \right)^{n-1} = (1 - \beta)^n (1 - 2\beta)^{-(n-1)/2},$$

where  $B = \{z \in \mathbb{R}^n : \|z\|_2 \leq 1\}$  is the unit ball. Taking logarithms,

$$\frac{d}{d\beta} \log(r s^{n-1}) = -\frac{n}{1 - \beta} + \frac{n-1}{1 - 2\beta} = \frac{(n+1)\beta - 1}{(1 - \beta)(1 - 2\beta)}.$$

Thus the minimizer is

$$\beta = \frac{1}{n+1}, \quad r = \frac{n}{n+1}, \quad s = \frac{n}{\sqrt{n^2 - 1}},$$

which is exactly the choice used above.

**Theorem 6.10** (Ellipsoid search complexity). *Assume  $n \geq 2$ . Let  $E_0 \subseteq \mathbb{R}^n$  be an ellipsoid and let  $K \subseteq E_0$  be a convex target set. At iteration  $t$ , query the center  $c_t$  of the current ellipsoid  $E_t$ . If the separation oracle does not certify that  $c_t \in K$ , let  $H_t$  be a valid cut for  $K$  whose boundary passes through  $c_t$ , and let  $E_{t+1}$  be the ellipsoid obtained by applying [Theorem 6.9](#) to  $E_t$  and  $H_t$ . Then*

$$\text{vol}(E_T) \leq e^{-T/(4n)} \text{vol}(E_0),$$

Consequently, if

$$\text{vol}(K) \geq \eta,$$

then

$$\left\lceil 4n \log \frac{\text{vol}(E_0)}{\eta} \right\rceil + 1$$

separation-oracle queries suffice to find a point in  $K$ .

*Proof of [Theorem 6.10](#).* As long as the algorithm has not yet queried a point in  $K$ , each returned cut  $H_t$  is valid for  $K$ , so

$$K \subseteq E_t \cap H_t \subseteq E_{t+1}.$$

Applying [Theorem 6.9](#) at each iteration yields

$$\text{vol}(E_T) \leq e^{-T/(4n)} \text{vol}(E_0).$$

If the algorithm had not yet queried a point in  $K$ , then

$$K \subseteq E_T, \quad \text{hence} \quad \text{vol}(E_T) \geq \text{vol}(K) \geq \eta.$$

Therefore the algorithm must have found a point of  $K$  once

$$e^{-T/(4n)} \text{vol}(E_0) < \eta,$$

which is equivalent to

$$T > 4n \log \frac{\text{vol}(E_0)}{\eta}.$$

Therefore the stated integer bound on  $T$  suffices.  $\square$

**Corollary 6.11** (Ellipsoid optimization complexity). *Assume  $0 < \delta \leq \Delta$ . Assume  $X \subseteq E_0$  for some ellipsoid  $E_0$ . Run the ellipsoid method on  $\min_{x \in X} f(x)$ , using the separation oracle for  $K_\delta$  from Proposition 6.4. Then after*

$$\left\lceil 4n \log \frac{\text{vol}(E_0)}{\text{vol}(X)} + 4n^2 \log \frac{\Delta}{\delta} \right\rceil + 1$$

*iterations, and hence at most that many first-order oracle calls, one of the queried points is  $\delta$ -optimal.*

*Proof of Corollary 6.11.* Choose

$$\alpha := \frac{\delta}{\Delta}.$$

By Proposition 6.5,

$$\text{vol}(K_\delta) \geq \alpha^n \text{vol}(X).$$

Applying Theorem 6.10 with  $K = K_\delta$  and  $\eta = \alpha^n \text{vol}(X)$  gives

$$\begin{aligned} T &= \left\lceil 4n \log \frac{\text{vol}(E_0)}{\alpha^n \text{vol}(X)} \right\rceil + 1 \\ &= \left\lceil 4n \log \frac{\text{vol}(E_0)}{\text{vol}(X)} + 4n^2 \log \frac{1}{\alpha} \right\rceil + 1 \\ &= \left\lceil 4n \log \frac{\text{vol}(E_0)}{\text{vol}(X)} + 4n^2 \log \frac{\Delta}{\delta} \right\rceil + 1. \end{aligned}$$

Any point in  $K_\delta$  is  $\delta$ -optimal by definition of  $K_\delta$ .  $\square$

*Proof of Theorem 6.9.* Let

$$\alpha := \sqrt{a^\top Q a} \quad \text{and} \quad \nu := \frac{Q^{1/2} a}{\alpha}.$$

Then  $\|\nu\|_2 = 1$ . Apply the affine change of variables

$$x = c + Q^{1/2} z.$$

In the  $z$ -coordinates, the ellipsoid  $E(c, Q)$  becomes the Euclidean unit ball

$$B := \{z \in \mathbb{R}^n : \|z\|_2 \leq 1\},$$

and the central cut becomes

$$B \cap \{z \in \mathbb{R}^n : \langle \nu, z \rangle \leq 0\}.$$

The proposed covering ellipsoid becomes

$$E_z := \left\{ z \in \mathbb{R}^n : (z + \nu/(n+1))^\top M^{-1} (z + \nu/(n+1)) \leq 1 \right\},$$

where

$$M := \frac{n^2}{n^2 - 1} \left( I - \frac{2}{n+1} \nu \nu^\top \right).$$

Fix  $z = s\nu + w \in B$  with  $s \leq 0$  and  $w \perp \nu$ . Since  $\|z\|_2 \leq 1$ , one has  $s^2 + \|w\|_2^2 \leq 1$ . The matrix  $M$  has eigenvalue  $n^2/(n+1)^2$  along  $\nu$  and eigenvalue  $n^2/(n^2-1)$  on  $\nu^\perp$ . Therefore

$$\begin{aligned} (z + \nu/(n+1))^\top M^{-1} (z + \nu/(n+1)) &= \frac{(n+1)^2}{n^2} \left( s + \frac{1}{n+1} \right)^2 + \frac{n^2-1}{n^2} \|w\|_2^2 \\ &\leq \frac{(n+1)^2}{n^2} \left( s + \frac{1}{n+1} \right)^2 + \frac{n^2-1}{n^2} (1-s^2) \\ &= 1 + \frac{2(n+1)}{n^2} s(s+1) \leq 1, \end{aligned}$$

because  $-1 \leq s \leq 0$ . Thus the cut ball is contained in  $E_z$ , hence

$$E(c, Q) \cap H \subseteq E(c^+, Q^+).$$

For the volume estimate, note that

$$Q^+ = Q^{1/2} M Q^{1/2}, \quad \det(Q^+) = \det(M) \det(Q).$$

Since  $M$  has one eigenvalue  $n^2/(n+1)^2$  and  $n-1$  eigenvalues  $n^2/(n^2-1)$ ,

$$\det(M) = \left( \frac{n^2}{n^2-1} \right)^n \frac{n-1}{n+1}.$$

Hence

$$\frac{\text{vol}(E(c^+, Q^+))}{\text{vol}(E(c, Q))} = \sqrt{\det(M)} = \left( \frac{n^2}{n^2-1} \right)^{n/2} \sqrt{\frac{n-1}{n+1}}.$$

A direct logarithmic estimate shows that this factor is at most  $e^{-1/(4n)}$  for all  $n \geq 2$ .  $\square$

*Remark 6.2* (Historical note). Cutting-plane ideas for convex programming long predate the ellipsoid method; a classical early example is Kelley's cutting-plane method [Kel60]. The ellipsoid method itself emerged from the 1970s Soviet line of work on nonsmooth convex optimization and became widely influential after Khachiyan's polynomial-time algorithm for linear programming [Kha79]. Its role in combinatorial optimization was then crystallized by Grötschel, Lovász, and Schrijver through the separation–optimization equivalence and related applications [GLS88]. Here the key point is not that the ellipsoid method somehow resolves integrality on its own; rather, certain discrete problems admit exact convex formulations, often integral polytopes or exact convex extensions, together with efficient separation oracles. For the broader oracle and information-based complexity viewpoint, see [NY83].

## Exercises

1. Show that the one-dimensional method above can be viewed as a cutting-plane method for the unknown target interval

$$\arg \min_{x \in [a, b]} f(x).$$

2. Suppose now that  $f^*$  is not known and that only a subgradient oracle for  $f$  is available. Show that the construction in [Proposition 6.4](#) no longer gives a genuine separation oracle for  $K_\delta$ , but still yields the following offline guarantee: if none of the first  $T$  query points is  $\delta$ -optimal, then every cut produced in those  $T$  steps is valid for  $K_\delta$ .
3. Compare the complexity statements of [Corollaries 6.8](#) and [6.11](#). Why does the ellipsoid method lose an extra factor of  $n$ ?

## References

- [GLS88] Martin Grötschel, László Lovász, and Alexander Schrijver. The ellipsoid method. In *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics*, pages 64–101. Springer, Berlin, Heidelberg, 1988.
- [Grü60] Branko Grünbaum. Partitions of mass-distributions and of convex bodies by hyperplanes. *Pacific Journal of Mathematics*, 10(4):1257–1261, 1960.
- [Kel60] J. E. Kelley. The cutting-plane method for solving convex programs. *Journal of the Society for Industrial and Applied Mathematics*, 8(4):703–712, 1960.
- [Kha79] L. G. Khachiyan. A polynomial algorithm in linear programming. *Soviet Mathematics Doklady*, 20(1):191–194, 1979.
- [MSZ18] Sergiy Myroshnychenko, Matthew Stephen, and Nicole Zhang. Grünbaum’s inequality for sections. *Journal of Functional Analysis*, 275(9):2516–2537, 2018.
- [NY83] Arkadi S. Nemirovski and David B. Yudin. *Problem Complexity and Method Efficiency in Optimization*. Wiley-Interscience, New York, 1983.