Convex Optimization and Applications

5 - Ellipsoid Methods

Guillaume Sagnol

Technische Universität Berlin

Outline

History

Ellipsoid Method

- Introduced in the 70's by Shor, and Yudin & Nemirovski
- Modifications by Khachian (1979), so it can solve LPs in polynomial time, i.e., an algorithm that finds an optimal solution of $\min\{\boldsymbol{c}^T\boldsymbol{x} : A\boldsymbol{x} \leq \boldsymbol{b}\}$ in time polynomial w.r.t. bit-size of $(A, \boldsymbol{b}, \boldsymbol{c})$
- Essential contributions of Grötschel, Lovász and Schrijver (1981):
  - Weak separation + finite-precision arithmetics
  - Applications to combinatorial optimization
- Not a practical method, but formidable tool:

$$\text{“separation of } C\text{”} \iff \text{“optimization over } C\text{”}$$

Warm-up: Bisection method

Consider the one-dimensional minimization problem for a convex function $f : [\ell_0, u_0] \to \mathbb{R}$:

$$\underset{x \in [\ell_0, u_0]}{\textbf{minimize}} \quad f(x).$$

At iteration $k \geq 1$:

- $x_k \leftarrow \frac{1}{2}(\ell_{k-1} + u_{k-1})$
- Evaluate $f'(x_k)$
- If $f'(x_k) \leq 0$:
$$(\ell_k, u_k) \leftarrow (x_k, u_{k-1})$$
- Else:
$$(\ell_k, u_k) \leftarrow (\ell_{k-1}, x_k)$$

Interval is halved at each iteration $\to$ fast convergence.

From Bisection to Ellipsoid method

Bisection method

- 1-dimensional problems
- Intervals $I_k \supseteq$ optimal set
- Evaluate $f'(x_k)$
- $\text{len}(I_k) = \frac{1}{2}\,\text{len}(I_{k-1})$

Ellipsoid method

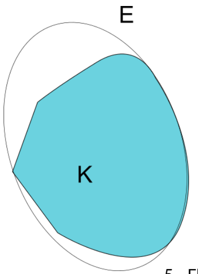
- n-dimensional problems
- Ellipsoids $E_k \supseteq$ optimal set
- Separation oracle
- $\text{vol}(E_k) \leq \alpha\,\text{vol}(E_{k-1})$, for some $\alpha < 1$.

Löwner-John Ellipsoid

Theorem (John, 1948).

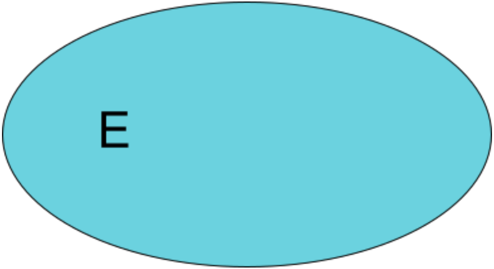
Every convex body $K \subset \mathbb{R}^n$ (i.e., compact convex, non-empty interior) is contained in a unique ellipsoid E of minimal volume, called the Löwner-John ellipsoid of K.

Moreover, the ellipsoid obtained by shrinking E by a factor $\frac{1}{n}$ around its center is contained in K



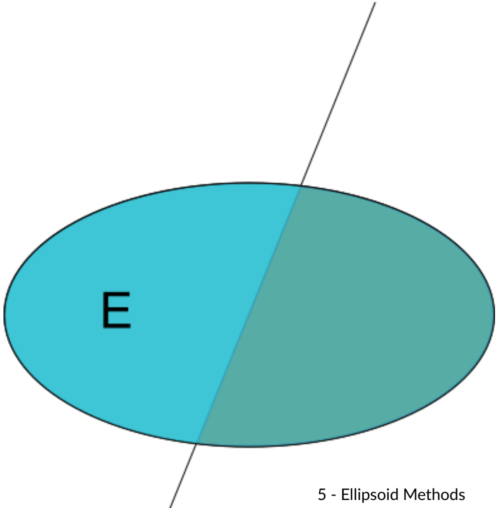
L-J ellipsoid of a half-ellipsoid

In the ellipsoid method, the operation corresponding to "halving intervals" is to "take the Löwner-John Ellipsoid of a half-ellipsoid"



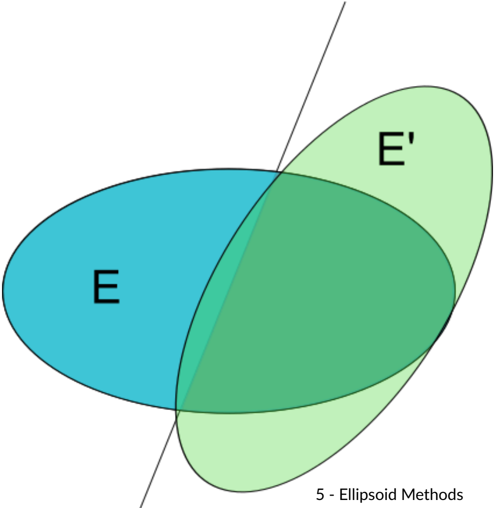
L-J ellipsoid of a half-ellipsoid

In the ellipsoid method, the operation corresponding to "halving intervals" is to "take the Löwner-John Ellipsoid of a half-ellipsoid"



L-J ellipsoid of a half-ellipsoid

In the ellipsoid method, the operation corresponding to "halving intervals" is to "take the Löwner-John Ellipsoid of a half-ellipsoid"



L-J ellipsoid of a half-ellipsoid

This can be done efficiently !

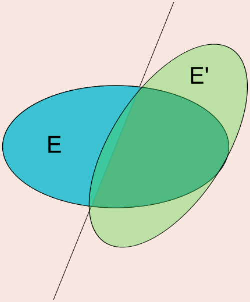
Proposition (L-J Ellipsoid of a half-ellipsoid).

Let $E = E(\boldsymbol{a}, Q) := \{\boldsymbol{x} \in \mathbb{R}^n : (\boldsymbol{x} - \boldsymbol{a})^T Q^{-1}(\boldsymbol{x} - \boldsymbol{a}) \leq 1\}$,
$H = E \cap \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{h}^T \boldsymbol{x} \leq \boldsymbol{h}^T \boldsymbol{a}\}, \quad \boldsymbol{b} := \frac{1}{\sqrt{\boldsymbol{h}^T Q \boldsymbol{h}}} Q\boldsymbol{h}.$

Then, the L-J ellipsoid of H is $E' = E(\boldsymbol{a}', Q')$, where

$$\boldsymbol{a}' := \boldsymbol{a} - \frac{1}{(n+1)}\boldsymbol{b}$$

$$Q' := \frac{n^2}{n^2-1}\left(Q - \frac{2}{n+1}\boldsymbol{b}\boldsymbol{b}^T\right)$$



Volume reduction

The volume of the Löwner-John ellipsoid of a half-ellipsoid is within a constant fraction of the original volume:

Lemma

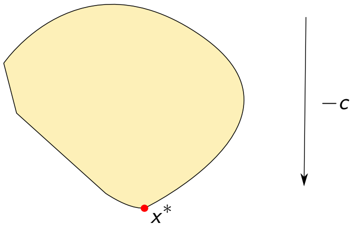
Let $E' = E(\boldsymbol{a}', Q')$ be the Löwner-John ellipsoid of $E(\boldsymbol{a}, Q) \cap \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{h}^T \boldsymbol{x} \leq \boldsymbol{h}^T \boldsymbol{a}\}$. Then,

$$\text{volume}(E') < e^{-\frac{1}{2(n+1)}} \, \text{volume}(E).$$

Separation Oracle

Framework:

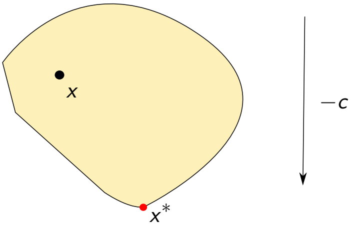
- Minimize a linear function $f(\boldsymbol{x}) = \langle \boldsymbol{c}, \boldsymbol{x} \rangle$ over a convex body $K \in \mathbb{R}^n$.
- The feasible set K is not given by constraints, but instead we assume that a *separation oracle* is available.



Separation Oracle

Framework:

- Minimize a linear function $f(\boldsymbol{x}) = \langle \boldsymbol{c}, \boldsymbol{x} \rangle$ over a convex body $K \in \mathbb{R}^n$.
- The feasible set K is not given by constraints, but instead we assume that a *separation oracle* is available.



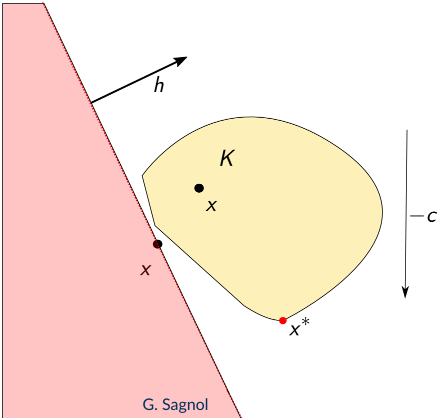
Case 1: $\boldsymbol{x} \in K$

Oracle returns "yes"

Separation Oracle

Framework:

- Minimize a linear function $f(\boldsymbol{x}) = \langle \boldsymbol{c}, \boldsymbol{x} \rangle$ over a convex body $K \in \mathbb{R}^n$.
- The feasible set K is not given by constraints, but instead we assume that a *separation oracle* is available.



Case 2: $\boldsymbol{x} \notin K$

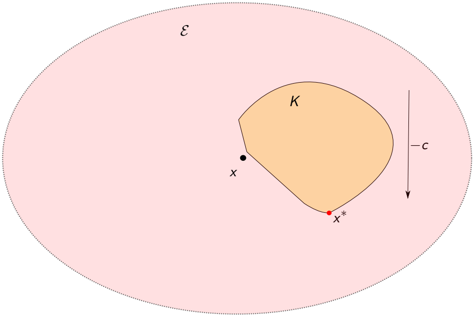
Oracle returns separating hyperplane $\boldsymbol{h}$:

- $\langle \boldsymbol{h}, \boldsymbol{x} \rangle < \langle \boldsymbol{h}, \boldsymbol{z} \rangle, \forall \boldsymbol{z} \in K$

The Ellipsoid method: Description

▷ Start with large ellipsoid $\mathcal{E}$ that contains K. Its center is $\boldsymbol{x}$.

▷ Repeat until convergence:

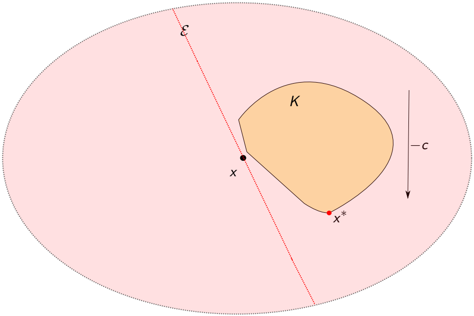


The Ellipsoid method: Description

▷ Start with large ellipsoid $\mathcal{E}$ that contains K. Its center is $\boldsymbol{x}$.

▷ Repeat until convergence:

1. Query Separation Oracle at $\boldsymbol{x}$.



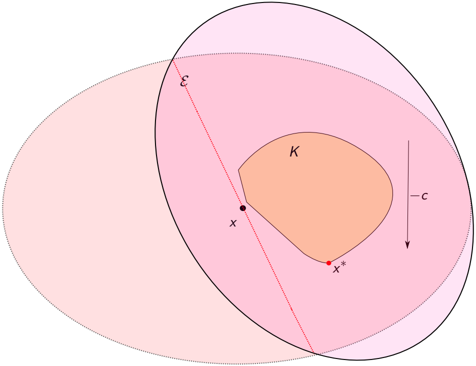
The Ellipsoid method: Description

▷ Start with large ellipsoid $\mathcal{E}$ that contains K. Its center is $\boldsymbol{x}$.

▷ Repeat until convergence:

1. Query Separation Oracle at $\boldsymbol{x}$.
2. If $\boldsymbol{x} \notin K$, we get a halfspace H s.t. $K \subset H$.

   Compute min. volume ellipsoid that contains the half-ellipsoid $H \cap \mathcal{E}$.



The Ellipsoid method: Description

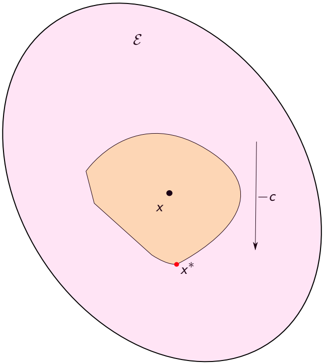
▷ Start with large ellipsoid $\mathcal{E}$ that contains K. Its center is $\boldsymbol{x}$.

▷ Repeat until convergence:

1. Query Separation Oracle at $\boldsymbol{x}$.
2. If $\boldsymbol{x} \notin K$, we get a halfspace H s.t. $K \subset H$.

   Compute min. volume ellipsoid that contains the half-ellipsoid $H \cap \mathcal{E}$.

   Update $\mathcal{E}$ and $\boldsymbol{x}$.



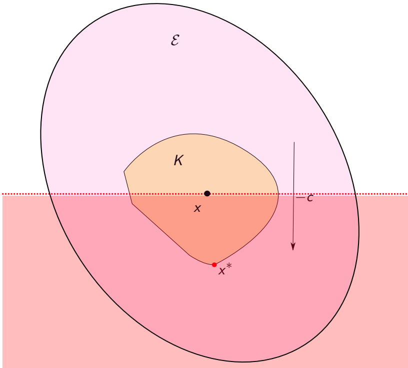
The Ellipsoid method: Description

- ▷ Start with large ellipsoid $\mathcal{E}$ that contains K. Its center is $\boldsymbol{x}$.
- ▷ Repeat until convergence:

1. Query Separation Oracle at $\boldsymbol{x}$.
2. If $\boldsymbol{x} \notin K$, we get a halfspace H s.t. $K \subset H$.

   Compute min. volume ellipsoid that contains the half-ellipsoid $H \cap \mathcal{E}$.

   Update $\mathcal{E}$ and $\boldsymbol{x}$.
3. Otherwise, define $H = \{\boldsymbol{z} : \langle \boldsymbol{c}, \boldsymbol{z} \rangle \leq \langle \boldsymbol{c}, \boldsymbol{x} \rangle\}$,



The Ellipsoid method: Description

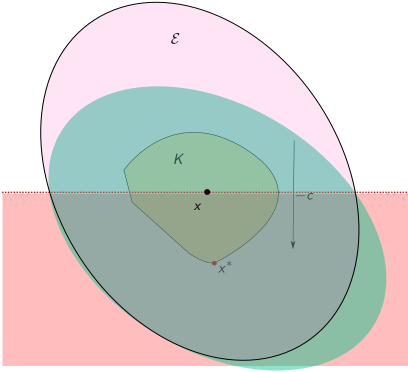
▷ Start with large ellipsoid $\mathcal{E}$ that contains K. Its center is $\boldsymbol{x}$.

▷ Repeat until convergence:

1. Query Separation Oracle at $\boldsymbol{x}$.
2. If $\boldsymbol{x} \notin K$, we get a halfspace H s.t. $K \subset H$.

   Compute min. volume ellipsoid that contains the half-ellipsoid $H \cap \mathcal{E}$.

   Update $\mathcal{E}$ and $\boldsymbol{x}$.
3. Otherwise, define $H = \{\boldsymbol{z} : \langle \boldsymbol{c}, \boldsymbol{z} \rangle \leq \langle \boldsymbol{c}, \boldsymbol{x} \rangle\}$, and proceed as above.



The Ellipsoid method: Description

▷ Start with large ellipsoid $\mathcal{E}$ that contains K. Its center is $\boldsymbol{x}$.

▷ Repeat until convergence:

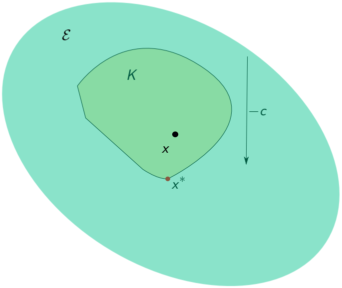
1. Query Separation Oracle at $\boldsymbol{x}$.

2. If $\boldsymbol{x} \notin K$, we get a halfspace H s.t. $K \subset H$.

   Compute min. volume ellipsoid that contains the half-ellipsoid $H \cap \mathcal{E}$.

   Update $\mathcal{E}$ and $\boldsymbol{x}$.

3. Otherwise, define $H = \{\boldsymbol{z} : \langle \boldsymbol{c}, \boldsymbol{z} \rangle \leq \langle \boldsymbol{c}, \boldsymbol{x} \rangle\}$, and proceed as above.



Analysis for feasibility problems

Assumptions: We are given $R, r > 0$ such that

(i) $K \subseteq B(\mathbf{0}, R)$

(ii) either $K = \emptyset$, or $\exists \boldsymbol{x} \in K : \supseteq B(\boldsymbol{x}, r) \subseteq K$.

Under (i) and (ii), we can solve the feasibility problem (find $\boldsymbol{x} \in K$, or assert that $K = \emptyset$) by calling the separation oracle $O(n^2 \log(R/r))$ times.

Theorem

If, after $N = \lfloor 2n(n+1) \log(R/r) \rfloor$ iterations, the ellipsoid algorithm didn't find a point $\boldsymbol{x} \in K$, then K is empty.

Analysis for optimization problems

$$p^* = \inf_{\mathbf{x} \in K} \mathbf{c}^T \mathbf{x}, \qquad (P)$$

- The solution of (P) can be irrational. Hence, we search $\epsilon-$suboptimal solutions.
- If we can solve feasibility problems, then we can solve the optimization problem to arbitrary precision, by binary search: Find the largest δ such that $\{\mathbf{x} \in K : \mathbf{c}^T \mathbf{x} \leq \delta\} \neq \emptyset$.
- But as δ approaches p^*, the $(\delta - p^*)$-suboptimal set becomes very small: will assumption (ii) still hold?

Analysis for optimization problems

Under (i) and (ii), the ϵ-suboptimal set cannot be too small:

Proposition

Let K be a convex body satifying (i) and (ii) for $r, R > 0$, and let $0 < \epsilon < R$. Then, either K is empty, or the ϵ-suboptimal set for (P) contains a ball of radius $\frac{r\epsilon}{2R+r}$.

Analysis for optimization problems

Under (i) and (ii), the ϵ-suboptimal set cannot be too small:

Proposition

Let K be a convex body satifying (i) and (ii) for $r, R > 0$, and let $0 < \epsilon < R$. Then, either K is empty, or the ϵ-suboptimal set for (P) contains a ball of radius $\dfrac{r\epsilon}{2R + r}$.

This allows us to show the following result:

Theorem

If constants R and r are known such that K satisfies (i)-(ii), then we can find an $\epsilon-$suboptimal solution of (P), or assert that this problem is infeasible, by making $O\left(n^2 \log \dfrac{R}{\min(r, \epsilon)}\right)$ calls to the separation oracle.

Weak separation & Optimization

Exact separators: not realistic using finite-precision arithmetics. Moreover, there is a square-root in the formula for the L-J ellipsoid of a half-ellipsoid, which must be approximated.

Definition

Let $K \subset \mathbb{R}^n$.

- We say that $\boldsymbol{x}$ is ϵ-*almost* in K, and we write $\boldsymbol{x} \in K^{+\epsilon}$, if $\exists \boldsymbol{z} \in K, \|\boldsymbol{x} - \boldsymbol{z}\| \leq \epsilon$.
- We say that $\boldsymbol{x}$ is ϵ-*deep* in K, and we write $\boldsymbol{x} \in K^{-\epsilon}$, if $B(\boldsymbol{x}, \epsilon) \subseteq K$.

Weak separation & Optimization

For weak optimization and separation problems, it is sufficient to distinguish between points that are almost/deep in K:

Definition (Weak optimization).

Given $K, \boldsymbol{c}, \epsilon$, either

- Return $\boldsymbol{x}^* \in K^{+\epsilon}$ such that $\boldsymbol{c}^T\boldsymbol{x}^* \leq \boldsymbol{c}^T\boldsymbol{y} + \epsilon,\ \forall \boldsymbol{y} \in K^{-\epsilon}$;
- or assert that $K^{-\epsilon}$ is empty.

Weak separation & Optimization

For weak optimization and separation problems, it is sufficient to distinguish between points that are almost/deep in K:

Definition (Weak optimization).

Given $K, \boldsymbol{c}, \epsilon$, either

- Return $\boldsymbol{x}^* \in K^{+\epsilon}$ such that $\boldsymbol{c}^T\boldsymbol{x}^* \leq \boldsymbol{c}^T\boldsymbol{y} + \epsilon,\ \forall \boldsymbol{y} \in K^{-\epsilon}$;
- or assert that $K^{-\epsilon}$ is empty.

Definition (Weak separation).

Given $K, \boldsymbol{x}, \epsilon$, either

- Assert that $\boldsymbol{x} \in K^{+\epsilon}$;
- or return $\boldsymbol{h}$ with $\|\boldsymbol{h}\|_\infty = 1$ such that $\boldsymbol{h}^T\boldsymbol{x}^* \leq \boldsymbol{h}^T\boldsymbol{y} + \epsilon,\ \forall \boldsymbol{y} \in K^{-\epsilon}$

Grötschel, Lovász & Schrijver's theorems

Theorem

Given R and a polynomial-time weak separation oracle for $K \subseteq B(0, R)$, we can solve the weak optimization problem in polynomial time.

Grötschel, Lovász & Schrijver's theorems

Theorem

Given R and a polynomial-time weak separation oracle for $K \subseteq B(0, R)$, we can solve the weak optimization problem in polynomial time.

Moreover, we have a converse, so that weak separation and weak optimization are essentially equivalent:

Theorem

Given a polynomial-time weak optimization oracle for a convex set K, we can solve the weakly separation problem for K in polynomial time.