Convex Optimization and Applications
5 - Ellipsoid Methods

Guillaume Sagnol
Outline

1. Introduction
2. Halfving Ellipsoids
3. Feasibility Problems
4. Convex Optimization Problems
5. Weak separation & optimization
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 \( \min \{ c^T x : Ax \leq b \} \) in time polynomial w.r.t. bit-size of \((A, b, c)\)
- Essential contributions of Grötschel, Lovász and Schrijver (1981):
  - Weak separation + finite-precision arithmetics
  - Applications to combinatorial optimization
- Not a practical method, but formidable tool:

  “separation of \( C \)” \(\iff\) “optimization over \( C \)”
Warm-up: Bisection method

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

\[
\begin{align*}
\text{minimize} & \quad f(x). \\
\text{subject to} & \quad x \in [\ell_0, u_0]
\end{align*}
\]

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 $\rightarrow$ 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 \( E = E(a, Q) := \{ x \in \mathbb{R}^n : (x - a)^T Q^{-1} (x - a) \leq 1 \} \),

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

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

\[
    a' := a - \frac{1}{(n + 1)} b \\
    Q' := \frac{n^2}{n^2 - 1} \left( Q - \frac{2}{n + 1} b b^T \right)
\]
The volume of the Löwner-John ellipsoid of a half-ellipsoid is within a constant fraction of the original volume:

**Lemma**

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

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

Framework:

- Minimize a linear function $f(x) = \langle c, x \rangle$ 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 $f(x) = \langle c, x \rangle$ 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: $x \in K$
Oracle returns “yes”
Separation Oracle

Framework:

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

Case 2: \( x \not\in K \)

Oracle returns separating hyperplane \( h \):

- \( \langle h, x \rangle < \langle h, z \rangle, \forall z \in K \)
The Ellipsoid method: Description

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

1. Query Separation Oracle at $x$.
2. If $x \not\in K$, 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 $x$.
3. Otherwise, define $H = \{ z : \langle c, z \rangle \leq \langle c, x \rangle \}$, and proceed as above.

$x^*$
The Ellipsoid method: Description

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

1. Query Separation Oracle at $x$.
The Ellipsoid method: Description

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

▷ Repeat until convergence:

1. Query Separation Oracle at $x$.

2. If $x \not\in K$, 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 \( x \).
- Repeat until convergence:
  
  1. Query Separation Oracle at \( x \).
  2. If \( x \notin K \), 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 \( x \).
The Ellipsoid method: Description

▷ Start with large ellipsoid $E$ that contains $K$. Its center is $x$.
▷ Repeat until convergence:

1. Query Separation Oracle at $x$.
2. If $x \notin K$, we get a halfspace $H$ s.t. $K \subset H$.
   Compute min. volume ellipsoid that contains the half-ellipsoid $H \cap E$.
   Update $E$ and $x$.
3. Otherwise, define
   $$H = \{ z : \langle c, z \rangle \leq \langle c, x \rangle \},$$
The Ellipsoid method: Description

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

1. Query Separation Oracle at $x$.

2. If $x \notin K$, 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 $x$.

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

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

1. Query Separation Oracle at $x$.

2. If $x \not\in K$, 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 $x$.

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

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

(i) $K \subseteq B(0, R)$

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

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

Theorem

If, after $N = \lceil 2n(n + 1) \log(R/r) \rceil$ iterations, the ellipsoid algorithm didn’t find a point $x \in K$, then $K$ is empty.
Analysis for optimization problems

\[ p^* = \inf_{x \in K} c^T x , \quad (P) \]

- 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 \(\delta\) such that \(\{x \in K : c^T x \leq \delta\} \neq \emptyset\).

- But as \(\delta\) approaches \(p^*\), the \((\delta - p^*)\)-suboptimal set becomes very small: will assumption (ii) still hold?
Under (i) and (ii), the $\epsilon$-suboptimal set cannot be too small:

**Proposition**

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

Analysis for optimization problems

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

**Proposition**

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

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

\[
O \left( n^2 \log \frac{R}{\min(r, \epsilon)} \right)
\]

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

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 $K$, $c$, $\epsilon$, either

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

**Definition (Weak separation).**

Given $K$, $x$, $\epsilon$, either
- Assert that $x \in K^{+\epsilon}$;
- or return $h$ with $\|h\|_{\infty} = 1$ such that $h^T x^* \leq h^T y + \epsilon$, $\forall y \in K^{-\epsilon}$.
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 (w.r.t. polytime complexity):

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.
Alternative cutting-plane approaches

- Historically, the ellipsoid method was the first method that could solve a convex feasibility problem within polytime using a separation oracle.
- Total complexity:
  \[ O(n^2 \log \frac{R}{r} (SO + n^2)) \]
- Nowadays, new cutting-planes methods exist, differing on which set \( E_k \supseteq K \) is maintained, and at which point \( x_k \) we query the separation oracle. In particular,
  - Inscribed ellipsoid
  - Analytic center
  - Random walk
- Best method to date [Lee, Sidford & Wong, 2015]:
  \[ O(n \log \frac{nR}{r} SO + n^3 \log^{O(1)} \frac{nR}{r}) \).