1 Overview

In our previous lecture we discussed several motivating examples for linear optimization. We used the problem about optimizing ad budget to show that finding an optimal solution is highly nontrivial, even in two dimensions. We then discussed four major questions which are central to our study of linear optimization, and optimization in general.

1. When is an LP feasible / infeasible?
2. When is an LP bounded / unbounded?
3. What is a characterization of optimality?
4. How do we design algorithms that find optimal solutions to (bounded and feasible) LPs (and optimization problems in general)?

Today we will describe a criterion for checking when an LP is infeasible. We will then give a simple algorithm for testing when an LP is unbounded. We will conclude by discussing the third question on characterization of optimality. We will use a fundamental theorem from calculus which gives a sufficient (though not necessary) condition for when an optimization problem has an optimal solution.

2 Certificate for Infeasibility

Let’s recall the example we used in the previous lecture:

\[
\begin{align*}
\text{max} & \quad x_1 + x_2 + x_3 \\
\text{s.t.} & \quad x_1 + x_2 + x_3 = 6 \\
& \quad 2x_1 + 3x_2 + x_3 = 8 \\
& \quad 2x_1 + x_2 + 3x_3 = 0
\end{align*}
\]

We were interested in understanding whether the LP has a feasible region, and we noticed that if we multiply the first, second, and third row by 4, −1, −1 respectively, we get that an inconsistency, namely that \(0x_1 + 0x_2 + 0x_3 = 16\).

We can write the set of constraints as the matrix \(A\) and vector \(b\) below:
\[
\begin{pmatrix}
1 & 1 & 1 \\
2 & 3 & 1 \\
2 & 1 & 3
\end{pmatrix},
\begin{pmatrix}
6 \\
8 \\
0
\end{pmatrix}
\] (5)

Stated in these terms, we are interested in knowing whether there exists an \(x \in \mathbb{R}^3\) s.t. \(Ax = b\). By performing raw operations on the matrix (Gaussian elimination), we saw that there exists a vector \(p = (4, -1, -1)\) s.t. \(p^T Ax = 0\) and \(p^T b = 16 \neq 0\). This was the certificate of infeasibility. We know this procedure as the fundamental theorem in linear algebra.

**Theorem 1** (Gauss). Let \(A\) be a \(m \times n\) matrix, and let \(b\) be a \(m \times 1\) vector. Then exactly one of the following statements hold, but not both:

(I) \(\exists x \in \mathbb{R}^n : Ax = b\) or

(II) \(\exists p \in \mathbb{R}^m\) s.t. \(p^T A = 0\) and \(p^T b \neq 0\).

So whenever the constraints can be encoded as \(Ax = b\), we can use the Gaussian elimination process to tell us whether an LP is feasible; if it is not feasible we are able to easily produce the certificate \(p\) which tells us that no solutions that satisfy the constraints exist. In general, we would like to be able to perform such tests for the inequality \(Ax \leq b\). This is Farkas Lemma\(^1\).

**Theorem 2** (Farkas). Let \(A\) be a \(m \times n\) matrix, and \(b\) is a \(m \times 1\) vector. Then exactly one of the following statements hold, but not both:

(I) \(\exists x \in \mathbb{R}^n : Ax \leq b\) or

(II) \(\exists p \in \mathbb{R}^m\) s.t. \(p^T A = 0\) and \(p^T b < 0\), \(p \geq 0\).

As we will soon see, we also have a constructive manner to find this vector \(p\) which serves as an infeasibility certificate in cases where the set is indeed infeasible. This will therefore give us an algorithm to test whether a set is infeasible. The main component in proving Farkas Lemma is the *separating hyperplane theorem* which we leave as a homework exercise in your first problem set.

Let’s introduce some definitions and see how we use this result.

**Definition 3** (hyperplane). Given a vector \(a \in \mathbb{R}^n\) and constant \(\alpha \in \mathbb{R}\) the hyperplane \(H\) defined by \(a\) and \(\alpha\) is \(H = \{x \in \mathbb{R}^n \mid a^T x = \alpha\}\).

The hyperplane is simply a generalization of a line. A *half space* is the set of all points above or below the hyperplane.

**Definition 4** (half space). Every hyperplane \(H\) defines half spaces \(H^- := \{x \in \mathbb{R}^n \mid a^T x \leq \alpha\}\) and \(H^+ := \{x \in \mathbb{R}^n \mid a^T x \geq \alpha\}\).

**Definition 5.** A set is closed if and only if it contains all its limit points.

The separating hyperplane theorem can be stated as follows:

\(^1\)The result is known as Farkas Lemma, but we write it here as a theorem.
**Theorem 6** (Separating Hyperplanes). For every nonempty, closed, convex set $C$ and $b \notin C$, there exists a hyperplane separating $C$ and $b$. That is, there exists a point $p \in \mathbb{R}^n, p \neq 0$ s.t. $p^T y \leq 0 < p^T b \forall y \in C$. Furthermore, this point $p$ is $b - P_C(b)$ where $P_C(b)$ is the projection of $b$ on $C$ (see definition in the problem set).

The fact that we know how to characterize $p$ in the above theorem statement and that it can be represented as $b - P_C(b)$ implies we can test whether a set is infeasible by finding $P_C(b)$. Given the above theorem, we can now prove Farkas Lemma. We will prove the equivalent form of the lemma, stated as follows.

**Theorem 7** (Farkas). Let $A$ be a $m \times n$ matrix, and $b$ is a $m \times 1$ vector. Then exactly one of the following statements hold, but not both:

(I) $\exists x \in \mathbb{R}^n : Ax = b$, $x \geq 0$ or

(II) $\exists p \in \mathbb{R}^m \setminus \{0\}$ s.t. $p^T A \leq 0$ and $p^T b > 0$.

**Proof of Farkas Lemma.** We will first show that if (I) is true then (II) is necessarily false. Assume $Ax = b$ for some $x \geq 0$. If $p^T A \leq 0$ then for $x \geq 0$ we have that $p^T Ax \leq 0$. Since $Ax = b$ this implies that $p^T b \leq 0$, and thus it cannot be that both $p^T Ax \leq 0$ and $p^T b > 0$.

Now we’ll prove that if (I) is false then (II) is necessarily true. Define:

$$C = \{ y \in \mathbb{R}^n : \exists x \geq 0 \ Ax = y \};$$

Since (I) is false $b \notin C$. From the separating hyperplane theorem, we know there exists $p \in \mathbb{R}^m, p \neq 0$ s.t. $p^T y \leq 0$ and $p^T b > 0$, for all $y \in C$. Since $y = Ax$ that implies that $\forall x \geq 0$ we have that $p^T Ax \leq 0$ and $p^T b > 0$, as required.
3 Certificate for Unboundedness

3.1 Example of Unboundedness

Suppose we want to minimize a function.

\[
\begin{align*}
\min & \quad x_1 \\
\text{s.t.} & \quad x_1 + 3x_2 + 2x_3 = 6 \\
& \quad x_2 - x_3 = 0
\end{align*}
\]

The general solution can be written as follows for \( w \in \mathbb{R} \):

\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{bmatrix} =
\begin{bmatrix}
  6 - 5w \\
  w \\
  w
\end{bmatrix} =
\begin{bmatrix}
  6 \\
  0 \\
  0
\end{bmatrix} + w \cdot
\begin{bmatrix}
  -5 \\
  1 \\
  1
\end{bmatrix}
\]

Note that the objective value is \( x_1 = 6 - 5w \to -\infty \) as \( w \to \infty \), and it is therefore unbounded.

To understand when the LP is unbounded, let’s consider a geometric interpretation of the set of constraints.

**Definition 8 (polyhedron).** An intersection of a finite number of half spaces is called a polyhedron, and can be written as \( P = \{ x \in \mathbb{R}^n \mid Ax \leq b \} \).

Observe that the set of constraints of an LP is a polyhedron.

**Definition 9 (Recession cone).** Let \( P \) be a non-empty polyhedron. The recession cone of \( P \), denoted \( P^\circ \) is the set of all directions \( d \) that are contained in \( P \). That is, \( P^\circ := \{ d \in \mathbb{R}^n \mid \exists x \in P : x + \lambda d \in P \ \forall \lambda > 0 \} \).

**Proposition 10.** Let \( P = \{ x \in \mathbb{R}^n \mid Ax \leq b \} \) and \( Q := \{ d \in \mathbb{R}^n \mid Ad \leq 0 \} \). Then, \( P^\circ = Q \).

**Proof.** Let \( d \in Q \), then: \( Ad \leq 0 \). If \( Ax \leq b \) then for any \( \lambda \geq 0 \) we have:

\[
A(x + \lambda d) = Ax + \lambda Ad \leq Ax + 0 = Ax \leq b.
\]

And thus, by definition of \( P^\circ \), \( d \in P^\circ \). Conversely, if \( d \in P^\circ \) then there exists some \( x \) for which \( A(x + \lambda d) \leq b \). If we assume, for purpose of contradiction that \( Ad > 0 \), then for any \( b \), we can find a large enough \( \lambda \) which contradicts \( A(x + \lambda d) \leq b \). Thus, \( Ad \leq 0 \).

**Theorem 11.** Let \( P \) be a non-empty polyhedron, and consider \( \max \{ c^T x : x \in P \} \). Then, the LP is unbounded if and only if \( \exists d \in P^\circ \text{ s.t. } c^T d > 0 \) (for minimization problems: \( c^T d < 0 \)).

We will now only prove one of the directions of this theorem, and leave the second direction after we prove the duality theorem next week.

**Proof.** Let \( d \in P^\circ \text{ s.t. } c^T d > 0 \). Then:

\[
\lim_{\lambda \to \infty} c^T(x + \lambda d) = \lim_{\lambda \to \infty} c^T x + \lambda c^T d = \infty.
\]

\( \Box \)
An algorithm to test unboundedness. From the above two lemmas we have a simple algorithm to test unboundedness of an LP. Given the LP \( \max \{ c^T x : x \in P \} \) we simply try to find a solution to the system of equations \( c^T d > 0 \) and \( A d \leq 0 \).

From now on, we will only consider bounded polyhedrons, called polytope.

Definition 12 (polytope). A bounded polyhedron is called a polytope.

4 Sufficient Condition for Optimality

Let’s first recall some basics from real analysis.

Definition 13. The value \( \alpha \) is the infimum of \( f \) over \( F \) if it is the largest value for which any point \( x \in F \) respects \( f(x) \geq \alpha \).

Theorem 14 (Bolzano-Weierstrass). Every bounded sequence in \( \mathbb{R}^n \) has a convergent subsequence.

The following theorem, which is a fundamental theorem in real analysis, gives us a sufficient (though not necessary) condition for optimality.

Theorem 15 (Weierstrass). Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a continuous function and \( F \subseteq \mathbb{R}^n \) be nonempty, bounded, and closed. Then, \( \min \{ f(x) : x \in F \} \) has an optimal solution.

Proof. Let \( \alpha \) be the infimum of \( f \) over \( F \), and for some \( \epsilon \in (0, 1) \) define \( F^k := \{ x \in F : \alpha \leq f(x) \leq \alpha + \epsilon^k \} \). Notice that \( F^k \) is bounded, for any \( k \). By the Bolzano-Weierstrass theorem we know that there is a convergent subsequence. Now, consider a convergent subsequence of points \( \{ x^k \}_{k=1}^{\infty} \). Since \( F \) is closed we have that:

\[
\lim_{k \to \infty} \{ x^k \} = \bar{x} \in F.
\]

Since \( f \) is continuous we have that \( \lim_{k \to \infty} f(x^k) = f(\bar{x}) \). The optimal solution is \( \bar{x} \). \( \square \)