AM 121: Introduction to Optimization
Models and Methods

Lecture 4: Basic Feasible Solutions, Extreme points

Yiling Chen
Harvard SEAS

Lesson Plan

- Polyhedron, convexity and optimality
- Basis, basic feasible solutions (BFS), optimality

- Optimization ↔ Geometry ↔ Algebra

- Jensen & Bard: 3.1, 3.3
Notation

- Vector will usually mean column vector
- $m \times n$ matrix:

\[
A = (A_1 \ A_2 \ \ldots \ \ A_n) = \begin{pmatrix}
\mathbf{a}_1^\top \\
\ldots \\
\mathbf{a}_m^\top
\end{pmatrix}
\]

$A_i$ is the $i$-th column, $a_j^\top$ is the $j$-th row.
- $A_B$ is matrix formed from columns in set $B$ (e.g., $B = \{1, 3\}$).
- $S \setminus T$ is set of elements of $S$ that do not belong to $T$
- $\mathbb{R}^n$ is set of $n$-dimensional vectors
- $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top \in \mathbb{R}^n; \ A\mathbf{x} = \sum_i x_i A_i$
- $\mathbf{x} \geq 0$ and $\mathbf{x} > 0$: every component of $\mathbf{x}$ is nonnegative (respectively, positive)

Recall: LP Standard Forms

- Standard equality form:

\[
\begin{align*}
\text{max} \quad & \mathbf{c}^\top \mathbf{x} \\
\text{s.t.} \quad & A\mathbf{x} = b \\
& \mathbf{x} \geq 0
\end{align*}
\]

- maximization problem
- $m$ equality constraints
- $n$ inequality constraints

- Standard inequality form:

\[
\begin{align*}
\text{max} \quad & \mathbf{c}^\top \mathbf{x} \\
\text{s.t.} \quad & A\mathbf{x} \leq b \\
& \mathbf{x} \geq 0
\end{align*}
\]

- Can transform any LP into standard forms.
Convex Sets

- Given \( y, z \in \mathbb{R}^n, y \neq z \) define **open line segment** 
  \((y, z) = \{ \lambda y + (1 - \lambda)z \mid 0 < \lambda < 1 \} \).

- **Definition:** A set \( F \subseteq \mathbb{R}^n \) is **convex** when 
  \( y, z \in F, y \neq z \Rightarrow (y, z) \subseteq F \).

---

Polyhedron

**Definition**

A **polyhedron** is a set that can be described in the form 
\( P = \{ x \in \mathbb{R}^n \mid Ax \leq b \} \).

**Fact**

The feasible set of any LP can be described as a polyhedron.
Definition

Let $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} \neq \mathbf{0}$, and let $b$ be a scalar. Then,

- $\{x \in \mathbb{R}^n \mid \mathbf{a}^\top x = b\}$ is a set of points that forms a **hyperplane**
- $\{x \in \mathbb{R}^n \mid \mathbf{a}^\top x \leq b\}$ is a set of points that forms a **halfspace**

- **Note:** vector $\mathbf{a}$ is perpendicular to the hyperplane
- If $\mathbf{x}$ and $\mathbf{y}$ are on the same hyperplane, then $\mathbf{a}^\top \mathbf{x} = \mathbf{a}^\top \mathbf{y}$ and $\mathbf{a}^\top (\mathbf{x} - \mathbf{y}) = 0$ and $\mathbf{a}$ is orthogonal to any vector on the hyperplane.
- $\mathbf{x} - \mathbf{y}$ is a vector on the hyperplane
**Theorem**

A halfspace is convex.

**Theorem**

The intersection of convex sets is convex.

**Theorem**

Every polyhedron is an intersection of halfspaces, and convex.

---

**Intuition: Optimality**

In two dimensions, convexity $\equiv$ angle inside feasible region at corner is $\leq 180^\circ$.

Let $v$ denote value of optimal solution $x^*$. Hyperplane $\mathbf{c}^\top \mathbf{x}^* = v$ “separates” all points in the feasible polyhedron from $x^*$, showing they have less objective value.
Extreme Points

**Definition**
Let $P$ be a polyhedron. $x \in P$ is an extreme point if we cannot find $y, z \in P$, both different from $x$, such that $x \in (y, z)$.

\[
x \text{ is an extreme point: if } x = \lambda y + (1-\lambda) z \text{ and } \\
\lambda \in (0,1) \text{ for } y, z \neq x \text{ then either } \\
y \not\in P \text{ or } z \not\in P.
\]

**Existence of Extreme Points**

**Definition**
A polyhedron contains a line if there exists a $x \in P$ and a nonzero vector $d \in P$ s.t. $x + \lambda d \in P$ for all scalars $\lambda$.

**Theorem**
A polyhedron $P$ contains an extreme point if and only if it does not contain a line.

e.g., $P$ contains a line, but $Q$ does not (figure from MIT 6.251).
Optimality of Extreme Points

Theorem

Consider $\max \ c^\top x$ over a polyhedron $P$. Suppose $P$ has at least one extreme point, and there exists an optimal solution. Then there exists an optimal, extreme solution.

Proof.

- $P = \{x \in \mathbb{R}^n | Ax \leq b\}$, $x^*$ optimal with $v = c^\top x^*$.
- Define $Q = \{x \in \mathbb{R}^n | Ax \leq b, c^\top x = v\} \subseteq P$; a polyhedron with an extreme point.
- Let $x'$ be an extreme point of $Q$. Suppose for contradiction that $x'$ is not extreme in $P$. Then, $x' = \lambda y + (1 - \lambda)z$, for $y, z \in P$, not equal to $x'$, with $\lambda \in (0, 1)$.
- We have $c^\top x' = \lambda c^\top y + (1 - \lambda)c^\top z = v$, and by optimality of $v$ we have $c^\top y \leq v$ and $c^\top z \leq v$. So, $c^\top y = c^\top z = v$, and $z, y \in Q$. But then $x'$ is not an extreme point in $Q$. Contradiction.

- How can an LP have an optimal solution but no extreme point?
- How can an LP have an extreme point but no optimal solution?
Review: Basis

- The span of vectors $y_1, \ldots, y_K$ in $R^m$ is the set of vectors $z$ of the form $z = \sum_k d_k \cdot y_k$, where $d_k$ is a scalar.

- Vectors $y_1, \ldots, y_K$ are linearly-independent if and only if the only solution of $\sum_k d_k \cdot y_k = 0$ is $d_k = 0$ for all $k$. (If linearly dependent, then one can be written as the linear combination of the others.)

- A basis of $R^m$ is a collection of linearly-independent vectors in $R^m$ that span $R^m$. (Any $m$ linearly-independent vectors will provide a basis.)

Examples

Consider $R^2$. What about:

- $\{ (1, 0)^T, (0, 1)^T \}$
- $\{ (1, 0)^T, (1, 1)^T \}$
- $\{ (1, 0)^T, (2, 0)^T \}$
Review: Matrix Properties

- Columns of $A = (A_1 \ldots A_n)$ are **linearly independent** if and only if the only solution of $Ax = 0$ is $x = 0$.
  
  Example:
  
  \[
  \begin{pmatrix}
  1 & 1 \\
  1 & 0
  \end{pmatrix}
  \quad ?
  \quad \begin{pmatrix}
  0 & 0 \\
  2 & 1
  \end{pmatrix}
  \quad ?
  \]

- Columns of $m$ by $n$ matrix $A$ span $R^m$ if $Ax = b$ has a solution for every $b \in R^m$.
  
  Example:
  
  \[
  \begin{pmatrix}
  1 & 0 & 1 & 0 \\
  1 & 2 & 0 & 1
  \end{pmatrix}
  \quad ?
  \quad \begin{pmatrix}
  0 & 0 \\
  2 & 1
  \end{pmatrix}
  \quad ?
  \]

- Rank of matrix $A$ is the size of largest collection of linearly independent columns (the **column rank**)
  
  - equivalently, the size of largest collection of linearly independent rows (the **row rank**)
  
  - Fact: column rank = row rank $\leq \min(m, n)$

Review: Invertible Matrices

- $m$ by $m$ matrix $A$ is **invertible** if there is some $A'$ such that $AA' = A'A = I_m$ (the identity matrix, 0s off-diagonal, 1s on-diagonal.)

- Following properties are equivalent for a square matrix:
  
  - $A$ is invertible
  
  - columns of $A$ span
  
  - columns of $A$ are linearly independent
  
  - for every $b \in R^m$, $Ax = b$ has a unique solution
Basis of a Matrix

- Consider an \( m \times n \) matrix \( A \)
- \( B \) is a basis for \( A \) if \( A_B \) is invertible (the columns of \( A_B \) are linearly independent and span \( \mathbb{R}^m \)). Need \(|B| = m\).
- Example: \( A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \)
  
  For \( B = \{1, 3\} \), obtain \( A_B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \).
  
  \( A_B \) is invertible, and so \( \{1, 3\} \) is a basis for \( A \).
- Extension rule: If columns of \( A \) span, and columns of \( A_C \) are linearly independent for \(|C| < m\), can extend \( C \) to form a basis \( B \) for \( A \).

Basic Solutions

**Definition**

\( x \) is a basic solution to \( Ax = b \) if the vectors of \( A \) with \( x_i \neq 0 \) are linearly independent.

- For basis \( B \), let \( B' \) denote \( \{1, \ldots, n\} \setminus B \). Call variables \( x_B \) basic and variables \( x_{B'} \) nonbasic.
  
  Example: for \( B = \{1, 3\} \), \( x_B = (x_1, x_3) \), \( x_{B'} = (x_2, x_4) \).

**Definition**

The basic solution corresponding to basis \( B \) is obtained by setting \( x_{B'} = 0 \) and solving for \( x_B \).

- \( A_Bx_B + A_{B'}x_{B'} = A_Bx_B + 0 = b \), and so \( x_B = A_{B}^{-1}b \). Unique solution (\( A_B \) invertible). Some \( x_B \) values may be 0!

**Fact**

\( x \) is a basic solution if and only if there is a basis \( B \) s.t. non-basic variables \( x_{B'} = 0 \). (via extension rule).
Example: Basic Solutions

Standard equality form (also canonical here):

\[
\begin{align*}
\text{max} & \quad x_1 + x_2 \\
\text{s.t.} & \quad x_1 \leq 2 \\
& \quad x_1 + 2x_2 \leq 4 \\
& \quad x_1, x_2 \geq 0
\end{align*}
\]

\[
\begin{align*}
\text{max} & \quad x_1 + x_2 \\
\text{s.t.} & \quad x_1 + x_3 = 2 \\
& \quad x_1 + 2x_2 + x_4 = 4 \\
& \quad x_1, x_2, x_3, x_4 \geq 0
\end{align*}
\]

- Five bases: \( \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{3, 4\} \).
- Corresponding basic solutions: \((x_1, x_2, x_3, x_4)^\top = (2, 1, 0, 0)^\top, (4, 0, -2, 0)^\top, (2, 0, 0, 2)^\top, (0, 2, 2, 0)^\top, (0, 0, 2, 4)^\top\)
- All feasible?

Basic Feasible Solution (BFS)

- Constraints \( Ax = b \), and \( x \geq 0 \).

**Definition**

A **basic feasible solution** is basic and feasible.

- In the example, there are 4 BFS, each of which corresponds to a feasible solution \((x_1, x_2)^\top = (2, 1)^\top, (2, 0)^\top, (0, 2)^\top, (0, 0)^\top\) of the original LP. The other basic solution is infeasible, not \( x \geq 0 \).
BFS occur at the “corners” of the feasible region
Geometrically, optimization finds a “corner” solution
Corners correspond exactly to BFS

Optimization ↔ geometry ↔ algebra

Let’s prove the correspondence between extreme points and BFS

Main LP Assumptions

- \( \max c^\top x \) s.t. \( Ax = b, x \geq 0 \)
- \( A \) has full row rank:
  - wlog because if not then have redundancy because one row can be written as linear combination of other rows
- Less rows than columns \((m < n)\)
  - wlog because this makes it an optimization problem!
  - \( n \) variables, \( m \) equations. \((n - m)\) is the “degree of freedom”
- **Columns of \( A \) span \( R^m \)**
  - meaning \( Ax = b \) has a solution for every \( b \)
  - follows from full row rank, and thus column rank = \( m \)
BFS and Extreme Points

Theorem

Consider \( P = \{ \mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0 \} \), for \( \mathbf{A} \) with columns that span. Then extreme points of \( P \) are exactly the BFS of \( P \).

Proof.

\((\Leftarrow)\) A BFS is an extreme point:
- Suppose \( \mathbf{x} \) is a BFS corresponding to basis \( B \).
- For contradiction, assume \( \mathbf{x} \) is not an extreme point, i.e. \( \mathbf{x} = \lambda \mathbf{y} + (1 - \lambda)\mathbf{z} \) for \( \mathbf{z}, \mathbf{y} \in P \), \( \mathbf{y} \neq \mathbf{z} \) and some \( \lambda \in (0,1) \).
- For all \( i \notin B \), we have \( x_i = 0 \), and because \( \mathbf{y}, \mathbf{z} \geq 0 \), we must have \( y_i = z_i = 0 \) for all \( i \notin B \).
- Conclude that \( \mathbf{y} \) and \( \mathbf{z} \) are the same BFS as \( \mathbf{x} \). Why? They have the same basic and non-basic variables.
- Contradiction!

\((\Rightarrow)\) An extreme point is a BFS:
- Let \( \mathbf{x} \) be an extreme point. Assume for contradiction it is not basic. Let \( C = \{ i : x_i > 0 \} \), with \( |C| = k \).
- \( \mathbf{A}_C \) must not have linearly independent columns:
  - if \( k \leq m \) and cols linearly ind., then \( \mathbf{x} \) would be basic.
  - if \( k > m \), would imply column rank larger than \( m \)!
- Let \( \mathbf{d}' \) denote a non-zero vector in \( \mathbb{R}^k \) such that \( \mathbf{A}_C \mathbf{d}' = \mathbf{0} \). Define \( \mathbf{d} \in \mathbb{R}^n \) with \( d_i = d_i' \) for \( i \in C \), and \( d_i = 0 \) otherwise.
- For small \( \epsilon > 0 \), points \( \mathbf{x} \pm \epsilon \mathbf{d} \) are distinct (since \( \mathbf{d} \neq \mathbf{0} \)) and both in \( P \) (since \( \mathbf{A}_C \mathbf{d}' = \mathbf{0} \) and thus \( \mathbf{A}\mathbf{d} = \mathbf{0} \), and with \( \mathbf{x} \pm \epsilon \mathbf{d} \geq \mathbf{0} \) for small \( \epsilon \) since \( d_i = 0 \) whenever \( x_i = 0 \)).
- A contradiction with \( \mathbf{x} \) being an extreme point.

\(\square\)
Summary

- If there’s an extreme point, and an optimal solution, then there’s an optimal solution at an extreme point.
- All non-zero variables in a basic solution $\mathbf{x}$ correspond to linearly independent columns of $\mathbf{A}$. There is also a basis $B$ ($|B| = m$, $A_B$ rank $m$) s.t. non-basic $\mathbf{x}_{B^c} = \mathbf{0}$.
- Extreme points of the polyhedron are exactly the basic feasible solutions (BFS) (basic and $\mathbf{x} \geq \mathbf{0}$)
- Suggests we can solve LPs by searching through BFS. *But can we do this efficiently?*