AM 121: Introduction to Optimization
Models and Methods

Lecture 4: Basic Feasible Solutions, Extreme points

Yiling Chen
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:

\[
A = \begin{pmatrix}
A_1 & A_2 & \ldots & A_n
\end{pmatrix}
= \begin{pmatrix}
a_1^T \\
\vdots \\
a_m^T
\end{pmatrix}
\]

where \( A_i \) is the \( i \)-th column and \( a_j^T \) 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 \( T \)
- Use \( \mathbb{R}^n \) for set of all \( n \)-dimensional vectors
- \( x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n \)
- \( x \geq 0 \) and \( x > 0 \): every component of \( x \) is nonnegative (respectively, positive)

Recall: LP Canonical Form

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

- maximization problem
- \( m \) equality constraints
- \( n \) inequality constraints
- \( b \geq 0 \)
- one “isolated variable” in each constraint

- Can transform any LP into canonical form.
- (Can also transform any LP into **standard inequality form**):

\[
\begin{align*}
\text{max} & \quad c^T x \\
\text{s.t.} & \quad Ax \leq b
\end{align*}
\]

where \( b \) can take any value, no need for isolated variables.)
Convex Sets

- Given $y, z \in \mathbb{R}^n$, define $(y, z) = \{ \lambda y + (1 - \lambda)z \mid 0 < \lambda < 1 \}$.
- If $y \neq z$, this is the open line segment between $y$ and $z$.
- **Definition:** A set $F \subseteq \mathbb{R}^n$ is convex when $y, z \in F \Rightarrow (y, z) \subseteq F$.

Polyhedron

**Definition**

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

**Fact**

The feasible set of any LP can be described in this way, and in particular $\{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \}$ is a polyhedron.

- Recall: Any $\{ x \in \mathbb{R}^n \mid Ax \geq b \}$ can be expressed as $\{ x \in \mathbb{R}^{n'} \mid Ax = b, x \geq 0 \}$ (for some $n'$)
**Halfspaces**

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

- $\{x \in \mathbb{R}^n \mid a^T x = b\}$ is a hyperplane
- $\{x \in \mathbb{R}^n \mid a^T x \geq b\}$ is a halfspace

**Note:** $a$ is perpendicular to the hyperplane

[If $x$ and $y$ are on the same hyperplane, then $a^T x = a^T y$ and $a^T(x - y) = 0$ and $a$ is orthogonal to any vector confined to 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: Convexity

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

We can draw a separating hyperplane through $x^*$ that places all of polyhedron in halfspace with value $\leq c^T x^*$. 
Extreme Points

**Definition**
Let $P$ be a polyhedron. A vector $x \in P$ is an extreme point of $P$ if we cannot find $y, z \in P$, both different from $x$, and $\lambda \in (0, 1)$ such that $x = \lambda y + (1 - \lambda)z$.

**Optimality of Extreme Points**

**Theorem**
Consider $\max c^T x$ 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.

Next, we will develop an algebraic notion of extreme points.
Review: Basis

- The \textbf{span} of vectors $y^{(1)}, \ldots, y^{(K)}$ in $R^m$ is the set of vectors $z$ of the form $z = \sum_{k=1}^{K} d_k y^{(k)}$, where $d_k$ is a scalar. (All vectors that are a linear combination of $y^{(1)}, \ldots$).
- Vectors $y^{(1)}, \ldots, y^{(K)}$ are \textbf{linearly-independent} if the only solution of $\sum_{k=1}^{K} d_k y^{(k)} = 0$ is $d_k = 0$, for all $k$.
- A \textbf{basis} of $R^m$ is a collection of \textbf{linearly-independent} vectors in $R^m$ that \textbf{span} $R^m$. (Must be $m$ vectors. Any $m$ linearly independent vectors provide a basis.)

\begin{itemize}
  \item \textbf{Examples}
  \end{itemize}

Consider $R^2$. What about:

- $\{(1, 0)^T, (0, 1)^T\}$
- $\{(1, 0)^T, (1, 1)^T\}$
- $\{(1, 0)^T, (2, 0)^T\}$

\begin{itemize}
  \item \textbf{Review: Matrix Properties}
  \end{itemize}

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

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

- \textbf{Rank} of matrix $A$ is the size of largest collection of linearly independent rows (the \textbf{row rank})
  \[
  \begin{itemize}
    \item\text{equivalently, the size of largest collection of linearly independent columns (the \textbf{column rank})}
    \item\text{column rank} = \text{row rank} \leq \min(m, n)
  \end{itemize}
Review: Invertible Matrices

- \( m \) by \( m \) matrix \( A \) is **invertible** if there is some \( B \) such that \( AB = BA = I_m \) (identity matrix, 0’s off-diagonal, 1’s 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 \( b \in \mathbb{R}^m \), \( Ax = b \) 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 (equivalently, 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 \( A \).
Basic Solutions

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

**Definition**

The basic solution to $Ax = b$ corresponding to basis $B$ is obtained by setting $x_{B'} = 0$ and solving for $x_B$.

- We have $A_B x_B + A_{B'} x_{B'} = A_B x_B + 0 = b$.
- Solve, obtain $x_B = A_B^{-1} b$, which has a (unique) solution because $A_B$ is invertible.
- Note: some of $x_B$ may be zero!

Main LP Assumptions

- $\max c^T x \quad \text{s.t.} \quad Ax = b, x \geq 0$ (canonical)
- Matrix $A$ has **full row rank**: all rows are linearly independent (needs $m \leq n$)
  - without loss of generality (wlog) since if not true then have redundancy (or infeasibility, with $m > n$)
- Matrix $A$ has **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**
  - meaning $Ax = b$ has a solution for every $b$
  - wlog because of the isolated variables in canonical form!
Existence of A Basis

- **Extension rule:** if cols of $A$ span, and for set $C$ ($|C| < m$), $A_C$ has linearly independent cols, then can extend $C$ to basis for $A$ by adding a vector not in span of $A_C$.

Example: $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}$. $C = \{3\}$. $A_C = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Clearly $A_C$ is linearly independent. Can extend $C = \{3\}$ to a basis; for example $\{2, 3\}$.

---

**Theorem**

* A matrix has a basis if and only if its columns span.

**Proof.**

$(\Rightarrow)$ *Basis then cols span:* trivial.

$(\Leftarrow)$ *Cols span then basis:* use extension rule, with $C = \{}$.

- An LP in canonical form will **always** have a basis!
Example

Canonical form:

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

Example

Canonical form:

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

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

- Constraints $Ax = b, x \geq 0$

**Definition**

A **basic feasible solution** is a basic solution that is feasible.

- In the example, there are 4 BFS, each of which corresponds to a feasible solution $(x_1, x_2)^T = (2, 1)^T, (2, 0)^T, (0, 2)^T, (0, 0)^T$ of the original LP.

BFS occur at the “corners” of the feasible region
- Geometrically, optimization finds a “corner” solution
- Corners correspond exactly to BFS

Optimization $\leftrightarrow$ geometry $\leftrightarrow$ algebra
BFS and Extreme Points

Theorem

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

Proof.

(\( \Leftarrow \)) BFS is extreme:

- Suppose \( x^* \) is a BFS corresponding to basis \( B \).
- For contradiction, assume \( x^* \) is not an extreme point, i.e. \( x^* = \lambda y + (1 - \lambda)z \) for \( z, y \in P \), \( y \neq z \) and some \( \lambda \in (0, 1) \).
- For all \( i \notin B \), we have \( x^*_i = 0 \), and because \( y, z \geq 0 \), we must have \( y_i = z_i = 0 \) for all \( i \notin B \).
- Because \( Ay = b \) (\( y \) is feasible) and \( y_i = 0 \) for all \( i \notin B \), then \( y \) is the same basic solution as \( x^* \).
- Similarly, \( z = x^* \). Contradiction with \( y \neq z \!).

(\( \Rightarrow \)) extreme is BFS:

- Let \( x^* \) be an extreme point (and thus feasible).
- For contradiction, assume \( x^* \) is not basic. Let \( C = \{ i : x^*_i > 0 \} \), with \( |C| = n' \).
- \( A_C \) must not have linearly independent columns:
  - if \( n' \leq m \) and cols linearly ind, then \( x^* \) would be basic! (For \( n' < m \), use ext. rule.)
  - if \( n' > m \), then \#col \( > \) \#rows, and cols not linearly ind.
- So, exists some \( d' \in R^{n'} \), \( d' \neq 0 \), with \( A_C d' = 0 \). Define \( d \in R^n \) with \( d_i = d'_i \) for \( i \in C \), and \( d_i = 0 \) otherwise.
- For small \( t > 0 \), vectors \( x^* + td, x^* - td \) are distinct (since \( d \neq 0 \) and in \( P \) (since \( Ad = 0 \), and \( d_i = 0 \) whenever \( x^*_i = 0 \)).
- \( x^* \) is on an open line segment. Contradiction with \( x^* \) extreme!
Summary

- Exists an optimal solution at an extreme point
- A basic solution is one where there is a basis (set of linearly independent columns that span) such that the non-basic variables are zero
- Extreme points are exactly the basic feasible solutions (BFS)
- Exists an optimal solution at a BFS
- Suggests we can solve LPs by searching through BFS