

AM 121: Introduction to Optimization Models and Methods

Lecture 4: Basic Feasible Solutions, Extreme points

David C. Parkes
SEAS

1

Lesson Plan

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

- Optimization \leftrightarrow Geometry \leftrightarrow Algebra

- Jensen & Bard: 3.1, 3.3

2

Notation

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

$$\mathbf{A} = (\mathbf{A}_1 \quad \mathbf{A}_2 \quad \dots \quad \mathbf{A}_n) = \begin{pmatrix} \mathbf{a}_1^\top \\ \dots \\ \mathbf{a}_m^\top \end{pmatrix}$$

\mathbf{A}_i is the i -th column, \mathbf{a}_j^\top is the j -th row.

- \mathbf{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
- R^n is set of n -dimensional vectors
- $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top \in R^n$; $\mathbf{Ax} = \sum_i x_i \mathbf{A}_i$
- $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{x} > \mathbf{0}$: every component of \mathbf{x} is nonnegative (respectively, positive)

3

Recall: LP Canonical Form

$$\begin{aligned} \max \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

- maximization problem
- m equality constraints
- n inequality constraints
- $\mathbf{b} \geq \mathbf{0}$
- one isolated variable in each equality constraint

- Can transform any LP into canonical form.
- (Standard inequality form:

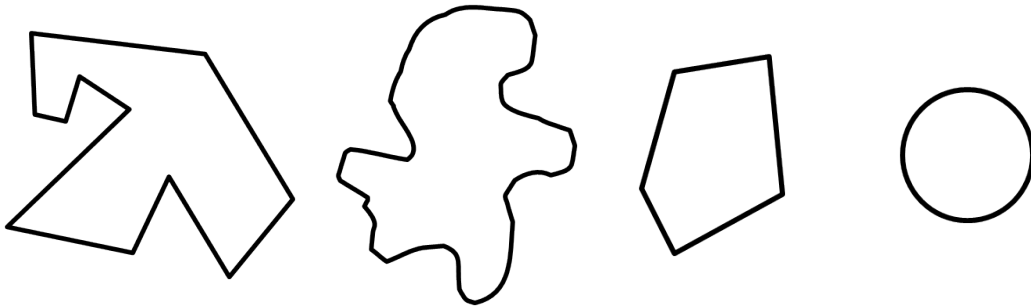
$$\begin{aligned} \max \quad & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

where \mathbf{b} can take any value, and no need for isolated variables.)

4

Convex Sets

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



5

Polyhedron, Polytopes

Definition

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

Definition

A **polytope** is a bounded polyhedron (i.e., there is some z such that $P \subseteq \{\mathbf{x} \in R^n : -z \leq x_j \leq z \text{ for } j = 1, \dots, n\}$)

Fact

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

6

Halfspaces

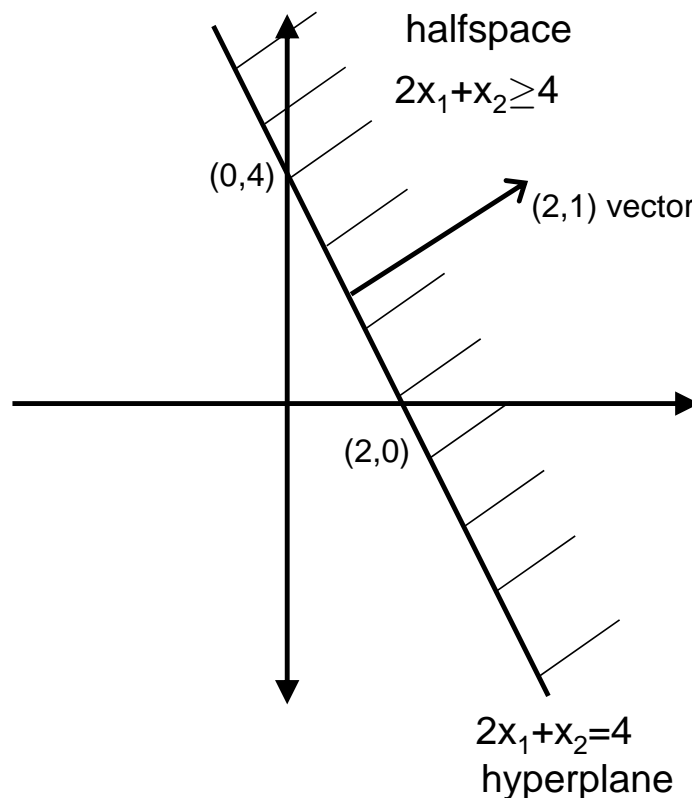
Definition

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

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

7



8

Theorem

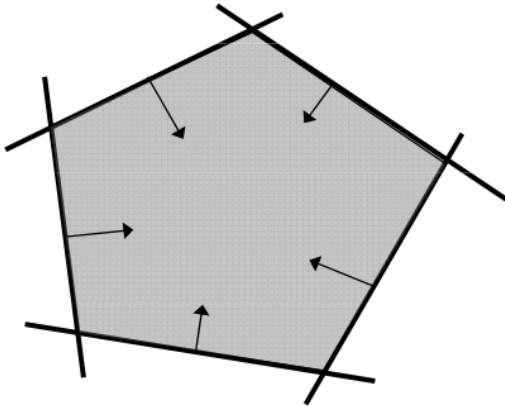
A halfspace is convex.

Theorem

The intersection of convex sets is convex.

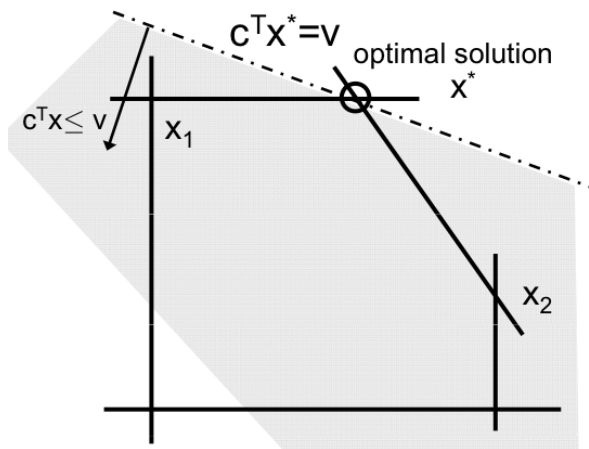
Theorem

Every polyhedron is an intersection of halfspaces, and convex.



9

Intuition: Optimality



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

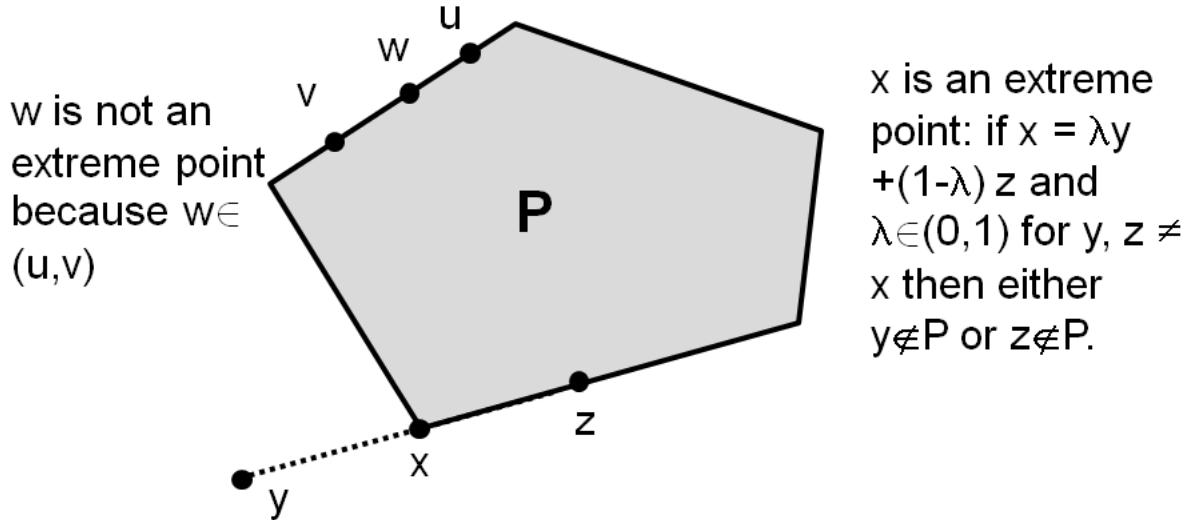
Let v denote value of optimal solution x^* . Hyperplane $c^T x^* = v$ “separates” all points in the feasible polyhedron from x^* , showing they have less objective value.

10

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 $x \in (y, z)$.



11

Existence of Extreme Points

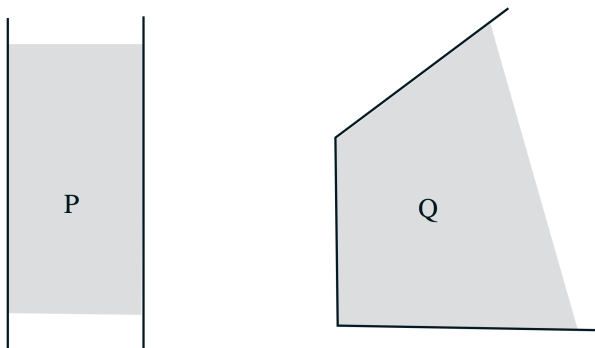
Definition

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 λ .

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).



12

Optimality of Extreme Points

Theorem

Consider $\max \mathbf{c}^\top \mathbf{x}$ 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.

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

□

13

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

Overview

Optimization \leftrightarrow Geometry \leftrightarrow Algebra

15

Review: Basis

- The **span** of vectors $\mathbf{y}_1, \dots, \mathbf{y}_K$ in R^m is **the set of vectors \mathbf{z} of the form $\mathbf{z} = \sum_k d_k \cdot \mathbf{y}_k$** , where d_k is a scalar.
- Vectors $\mathbf{y}_1, \dots, \mathbf{y}_K$ are **linearly-independent** if and only if the **only** solution of $\sum_k d_k \cdot \mathbf{y}_k = 0$ 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:

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

16

Review: Matrix Properties

- Columns of $\mathbf{A} = (\mathbf{A}_1 \dots \mathbf{A}_n)$ are **linearly independent** if and only if the **only** solution of $\mathbf{Ax} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.

Example:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} ? \quad \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix} ?$$

- Columns of m by n matrix \mathbf{A} **span** R^m if $\mathbf{Ax} = \mathbf{b}$ has a solution for every $\mathbf{b} \in R^m$.

Example:

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} ? \quad \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix} ?$$

- **Rank** of matrix \mathbf{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 $\leq \min(m, n)$

17

Review: Invertible Matrices

- m by m matrix \mathbf{A} is **invertible** if there is some \mathbf{A}' such that $\mathbf{AA}' = \mathbf{A}'\mathbf{A} = \mathbf{I}_m$ (the *identity matrix*, 0s off-diagonal, 1s on-diagonal.)
- Following properties are equivalent for a square matrix:
 - ▶ \mathbf{A} is invertible
 - ▶ columns of \mathbf{A} span
 - ▶ columns of \mathbf{A} are linearly independent
 - ▶ for every $\mathbf{b} \in R^m$, $\mathbf{Ax} = \mathbf{b}$ has a unique solution

18

Basis of a Matrix

- Consider an $m \times n$ matrix \mathbf{A}
- B is a **basis** for \mathbf{A} if \mathbf{A}_B is **invertible** (the columns of \mathbf{A}_B are linearly independent and span R^m). Need $|B| = m$.

- Example: $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}$

For $B = \{1, 3\}$, obtain $\mathbf{A}_B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

\mathbf{A}_B is invertible, and so $\{1, 3\}$ is a basis for \mathbf{A} .

- **Extension rule:** If columns of \mathbf{A} span, and columns of \mathbf{A}_C are linearly independent for $|C| < m$, can extend C to form a basis B for \mathbf{A} .

19

Basic Solutions

Definition

\mathbf{x} is a **basic solution** to $\mathbf{A}\mathbf{x} = \mathbf{b}$ if the vectors of \mathbf{A} with $x_i \neq 0$ are linearly independent.

- For basis B , let B' denote $\{1, \dots, n\} \setminus B$. Call variables \mathbf{x}_B **basic** and variables $\mathbf{x}_{B'}$ **nonbasic**.
 - ▶ Example: for $B = \{1, 3\}$, $\mathbf{x}_B = (x_1, x_3)$, $\mathbf{x}_{B'} = (x_2, x_4)$.

Definition

The **basic solution corresponding to basis B** is obtained by setting $\mathbf{x}_{B'} = \mathbf{0}$ and solving for \mathbf{x}_B .

- $\mathbf{A}_B \mathbf{x}_B + \mathbf{A}_{B'} \mathbf{x}_{B'} = \mathbf{A}_B \mathbf{x}_B + \mathbf{0} = \mathbf{b}$, and so $\mathbf{x}_B = \mathbf{A}_B^{-1} \mathbf{b}$. Unique solution (\mathbf{A}_B invertible). Some \mathbf{x}_B values may be 0!

Fact

\mathbf{x} is a basic solution if and only if there is a basis B s.t. non-basic variables $\mathbf{x}_{B'} = \mathbf{0}$. (via extension rule).

20

Example: Basic Solutions

Canonical form:

$$\begin{array}{llll} \max & x_1 + x_2 & & \\ \text{s.t.} & x_1 & \leq & 2 \\ & x_1 + 2x_2 & \leq & 4 \\ & x_1, x_2 & \geq & 0 \end{array} \qquad \begin{array}{llll} \max & x_1 + x_2 & & \\ \text{s.t} & x_1 & +x_3 & = 2 \\ & x_1 + 2x_2 & + x_4 & = 4 \\ & x_1, x_2, & x_3, x_4 & \geq 0 \end{array}$$

- 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?

21

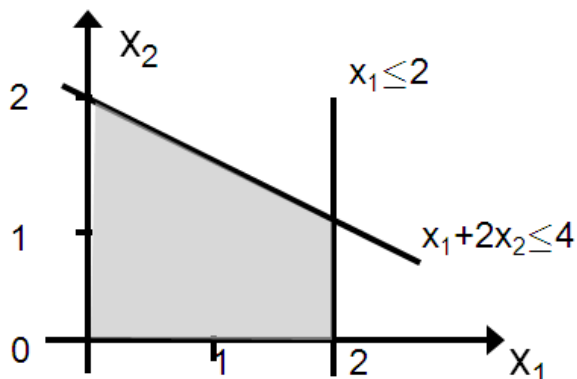
Basic Feasible Solution (BFS)

- Constraints $A\mathbf{x} = \mathbf{b}$, and $\mathbf{x} \geq \mathbf{0}$.

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 $\mathbf{x} \geq \mathbf{0}$.



22

- 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

- Let’s prove the correspondence between extreme points and BFS

23

Main LP Assumptions

- $\max \mathbf{c}^\top \mathbf{x}$ s.t. $\mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0$ (and *canonical*)
- \mathbf{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 \mathbf{A} span \mathbb{R}^m**
 - ▶ meaning $\mathbf{Ax} = \mathbf{b}$ has a solution for every \mathbf{b}
 - ▶ follows from full row rank, and thus column rank = m

24

BFS and Extreme Points

Theorem

Consider $P = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, 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 $\mathbf{y}, \mathbf{z} \geq \mathbf{0}$, 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!

□

25

BFS and Extreme Points

Proof.

(\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 R^k such that $\mathbf{A}_C \mathbf{d}' = \mathbf{0}$. Define $\mathbf{d} \in R^n$ 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 ϵ since $d_i = 0$ whenever $x_i = 0$).
- A contradiction with \mathbf{x} being an extreme point.

□

26

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 ($|B| = m$, \mathbf{A}_B rank m) s.t. non-basic $\mathbf{x}_{B'} = \mathbf{0}$.
- Extreme points of the polytope 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?*