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:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \ldots & \boldsymbol{A}_n \end{pmatrix} = \begin{pmatrix} {\boldsymbol{a}_1}^\top \\ \ldots \\ {\boldsymbol{a}_m}^\top \end{pmatrix}$$

  $\boldsymbol{A}_i$ is the i-th column, ${\boldsymbol{a}_j}^\top$ is the j-th row.
- $\boldsymbol{A}_B$ 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
- $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^\top \in R^n$; $\boldsymbol{A}\boldsymbol{x} = \sum_i x_i \boldsymbol{A}_i$
- $\boldsymbol{x} \geq \boldsymbol{0}$ and $\boldsymbol{x} > \boldsymbol{0}$: every component of $\boldsymbol{x}$ is nonnegative (respectively, positive)

Recall: LP Standard Forms

- **Standard inequality form**:

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

  - maximization problem
  - m equality constraints
  - n inequality constraints

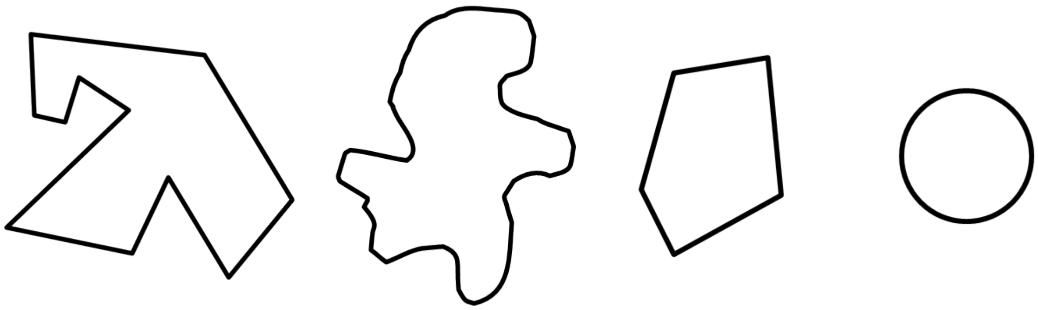
- **Standard inequality form**:

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

- Can transform any LP into standard forms.

Convex Sets

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



Polyhedron

Definition

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

Fact

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

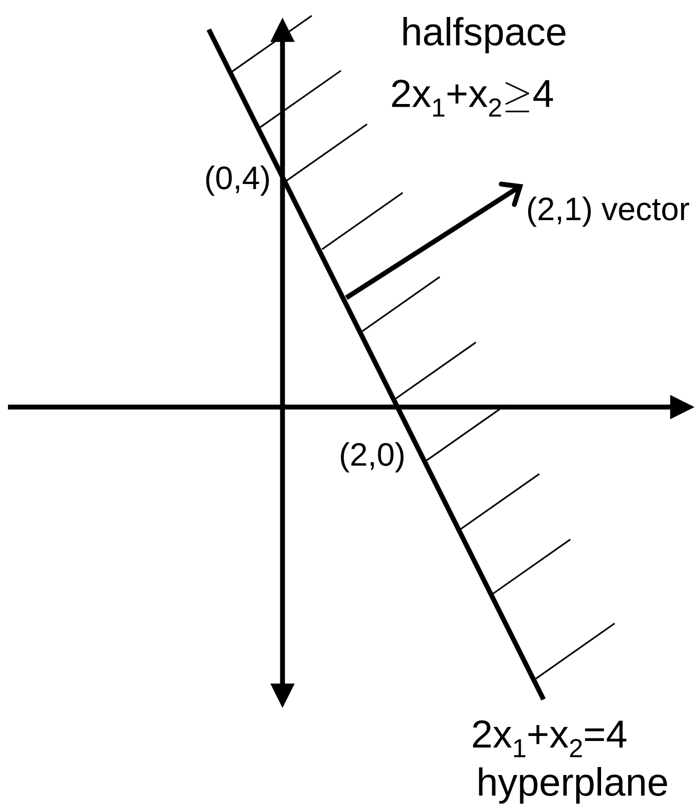
Halfspaces

Definition

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

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

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



Theorem

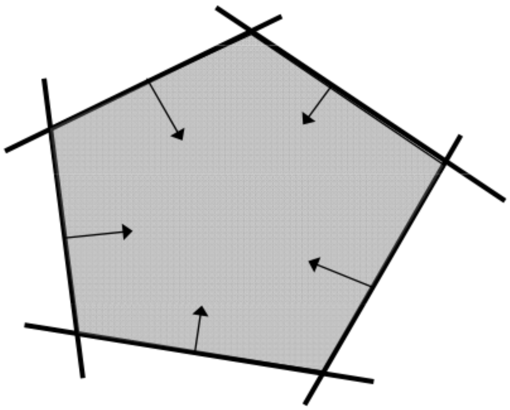
A halfspace is convex.

Theorem

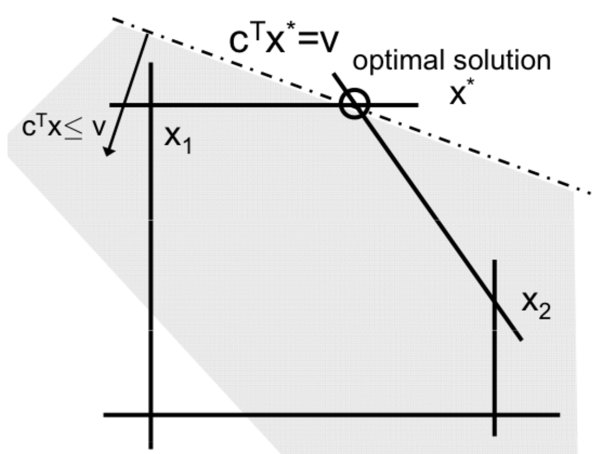
The intersection of convex sets is convex.

Theorem

Every polyhedron is an intersection of halfspaces, and convex.



Intuition: Optimality



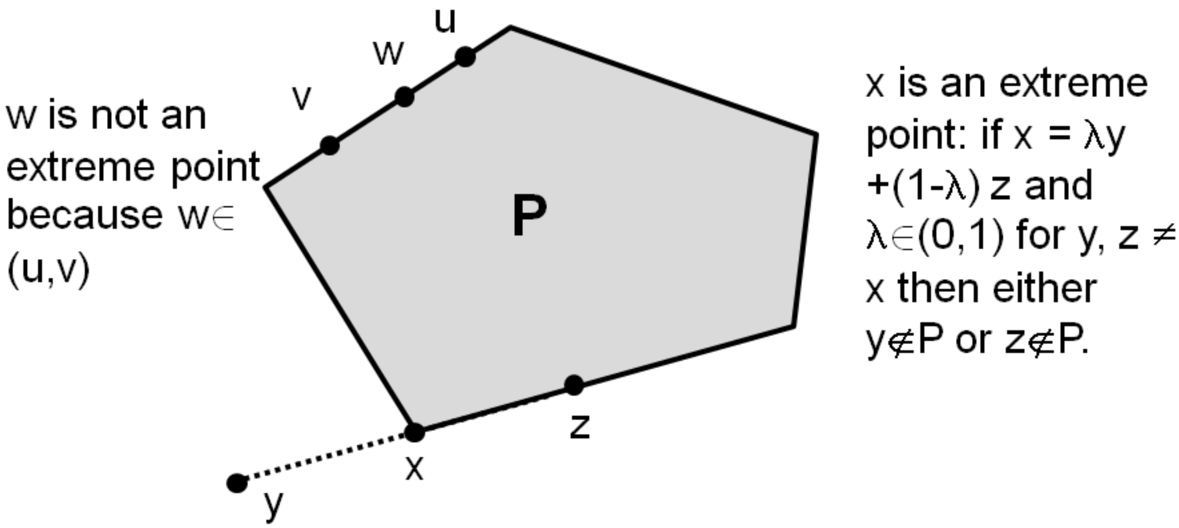
In two dimensions, convexity $\equiv$ angle inside feasible region at corner is $\leq 180^o$.

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

Extreme Points

Definition

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



Existence of Extreme Points

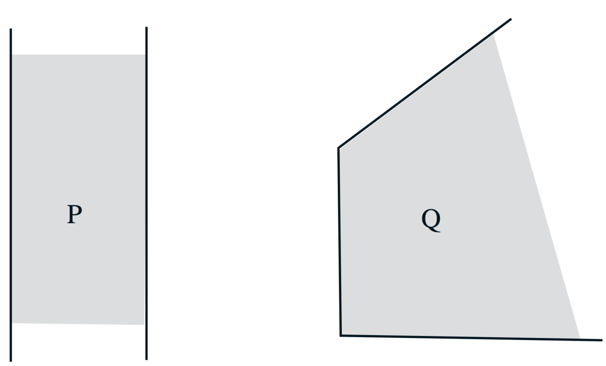
Definition

A polyhedron **contains a line** if there exists a $\boldsymbol{x} \in P$ and a nonzero vector $\boldsymbol{d} \in P$ s.t. $\boldsymbol{x} + \lambda \boldsymbol{d} \in P$ 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).



Optimality of Extreme Points

Theorem

Consider $\max \boldsymbol{c}^\top \boldsymbol{x}$ *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.

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

$\square$

- 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

Review: Basis

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

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

Review: Matrix Properties

- Columns of $\boldsymbol{A} = (\boldsymbol{A}_1 \ldots \boldsymbol{A}_n)$ are **linearly independent** if and only if the **only** solution of $\boldsymbol{Ax} = \mathbf{0}$ is $\boldsymbol{x} = \mathbf{0}$.
  Example:
  $$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \ ? \qquad \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix} \ ?$$
- Columns of m by n matrix $\boldsymbol{A}$ **span** R^m if $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution for every $\boldsymbol{b} \in R^m$.
  Example:
  $$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \ ? \qquad \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix} \ ?$$
- **Rank** of matrix $\boldsymbol{A}$ 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 $\boldsymbol{A}$ is **invertible** if there is some $\boldsymbol{A}'$ such that $\boldsymbol{AA}' = \boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I}_m$ (the *identity matrix*, 0s off-diagonal, 1s on-diagonal.)
- Following properties are equivalent for a square matrix:
  - $\boldsymbol{A}$ is invertible
  - columns of $\boldsymbol{A}$ span
  - columns of $\boldsymbol{A}$ are linearly independent
  - for every $\boldsymbol{b} \in R^m$, $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution

Basis of a Matrix

- Consider an $m \times n$ matrix $\boldsymbol{A}$
- B is a **basis** for $\boldsymbol{A}$ if $\boldsymbol{A}_B$ **is invertible** (the columns of $\boldsymbol{A}_B$ are linearly independent and span R^m). Need $|B| = m$.
- Example: $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}$

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

  $\boldsymbol{A}_B$ is invertible, and so $\{1, 3\}$ is a basis for $\boldsymbol{A}$.
- **Extension rule**: If columns of $\boldsymbol{A}$ span, and columns of $\boldsymbol{A}_C$ are linearly independent for $|C| < m$, can extend C to form a basis B for $\boldsymbol{A}$.

Basic Solutions

Definition

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

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

Definition

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

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

Fact

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

Example: Basic Solutions

$$\begin{array}{lll} \max & x_1 + x_2 & \\ \text{s.t.} & x_1 & \le 2 \\ & x_1 + 2x_2 & \le 4 \\ & x_1, x_2 & \ge 0 \end{array}$$

Standard equality form (also canonical here):

$$\begin{array}{llllll} \max & x_1 + x_2 & & & \\ \text{s.t} & x_1 & +x_3 & & = 2 \\ & x_1 + 2x_2 & & +x_4 & = 4 \\ & x_1,\ x_2, & x_3,\ x_4 & & \ge 0 \end{array}$$

- 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?

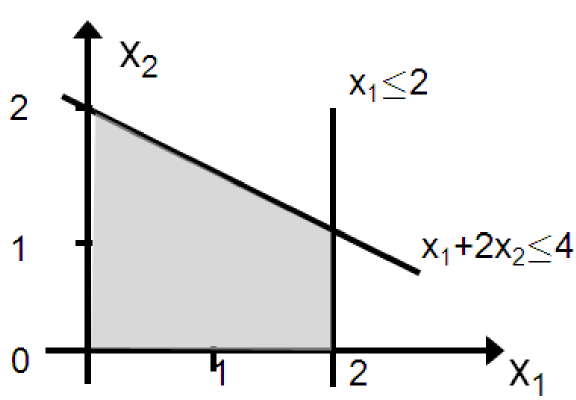
Basic Feasible Solution (BFS)

- Constraints $\boldsymbol{Ax} = \boldsymbol{b}$, and $\boldsymbol{x} \ge 0$.

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 $\boldsymbol{x} \ge \boldsymbol{0}$.



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

$$\text{Optimization} \leftrightarrow \text{geometry} \leftrightarrow \text{algebra}$$

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

Main LP Assumptions

- $\max \boldsymbol{c}^\top \boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0$
- $\boldsymbol{A}$ 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 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a solution for every $\boldsymbol{b}$
  - follows from full row rank, and thus column rank $= m$

BFS amd Extreme Points

Theorem

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

Proof.

($\Leftarrow$) *A BFS is an extreme point*:

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

□

BFS and Extreme Points

Proof.

($\Rightarrow$) *An extreme point is a BFS*:

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

□

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 $\boldsymbol{x}$ correspond to linearly independent columns of $\boldsymbol{A}$. There is also a basis B ($|B| = m$, $\boldsymbol{A}_B$ rank m) s.t. non-basic $\boldsymbol{x}_{B'} = \mathbf{0}$.
- Extreme points of the polyhedron are exactly the basic feasible solutions (BFS) (basic and $\boldsymbol{x} \geq \mathbf{0}$)
- Suggests we can solve LPs by searching through BFS. *But can we do this efficiently?*