CO 255

Instructor: Jim Geelen

FALL 2018

# Contents

# Info

## Instructor

Jim Geelen (Instructor)

- Office MC 5124
- Phone x35594
- Email jim.geelen@uwaterloo.ca
- Office hour: 3:00 to 4:00 p.m. Tuesday (or by appointment)

## TA

Adam Brown (TA)

- Office MC 5462
- Email ajmbrown@uwaterloo.ca
- Office hours: By appointment

## Overview

This course serves as an introduction to optimization, with particular emphasis on convex optimization, linear optimization, and combinatorial optimization.

### Topics

- Introduction
- Linear Programming: feasibility, unboundedness, duality
- Polyhedra: polyhedral cones, extreme points, faces, constructing polyhedra
- Solving Linear Programs: Simplex Algorithm, testing feasibility, finding extreme points, perturbation method
- Combinatorial Optimization: integer programming, total unimodularity, weighted bipartite matching

- Convex Geometry: Separating Hyperplane Theorem, duality for cones, extreme points
- Convex Optimization: convex functions, normal cones and tangent cones, optimality conditions, Ellipsoid Method
- Complexity Theory: linear algebra, linear programming, integer linear programming

**Suggested reading**

- A. Schrijver, Theory of Integer and Linear Programming, Wiley 1998.
- V. Chvatal, Linear Programming, W.H. Freeman and Company, 1983.
- J.M. Borwein and A.S. Lewis, Convex Analysis and Nonlinear Optimization, Second Edition, Springer, 2006. (Electronic copy.)

## Assignments

Assignments 40%
There will be six assignments:

- Assignment 1, due September 19
- Assignment 2, due October 3
- Assignment 3, due October 17
- Assignment 4, due October 31
- Assignment 5, due November 14
- Assignment 6, due November 28

Solutions will not be posted.

**Late policy:** You may submit one assignment late (but not Assignment 6) provided that you email the instructor before the start of class in which the assignment is due. Late assignments should be submitted at the start of the following class.

Uncollected assignments will be disposed of after the final exam.

## Final

Final exam 60%
Information will be posted in November.

# Intro

Given a set $S$ (the feasible region) and a function $f : S \to \mathbb{R}$ (the objective function) solve

$$\min(f(x) : x \in S) \tag{1}$$

$$\max(f(x) : x \in S) \tag{2}$$

Note that

$$\max(f(x) : x \in S) = \min(-f(x) : x \in S)$$

Problem (1) may not be well posed. For example:

- (1) may be <u>infeasible</u>: that is $S = \varnothing$
- (1) may be <u>unbounded</u>: that is there may exist $x \in S$ with $f(x)$ arbitrarily small

Even if (1) is feasible and bounded it may not be well posed.
For example, $\min(x : x > 1)$

## Infinum and Supremum

Consider

$$\max(z \in \mathbb{R} : z \leq f(x) \text{ for all } x \in S) \tag{3}$$

If (1) is feasible and bounded, then (3) has an optimal solution.
We define,

$$\inf(f(x), x \in S) = \begin{cases} \infty & \text{(1) is infeasible} \\ -\infty & \text{(1) is unbounded} \\ \max\,(z \in \mathbb{R} : z \leq f(x) \text{ for all } x \in S) & \text{otherwise} \end{cases}$$

and $\sup(f(x) : x \in S) = -\inf(-f(x) : x \in S)$.

### Optimal Value

We let

$$\mathrm{OPT}(1) = \inf(f(x) : x \in S)$$

and

$$\mathrm{OPT}(2) = \sup(f(x) : x \in S)$$

## 0.1 Some optimization problems

**Linear programming**

$$f(x) = \mathbf{c}^T\mathbf{x} \qquad \text{and } S = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$$
$$\text{where } A \in \mathbb{R}^{m\times n}, \mathbf{b} \in \mathbb{R}^m \text{ and } \mathbf{c} \in \mathbb{R}^n$$

**Integer Linear programming**

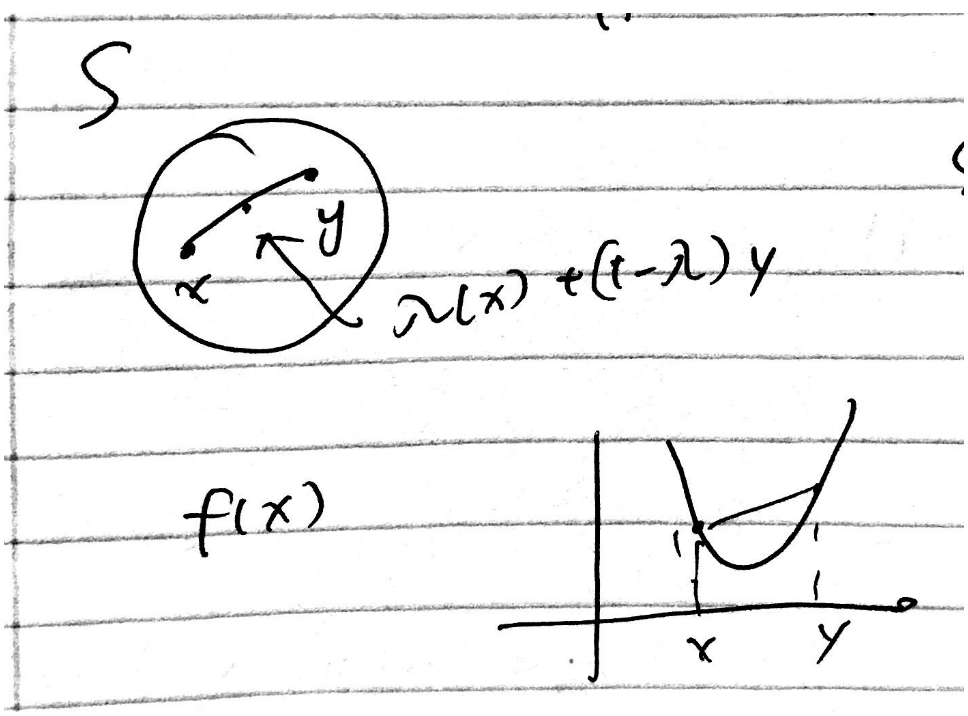
$$f(x) = \mathbf{c}^T\mathbf{x} \qquad \text{and } S = \{\mathbf{x} \in \mathbb{Z}^n : A\mathbf{x} \leq \mathbf{b}\}$$

**Convex Optimization**

$$S \subseteq \mathbb{R}^n \text{ is convex and } f : S \to \mathbb{R} \text{ is convex}$$

$S \subseteq \mathbb{R}^n$ is <u>convex</u> if for each $x,\ y \in S$ and $\lambda \in [0,\ 1]$,

$$\lambda x + (1-\lambda)y \in S$$




$f : S \to \mathbb{R}$ is <u>convex</u> if for each $x,\ y \in S$ and $\lambda \in [0,\ 1]$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

The <u>convex hull</u> of $S \subseteq \mathbb{R}^n$, denoted conv($S$), is (unique) minimal convex set contains $S$.

Consider an optimization problem $\min(f(x) : x \in S)$ where $S \subseteq \mathbb{R}^n$. and $f : \mathbb{R}^n \to \mathbb{R}$

We can "reduce" this to a convex optimization problem with a linear objective function.

<u>**Step 1**</u> Linearize the objective function.

Let $\hat{S} = \{\begin{pmatrix} x \\ y \end{pmatrix} : x \in S,\ y = f(x)\} \qquad \subseteq \mathbb{R}^{n+1}$

Then $\min(f(x) : x \in S) = \min(y : \begin{pmatrix} x \\ y \end{pmatrix} \in \hat{S})$

**Step 2** Convexify $S$
If $f:\mathbb{R}^n\to\mathbb{R}$ is linear, then

$$\min(f(x):x\in S)=\min(f(x):x\in\operatorname{conv}(S))$$

This one is theoretically true.

**Recall: (9.10)**

- Linear Programming
- Integer Programming
- Convex Optimization

**Examples:**

1. A two player game
Given $A\in\mathbb{R}^{m\times n}$, Rose chooses $i\in\{1,\cdots,m\}$, and Colin chooses $j\in\{1,\cdots,n\}$ (independently), then Colin pays Rose \$ $a_{ij}$.

   **eg**

$$A=\begin{bmatrix}2 & -2\\ 1 & 5\end{bmatrix}$$

   If Rose chooses 1, then her return is at least -2.
   If Rose chooses 2, then her return is at least 1.
   If Rose chooses with equal probability then her expected return is $\geq \min(\frac{1}{2}(2+1),\frac{1}{2}(-2+5))=\frac{3}{2}$
   Rose will choose her strategy maximizing her expected return in the worst case (i.e. that Colin guesses her strategy)

$$\begin{cases}\max(\min_{j\in\{1,\cdots,n\}}\sum_{i=1}^m p_i a_{ij})\\ p_1+\cdots+p_m=1\\ p_1,\cdots,p_m\geq 0\end{cases}$$

   or equivalently

$$(R)\begin{cases}\max z\\ \quad\downarrow \text{ subject to}\\ \begin{cases} z\leq \sum_{i=1}^m p_i a_{ij}, \quad j\in\{1,\cdots,n\}\\ p_1+\cdots+p_m=1\\ p_1,\cdots,p_m\geq 0\end{cases}\end{cases}$$

   Note that (R) is a linear program.
   Likewise Colin will choose his strategy using the following linear program:

$$(C)\begin{cases}\min z\\ \quad\downarrow \text{ subject to}\\ \begin{cases} z\geq \sum_{j=1}^n q_j a_{ij}, \quad i\in\{1,\cdots,m\}\\ q_1+\cdots+q_n=1\\ q_1,\cdots,q_m\geq 0\end{cases}\end{cases}$$

   Note that

   OPT(R) $\leq$ OPT(C)

   Surprising Fact: (R) and (C) have the same optimal value.
   Hence, it does not harm either Rose or Colin to reveal strategy.

2. Weighted bipartite matching.

**Problem:** Given $n$ jobs, $n$ workers, and a utility $a_{ij}$ of worker $i$ performing job $j$, find an assignment of workers to jobs of maximum total utility.

**Variables:** $x_{ij} \in \{0, 1\}$ where $x_{ij}$ indicates assigning worker $i$ to job $j$.

**Formulation:**

$$(P) \begin{cases} \max \sum_{i=1}^{n} \sum_{j=1}^{n} u_{ij} x_{ij} \\ \quad \downarrow \text{ subject to} \\ \begin{cases} \sum_{i=1}^{n} x_{ij} = 1, \quad j = 1, \cdots, n \\ \sum_{j=1}^{n} x_{ij} = 1, \quad i = 1, \cdots, n \\ x_{ij} \in \{0, 1\} \quad i, j \in \{1, \cdots, n\} \end{cases} \end{cases}$$

This is an integer linear program.
The linear relaxation of (P) is the linear program (P') obtained by replacing the last constraint with

$$0 \leq x_{ij} \leq 1, \quad i, j \in \{1, \cdots, n\}$$

Note that

OPT(P) $\leq$ OPT(P')

Surprising Fact: In this case OPT(P) = OPT(P')

3. 3D-Matching Problem

**Problem:** Given $A \in \mathbb{R}^{n \times n \times n}$ where $a_{ijk}$ is the utility of person $i$ performing job $j$ on machine $k$, find an assignment of maximum total utility.

**Formulation**

$$\max \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} a_{ijk} x_{ijk}$$

subject to

$$\sum_{j=1}^{n} \sum_{k=1}^{n} x_{ijk} = 1, \quad i = 1, \cdots, n$$

$$\sum_{i=1}^{n} \sum_{k=1}^{n} x_{ijk} = 1, \quad j = 1, \cdots, n$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} x_{ijk} = 1, \quad k = 1, \cdots, n$$

$$x_{ijk} \in \{0, 1\} \qquad i, \; j, \; k \in \{1, \cdots, n\}$$

$$- - - - - - - - - - - - - - - - - - - - - - - - - - (P)$$

In this case the inequality

OPT(P) $\leq$ OPT(P')

may be strict.

**Remark:** 3D matching is NP-hard, so integer linear programming is NP-hard. Note that we can replace $z \in \{0, 1\}$ with $z(z - 1) = 0$. so "quadratic programming" is NP-hard.

<u>**Recall:**</u> Integer linear programming is NP-hard.

4. Integer solutions to Diophantine Equations.

**Example**

$$(P)\begin{cases} \min \quad \sin(\pi x)^2 + \sin(\pi y)^2 + \sin(\pi z)^2 \\ \quad \downarrow \text{ subject to} \\ \begin{cases} x^3 + y^3 = z^3 \\ x, y, z \geq 1 \end{cases} \end{cases}$$

Note that OPT(P) $\geq 0$, and a feasible solution $(x, y, z)$ has objective value zero if and only if $x, y, z$ are positive integers satisfying $x^3 + y^3 = z^3$.

A Diophantine Equations is an equation

$$p(x_1, \cdots, x_n) = 0$$

where $p(x_1, \cdots, x_n)$ is a polynomial with integer coefficients.

**Hilbert's $\mathbf{10^{th}}$ problem:** Given a Diophantine equation, decide whether it has an integer solution.

**Formulation:**

$$(P)\begin{cases} \min \quad \sin(\pi x_1)^2 + \cdots + \sin(\pi x_n)^2 \\ \quad \downarrow \text{ subject to} \\ p(x_1, \cdots, x_n) = 0 \end{cases}$$

<u>**Remarks:**</u> Many famous problems are instances:

- The Four-Colour Theorem
- Riemann Hypothesis
- Goldbach's Conjecture

**Summary**

Optimization is hard. We need restrictive assumptions to develop theory and algorithms, even for convex optimization.

# 1 Linear Programming

## 1.1 Linear Programming Feasibility

**Problem:** Given $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$, does there exist $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} \geq \mathbf{b}$

**Remark:** For vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, we can have neither $\mathbf{a} \geq \mathbf{b}$ nor $\mathbf{a} \leq \mathbf{b}$. For example, $\mathbf{a} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

**Eliminating a variable.**

(Fourier-Motzkin Elimination)

**Example:**

$$\left.\begin{array}{l} x_2 \leq 2 \\ x_1 - x_2 \geq -1 \\ 2x_1 + x_2 \geq 2 \\ x_1 - 3x_2 \leq 1 \end{array}\right\} (1)$$

Rewrite as

$$\begin{array}{l} x_2 \leq 2 \quad \text{does not use } x_1 \\ \left.\begin{array}{l} x_1 \geq x_2 - 1 \\ x_1 \geq -\frac{1}{2}x_2 + 1 \end{array}\right\} \text{lower bounds on } x_1 \\ x_1 \leq 3x_2 + 1 \quad \text{upper bound on } x_1 \end{array}$$

So (1) has a solution if and only if there exists $x_2 \in \mathbb{R}$ satisfying:

(2.1) $x_2 \leq 2$
(2.2) $\max\,(x_2 - 1, -\frac{1}{2}x_2 + 1) \leq 3x_2 + 1$

or equivalently

$$\left.\begin{array}{l} x_2 \leq 2 \\ x_2 - 1 \leq 3x_2 + 1 \\ -\frac{1}{2}x_2 + 1 \leq 3x_2 + 1 \end{array}\right\} (3)$$

that is $0 \leq x_2 \leq 2$

More generally, consider (1) $\quad A\mathbf{x} \geq \mathbf{b} \qquad$ where $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$

Rewrite as

$$\begin{cases} f(x_1) \geq 0 \\ \vdots \\ f(x_m) \geq 0 \end{cases} \qquad \text{where } f_i(x) \geq a_{i1}x_1 + \cdots + a_{in}x_n - b_i$$

Scale so that $a_{in} \in \{0, 1, -1\}$ for each $i \in \{1, \cdots, m\}$

Define

$$A_1 = \{i \in \{1, \cdots, m\} : a_{in} \in A_1\}$$
$$A_{-1} = \{i \in \{1, \cdots, m\} : a_{in} \in A_{-1}\}$$
$$A_0 = \{i \in \{1, \cdots, m\} : a_{in} \in A_0\}$$

Rewrite (2) as:

(3.1) $\quad x_n \geq g_i(x_1, \cdots, x_{n-1}), \quad i \in A_1$
(3.2) $\quad x_n \leq g_i(x_1, \cdots, x_{n-1}), \quad i \in A_{-1}$
(3.3) $\quad 0 \leq g_i(x_1, \cdots, x_{n-1}), \quad i \in A_0$

Hence (1) has a solution if and only if the following system does

(4.1) $\quad 0 \leq g_i(x_1, \cdots, x_{n-1}), \quad i \in A_0$
(4.2) $\quad g_i(x_1, \cdots, x_{n-1}) \leq g_j(x_1, \cdots, x_{n-1})$
$\uparrow$—— for each $i \in A_1$ and $j \in A_{-1}$

The system (4) has only $n-1$ variables, but it has $O(m^{2^k})$ constraints, so this method is inefficient.

## 1.2 Polyhedra

A polyhedron is a set of the form $\{x \in \mathbb{R}^n : A\mathbf{x} \geq \mathbf{b}\}$ where $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. A polytope is a bounded polyhedron.

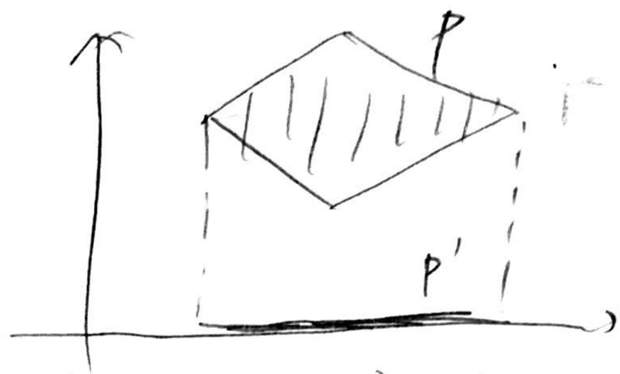
**Recall:**

- Fourier-Motzkin Elimination (Page 9)
- Polyhedra
  A polyhedron is a set of the form $\{x \in \mathbb{R}^n : A\mathbf{x} \geq \mathbf{b}\}$ where $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. A polytope is a bounded polyhedron.

## 1.3 Projection

Let $P \subseteq \mathbb{R}^n$, let $l < n$, and $P' = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_l \end{bmatrix} : x \in P \right\}$. We call $P'$ the projection of $P$ onto $x_1, \cdots, x_l$

Example




P' is the projection of P onto $x_1$

> **Theorem 1.1**
>
> If $P \subseteq \mathbb{R}^n$ is a polyhedron and $P'$ is the projection of $P$ onto $x_1, \cdots, x_l$, then $P'$ is a polyhedron.

**Proof:** Use Fourier-Motzkin elimination □

## 1.4 Certifying infeasibility

Recall

> **Fundamental Theorem of Linear Algebra**
>
> Let $\mathbb{F}$ be a field. For $A \in \mathbb{F}^{m \times n}$ and $\mathbf{b} \in \mathbb{F}^m$, exactly one of the following systems has a solution:
>
> (1) $(A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{F}^n)$
>
> (2) $(\mathbf{y}^T A = 0, \mathbf{y}^T \mathbf{b} = 1, \mathbf{y} \in \mathbb{F}^m)$

(That is if $A\mathbf{x} = \mathbf{b}$ is infeasible, then we can obtain the equation $0 = 1$ by taking a linear combination of the rows)

> **Farkas' Lemma**
>
> Let $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Exactly one of the following systems has a solution:
>
> (1) $(A\mathbf{x} \geq \mathbf{b}, \mathbf{x} \in \mathbb{R}^n)$
>
> (2) $(\mathbf{y}^T A = 0, \mathbf{y}^T \mathbf{b} = 1, \mathbf{y} \geq 0, \mathbf{y} \in \mathbb{R}^m)$

(That is if $A\mathbf{x} \geq \mathbf{b}$ has no solution then we can obtain the inequality $0 \geq 1$ as a non-negative combination of the constraints)

**Claim:** Easy direction (1) and (2) cannot both hold.

**Proof:** If (1) and (2) hold, then $0 = (\mathbf{y}^T A)\mathbf{x} = \mathbf{y}^T(A\mathbf{x}) \geq \mathbf{y}^T\mathbf{b} = 1$ □

**Example:**

$$\begin{array}{cl} x + 2y \leq 2 & (1.1) \\ x - y \geq 0 & (1.2) \\ 3x + 2y \leq 6 & (1.3) \\ y \geq 1 & (1.4) \end{array}$$

Consider (1.1)-(1.2)-3(1.4)

$$0 = (x + 2y) - (x - y) - 3y \leq 2 - 0 - 3 \times 1 = -1$$

Hence (1) is infeasible

**Implied inequalities**

A linear inequalities $\mathbf{a}^T\mathbf{x} \geq \mathbf{a}_0$ is implied by a system $A\mathbf{x} \geq \mathbf{b}$ if there is a non-negative vector $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{a} = A^T\mathbf{y}$ and $\mathbf{a}_0 = \mathbf{y}^T\mathbf{b}$

(This definition is non-standard; the more standard definition allows $\mathbf{a}_0 \leq \mathbf{y}^T\mathbf{b}$)

The hard direction of the Farkas' Lemma is that, if $A\mathbf{x} \geq \mathbf{b}$ is infeasible, then the inequality $0 \geq 1$ is implied by $A\mathbf{x} \geq \mathbf{b}$

Note that $A'\mathbf{x} \geq \mathbf{b}'$ is a set of implied inequalities of $A\mathbf{x} \geq \mathbf{b}$, and $A''\mathbf{x} \geq \mathbf{b}''$ is a set of implied inequalities of $A'\mathbf{x} \geq \mathbf{b}'$. Then $A''\mathbf{x} \geq \mathbf{b}''$ is implied by $A\mathbf{x} \geq \mathbf{b}$.

**Claim:** The system obtained from $A\mathbf{x} \geq \mathbf{b}$ by Fourier-Motzkin Elimination is implied by $A\mathbf{x} \geq \mathbf{b}$

**Proof:** Consider

$$x_n \leq g_1(x_1, \cdots, x_{n-1}) \qquad (1)$$

$$x_n \geq g_2(x_1, \cdots, x_{n-1}) \qquad (2)$$

Now (1) − (2) gives (1) minus (2)

$$g_2(x_1, \cdots, x_{n-1}) \leq g_1(x_1, \cdots, x_{n-1})$$

□

Hard direction of the Farkas' Lemma.

**Theorem**

If $A\mathbf{x} \geq \mathbf{b}$ is infeasible, then $0 \geq 1$ is an implied inequality.

**Proof:** An easy induction based on the previous claim. □

**Example**

$$\begin{array}{cl} x+2y \le 2 & (1.1) \\ x-y \ge 0 & (1.2) \\ 3x+2y \le 6 & (1.3) \\ y \ge 1 & (1.4) \end{array}$$

Eliminate $x$:

$$\begin{array}{cc} 0 \le -3y+2 & (1.1)-(1.2) \\ 0 \le -\frac{5}{3}y+2 & \frac{1}{3}(1.3)-(1.2) \\ y \ge 1 & (1.4) \end{array}$$

Thus

$$\begin{array}{cc} y \le \frac{2}{3} & \frac{1}{3}((1.1)-(1.2)) \\ y \ge 1 & (1.4) \end{array}$$

Eliminating $y$: $1 \le \dfrac{2}{3} \quad \dfrac{1}{3}((1.1)-(1.2))-(1.4)$

Thus $1 \le 0 \quad (1.1)-(1.2)-3(1.4)$

# Other forms:

**Theorem 1.2 (Another form of Farkas' Lemma)**

Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Then exactly one of the following systems has a solution.

(1) $(A\mathbf{x} = \mathbf{b},\ \mathbf{x} \ge \mathbf{0})$

(2) $(\mathbf{y}^T A \ge 0, \mathbf{y}^T\mathbf{b} = -1)$

If $x$ satisfies (1) and $y$ satisfies (2), then

$$0 \le (\mathbf{y}^T A)\mathbf{x} = \mathbf{y}^T(A\mathbf{x}) = \mathbf{y}^T\mathbf{b} = -1$$

— contradiction
So (1) and (2) cannot both have a solution.

Suppose that (1) has no solution, we can rewrite (1) as $(A\mathbf{x} \ge \mathbf{b}, A\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0})$

or equivalently

$$\begin{bmatrix} A \\ -A \\ I \end{bmatrix} \mathbf{x} \ge \begin{bmatrix} \mathbf{b} \\ \mathbf{b} \\ \mathbf{0} \end{bmatrix} \qquad \text{(i)}$$

By Farkas Lemma, if (i) has no solution, then there exist non-negative vectors $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^m$ and $\mathbf{z} \in \mathbb{R}^n$ such that

$$[\mathbf{y}_1{}^T, \mathbf{y}_2{}^T, \mathbf{z}^T] \begin{bmatrix} A \\ -A \\ I \end{bmatrix} = 0$$

and

$$[\mathbf{y}_1{}^T, \mathbf{y}_2{}^T, \mathbf{z}^T] \begin{bmatrix} \mathbf{b} \\ -\mathbf{b} \\ \mathbf{0} \end{bmatrix} = 1$$

That is

$$\mathbf{y}_1{}^T A - \mathbf{y}_2{}^T A + \mathbf{z} = 0, \qquad \mathbf{y}_1{}^T\mathbf{b} - \mathbf{y}_2{}^T\mathbf{b} = 1$$

So

$$(\mathbf{y}_2 - \mathbf{y}_1)^T A \geq 0, \qquad (\mathbf{y}_2 - \mathbf{y}_1)^T \mathbf{b} = -1$$

Setting $\mathbf{y} = \mathbf{y}_2 - \mathbf{y}_1$ gives a solution to (2) □

## 1.5 Geometric Intepretation

A set $\mathcal{K} \in \mathbb{R}^n$ is a cone if

- $0 \in \mathcal{K}$
- for each $x \in \mathcal{K}$ and $\lambda \geq 0$, we have $\lambda x \in \mathcal{K}$, and
- $\mathcal{K}$ has to be convex

For a set $\mathcal{S} \in \mathbb{R}^n$ we let cone($\mathcal{S}$) denote the smallest cone containing $\mathcal{S}$ (note that this is well-defined)

**Lemma 1.3**

If $a_1, \cdots, a_n \in \mathbb{R}^n$, then: $\text{cone}(\{a_1, \cdots, a_n\}) = \{\lambda_1 a_1 + \cdots + \lambda_k a_k : \lambda_1, \cdots, \lambda_k \geq 0\}$

**Proof:** Exercise. Similar to the assignment □

**Remark:** $b \in \text{conv}(\{a_1, \cdots, a_n\})$ if and only if $\begin{bmatrix} 1 \\ b \end{bmatrix} \in \text{conv}(\left\{ \begin{bmatrix} 1 \\ a_1 \end{bmatrix}, \cdots, \begin{bmatrix} 1 \\ a_k \end{bmatrix} \right\})$

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ and let $a_1, \cdots, a_n$ denote the columns of $A$.

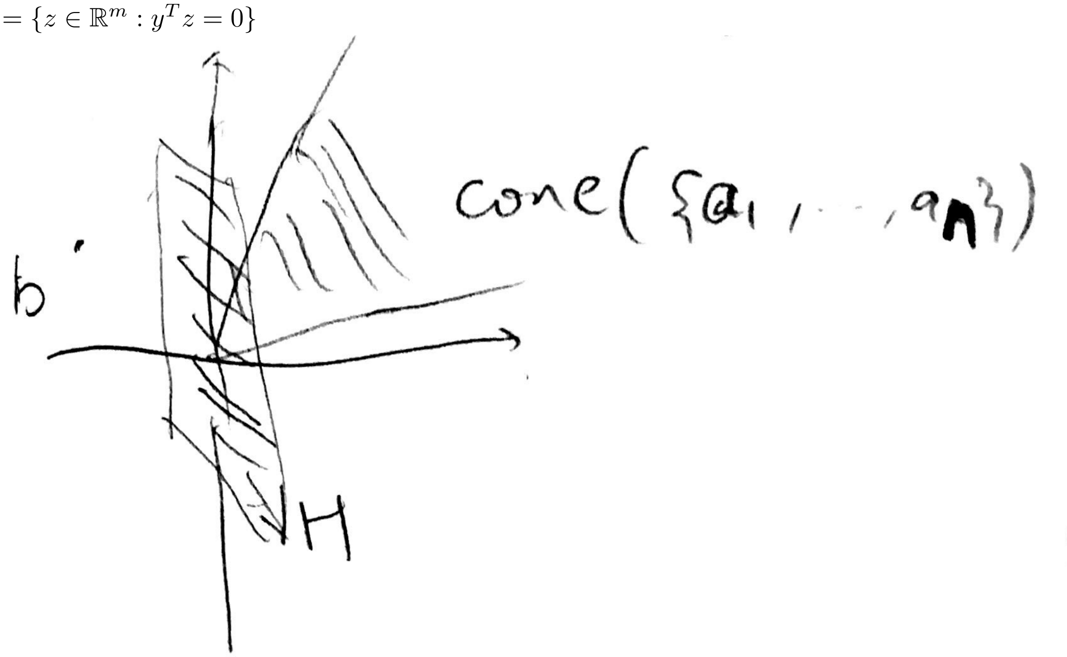
The following are equivalent:

(i) $b \notin \text{cone}(\{a_1, \cdots, a_n\})$

(ii) $(Ax = b, x \geq 0)$ has no solution

(iii) $(y^T A \geq 0, y^T b = -1)$ has a solution

Let $H = \{z \in \mathbb{R}^m : y^T z = 0\}$




$H$ is a hyperplane that separates $b$ from $\text{cone}(\{a_1, \cdots, a_n\})$

**Exercise:** prove the Farkas Lemma from Theorem 1.2

- Recall that an infeasible system of linear equations in $n$ variables contains an infeasible subsystem of size at most $n+1$.

## 1.6 Minimally Infeasible Subsystems

**Theorem 1.3**

Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. If the system $A\mathbf{x} \geq \mathbf{b}$ is infeasible, then there is an infeasible subsystem with at most $n+1$ inequalities.

**Lemma 1.4**

Let $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$. If $(A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0})$ has a solution, then there is a solution $x^*$ with at most $m$ non-zero entries.

**Proof** Assignment 1, Problem 5 □

**Geometric Interpretation:**

If $b, a_1, \ldots, a_n \in \mathbb{R}^n$ and $b \in \text{cone}(\{a_1, \ldots, a_n\})$, then there exists $X \subseteq \{a_1, \ldots, a_n\}$ with $|x| \leq m$ such that $b \in \text{cone}(X)$

**Proof of Theorem 1.3**
By the Farkas Lemma, if $A\mathbf{x} \geq \mathbf{b}$ is infeasible, then there exists $\mathbf{y} \in \mathbb{R}^m$ such that

$$\mathbf{y}^T A = \mathbf{0}, \mathbf{y}^T \mathbf{b} = 1, \quad \mathbf{y} \geq 0$$

That is $[A, \mathbf{b}]^T \mathbf{y} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}, \qquad \mathbf{y} \geq 0$
By Lemma 1.4, there is a solution $y^*$ with at most $n+1$ non-zero entries.
We get the result by taking the inequalities in $A\mathbf{x} \geq \mathbf{b}$ indexed by the support of $y^*$ □

# 2 Linear Programming

**Theorem 2.1 (Fundamental Theorem of Linear Programming)**

A linear program is either

- infeasible
- unbounded, or
- has an optimal solution

We'll prove it later.
Note that $\min(\frac{1}{x} : x \geq 1)$ is feasible and bounded but it has no optimal solution.
Consider a linear program
(LP) $\min(c^T x : Ax \geq b)$ where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$

- Let $P = \{x \in \mathbb{R}^n : Ax \geq b\}$

**Lemma 2.2**

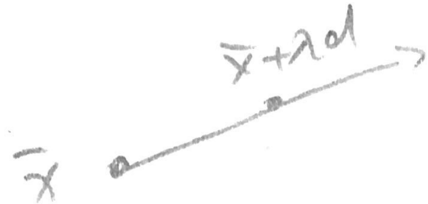
Let $\overline{x} \in P$ and $d \in \mathbb{R}^n$. Then

(i) $\{\overline{x} + \lambda d : \lambda \geq 0\} \subseteq P$ if and only if $Ad \geq 0$, and

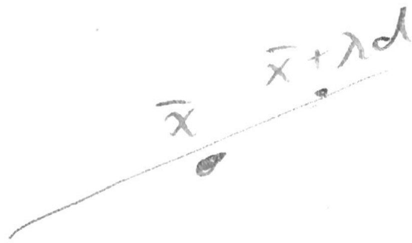
(ii) $\{\overline{x} + \lambda d : \lambda \in \mathbb{R}\} \subseteq P$ if and only if $Ad = 0$

**Proof:** see Assignment 1 □

$\{\bar{x} + \lambda d : \lambda \geq 0\}$ is a <u>ray</u>




$\{\bar{x} + \lambda d : \lambda \in \mathbb{R}\}$ is a line




**Theorem 2.3 (Unboundedness Theorem)**

(LP) is unbounded if and only if (LP) is feasible, and there exists $d \in \mathbb{R}^n$ such that $Ad \geq 0$ and $c^T d < 0$
(That is, P contains a ray $R$ and $\min(c^T x : x \in \mathbb{R})$ is unbounded)

We'll prove it later.

<u>**Claim (Easy direction):**</u>
If (LP) is feasible and $d \in \mathbb{R}^n$ satisfies $(Ad \geq 0, c^T d < 0)$, then (LP) is unbounded.

<u>**Proof:**</u> Let $\overline{x} \in P$ and let $\lambda \geq 0$. By Lemma 2.2 (i), $\overline{x} + \lambda d \in P$. Moreover $c^T(\overline{x} + \lambda d) = c^T\overline{x} + \lambda c^T d$. So $\lim\limits_{\lambda \to \infty} c^T(\overline{x} + \lambda d) = -\infty$ □

## 2.1 Duality

Question: How do we show that an (LP) is bounded?
Answer: use implied inequalities.

**Example**

$$\text{(LP)} \begin{cases} \min(x_1 + x_2) & \text{subject to} & \\ 2x_1 + x_2 \geq 4 & & (a) \\ 2x_1 + x_2 \geq 6 & & (b) \\ x_1 + 4x_2 \geq 4 & & (c) \end{cases}$$

$\frac{3}{7}(a) + \frac{1}{7}(c):\qquad x_1 + x_2 \geq \dfrac{16}{7}$

Hence OPT(LP) $\geq \dfrac{16}{7}$

**Question:** what is the best lower bound that we can get by using implied inequalities?
Each implied inequality has the form:

$$(2y_1 + 2y_2 + y_3)x_1 + (y_1 + 3y_2 + 4y_3)x_2 \geq 4y_1 + 6y_2 + 4y_3 \qquad \text{where } y_1, y_2, y_3 \geq 0$$

we multiply $y_1$ on both sides of $(a)$, $y_2$ on $(b)$, $y_3$ on $(c)$

If $\begin{array}{ll} 2y_1 + 2y_2 + y_3 & = 1 \\ y_1 + 3y_2 + 4y_3 & = 1 \end{array}$, then we have the inequality $\qquad x_1 + x_2 \geq 4y_1 + 6y_2 + 4y_3$

So to get the best lower bound, we want to solve

$$(D)\begin{cases} \max(4y_1 + 6y_2 + 4y_3) \qquad \text{subject to} \\ 2y_1 + 2y_2 + y_3 = 1 \\ y_1 + 3y_2 + 4y_3 = 1 \\ y_1, y_1, y_3 \geq 0 \end{cases}$$

By construction, OPT(D) $\leq$ OPT(LP)

Note that $x^* = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix}$ is a feasible solution to (LP) with objective value $\dfrac{5}{2}$, and $y^* = \begin{bmatrix} 1/4 \\ 1/4 \\ 0 \end{bmatrix}$ is a feasible solution to (D) with objective value $\dfrac{5}{2}$. Hence $x^*$ is optimal for (LP) and OPT(LP) $= \frac{5}{2}$.

More generally, consider

$$\text{(P)} \qquad \min(c^T x : Ax \geq b) \qquad \text{where } A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m \text{ and } c \in \mathbb{R}^n$$

The dual of (P)

$$\text{(D)} \qquad \max(b^T y : A^T y = c, y \geq 0)$$

**Weak Duality Theorem**

If $x$ is a feasible solution for (P) and $y$ is a feasible solution for (D), then $c^T x \geq b^T y$

**Proof:** $c^T x = (A^T y)^T x = y^T (Ax) \geq y^T b = b^T y$ □

**Corollary 2.4:** If (D) is feasible, then (P) is bounded.

**Proof:** Immediate □

**Corollary 2.5:** If (P) is feasible, then (D) is bounded.

**Proof:** Immediate □

**Corollary 2.6:** If $x$ is feasible for (P), $y$ is feasible for (D), and $c^T x = b^T y$, then $x$ is optimal for (P) and $y$ is optimal for (D).

**Proof:** Immediate □

**Strong Duality Theorem**

If (P) has an optimal solution, then (D) has an optimal solution and OPT(P) = OPT(D)

**Proof:** Later □

| (P) \ (D) | infeasible | unbounded | has optimal solution |
|---|---|---|---|
| infeasible | ✓ | ✓ | ✗ |
| unbounded | ✓ | ✗ | ✗ |
| has optimal solution | ✗ | ✗ | ✓ |

- Strong Duality Theorem
- Corollary 2.4
- Corollary 2.5
- Use the Farkas Lemma (Assignment 2)
- this can happen (example Assignment 3)

**Theorem 2.7 (LP Uber Theorem)**

Either

(I) (P) and (D) both have optimal solutions and OPT(P) = OPT(D), or

(II) There exists $y \in \mathbb{R}^m$ such that $(y^T A = 0, y^T b = 1, y \geq 0)$ and hence (P) is infeasible, or

(III) (P) is feasible and there exists $d \in \mathbb{R}^n$ such that $(c^T d < 0, Ad \geq 0)$ and hence (PP) is unbounded

**Remark:** this implies

- Fundamental Theorem
- the Unboundedness Theorem, and
- the Strong Duality Theorem

We'll assume that neither (I) nor (II) hold, and will show that (III) holds.
By the Farkas' Lemma, since (II) does not hold, (P) is feasible. We can rewrite (I) as
(A) $\quad (Ax \geq b, A^T y = c, y \geq 0, b^T y \geq c^T x)$

Consider the system:
(B) $\quad (A^T y = zc, Ax \geq zb, \quad b^T y - c^T x = 1, y \geq 0, z \geq 0)$

<u>**Claim 1**</u> Exactly one of (A) and (B) has a solution.

**Proof** Exercise □

**Claim 2** $\overline{z} = 0$ $(\overline{z} \in \mathbb{R})$

**Proof** Suppose that $\overline{z} > 0$. Let $\begin{cases} x' = \frac{1}{\overline{z}} \cdot \overline{x} \\ y' = \frac{1}{\overline{z}} \cdot \overline{y} \end{cases}$

Then

$$A^T y' = c, Ax' \geq b, \quad b^T y' - c^T x' = \frac{1}{\overline{z}}(b^T \overline{y} - c^T \overline{x}) = \frac{1}{\overline{z}} > 0, \quad y' \geq 0$$

Thus $x', y'$ satisfies (A), contrary to **Claim 1** □

Since $b^T \overline{y} - c^T \overline{x} = 1$, either $b^T \overline{y} > 0$ or $c^T \overline{x} < 0$

case 1: $b^T \overline{y} > 0$
Thus $A^T \overline{y} = 0, b^T \overline{y} > 0, \overline{y} \geq 0$
We can scale $\overline{y}$ to obtain a solution to $(A^T y = 0, b^T y = 1, y \geq 0)$ and hence (III) holds
**contradiction**

case 2: $c^T \overline{x} < 0$
Then $A\overline{x} \geq 0$, and $c^T \overline{x} < 0$
So $d = \overline{x}$ satisfies (III)
This proves Theorem 2.7 □

## 2.2 Complementary Slackness Conditions

Consider

$$\begin{array}{ll} (P) & \min(c^T x : Ax \geq b) \\ (D) & \max(b^T y : A^T y = c, y \geq 0) \end{array}$$

Let $a_1{}^T, \ldots, a_m{}^T$ be the rows of $A$
If $\overline{x}$ is feasible for (P) and $\overline{y}$ is feasible for (D)

$$\begin{aligned} c^T \overline{x} - b^T \overline{y} &= (A^T \overline{y})^T \overline{x} - \overline{y}^T b \\ &= \overline{y}^T (A\overline{x} - b) \\ &= \sum_{i=0}^{m} \underbrace{\overline{y}_i}_{\geq 0} \underbrace{(a_i{}^T - b_i)}_{\geq 0} \geq 0 \end{aligned} \tag{2.1}$$

Moreover, equality holds if and only if

(∗) for each $i \in \{1, \ldots, m\}$, either

$$a_i{}^T = b_i \qquad \text{or} \qquad \overline{y}_i = 0$$

(That is, at most one of the inequalities $a_i{}^T \geq b_i$ and $\overline{y}_i \geq 0$ is strict)

We call (∗) the complementary slackness conditions.

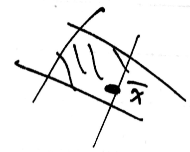
**Theorem 2.8**

Let $\overline{x}$ be a feasible solution for (D). Then $c^T \overline{x} = b^T \overline{y}$ if and only if for each $i \in \{1, \ldots, m\}$ either $a_i{}^T = b_i$ or $\overline{y}_i = 0$

**Proof** See above □

## 2.3 Certifying Optimality

An equality $a_i{}^T \geq b_i$ is an <u>equality constant</u> for $\overline{x}$ if $a_i{}^T = b_i$ and the set of all equality constraints is called the <u>equality subsystem</u> for $\overline{x}$




**Theorem 2.9**

Let $\overline{x}$ be a feasible solution for

$$\min(c^T x : Ax \geq b)$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^n$, and $c \in \mathbb{R}^n$, and let $A^= x \geq b^=$ be the equality subsystem for $\overline{x}$. Then there is an optional solution if and only if there is a non-negative vectors $y$ such that

$$c = (A^=)^T y$$

**Proof** Follows immediately from **Theorem 2.8** □

**Remark:** If $a_1, \ldots, a_n$ are the rows of $A^=$, then the following are equivalent:

(a) there is a non-negative vector $y$ such that $c = (A^=)^T y$, and

(b) $c \in \text{cone}(\{a_1, \ldots, a_k\})$

## 2.4 Cost Splitting

For each $i \in \{1, \ldots, m\}$. Let $P_1 = \{x \in \mathbb{R}^n : a_i^T x \geq b_i\}$. Then (P) can be rewritten as:

$$\min(c^T x : x \in P_1 \bigcap \ldots \bigcap P_m)$$

**Theorem 2.10 (Weak Cost-Splitting Theorem)**

Let $S_1, \ldots, S_m \subseteq \mathbb{R}^n$, $c \in \mathbb{R}^n$, and $\overline{x} \in S_1 \bigcap \ldots \bigcap S_m$. If there exist $c_1, \ldots, c_m \in \mathbb{R}^n$, such that $c = c_1 + \ldots + c_m$ and minimizes $(c_i^T x : x \in S_i)$, for each $i \in \{1, \ldots, m\}$, then $\overline{x}$ minimizes $(c^T x : x \in S_1 \bigcap \ldots \bigcap S_m)$.

**Proof** Let $\widetilde{x} \in S_1 \bigcap \ldots \bigcap S_m$
For each $i \in \{1, \ldots, m\}$, we have $c_i^T \widetilde{x} \geq c_i^T \overline{x}$.
So

$$\begin{aligned}
c^T \widetilde{x} &= (c_1^T + \ldots + c_m^T)\widetilde{x} \\
&= c_1^T \widetilde{x} + \ldots + c_m^T \widetilde{x} \\
&\geq c_1^T \overline{x} + \ldots + c_m^T \overline{x} \\
&= (c_1^T + \ldots + c_m^T)\overline{x} \\
&= c^T \overline{x}
\end{aligned}$$

Hence, $\overline{x}$ miminizes $(c^T x : x \in S_1 \bigcap \ldots \bigcap S_m)$ □

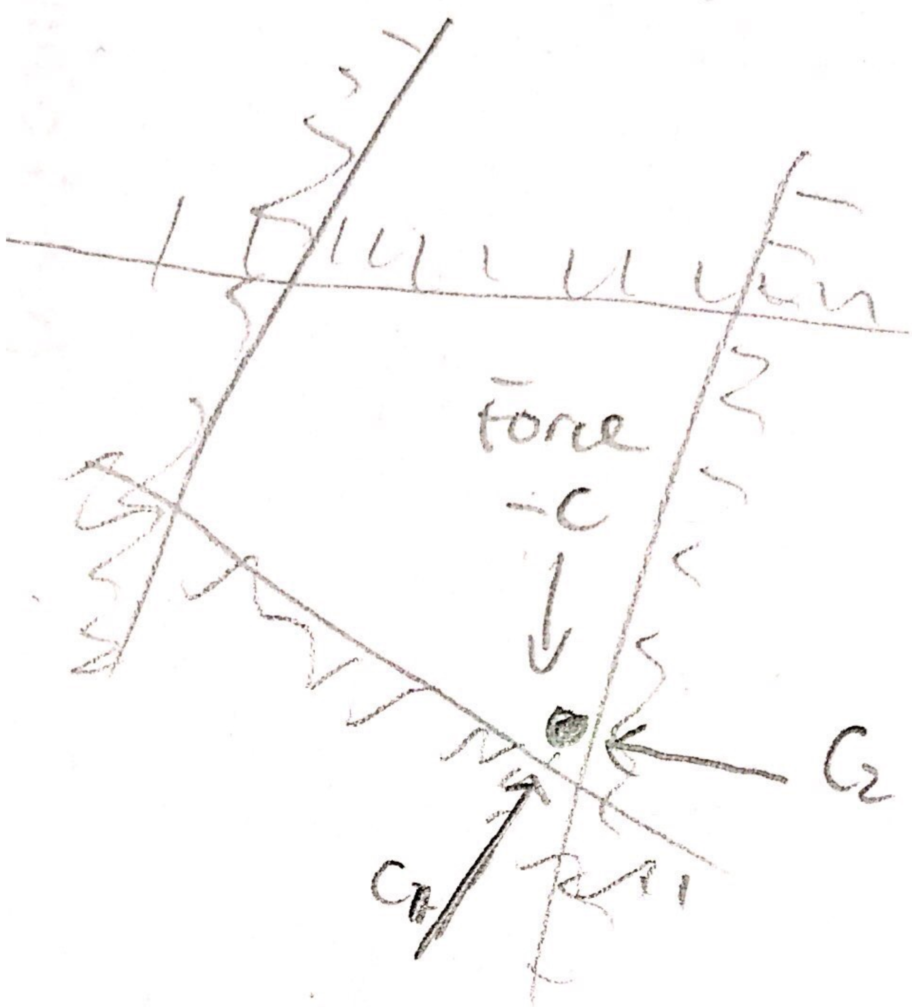
**Theorem 2.11 (Strong Cost-Splitting Theorem for Linear Programming)**

If $\overline{x}$ minimizes $(c^T x : x \in S_1 \bigcap \ldots \bigcap S_m)$, then there exists $c_1, \ldots, c_m \in \mathbb{R}^n$ such that $c = c_1 + \ldots + c_m$ and $\overline{x}$ minimizes $(c_i^T x : x \in P_i)$ for each $i \in \{1, \ldots, m\}$.

**Proof** Exercise □

**Economic Interpretation** Cost-Splitting has an economic interpretation, the cost $c^T \overline{x}$ of $\overline{x}$ can be divided up as $c_1^T \overline{x}, \ldots, c_m^T \overline{x}$ and apportioned to the constraints.

**Physical Interpretation** Consider an optional solution $\overline{x}$ and a cost-splitting $c = c_1 + \ldots + c_m$ given by by Theorem 2.11




- Newton's Third Law: for every action, there is an equal and opposite reaction.

## 2.5 Duality (other forms)

**Example**

$$(P')\begin{cases}\max\ c^Tx & \text{subject to} & \\ Ax \le b & & y \ge 0\\ x \ge 0 & & \end{cases}$$

$$(D')\begin{cases}\max\ b^Ty & \text{subject to} & \\ A^Ty \ge c & & x \ge 0\\ y \ge 0 & & \end{cases}$$

If $x$ is feasible for $(P')$ and $y$ is feasible for $(D')$, then

$$\begin{aligned} c^Tx &\le (A^Ty)^Tx \\ &= y^T(Ax) \\ &\le y^Tb \\ &= b^Ty \end{aligned}$$

Hence $\text{OPT}(P') \le \text{OPT}(D')$

**Theorem 2.12**

If $(P')$ has an optimal solution, then $(D')$ has an optimal solution and $\text{OPT}(P') = \text{OPT}(D')$

**Proof** We can rewrite $(P')$ as

$$(P'')\begin{cases}\min\ -c^Tx & \text{subject to} & \\ -Ax \ge -b & & y \ge 0\\ Ix \ge 0 & & s \ge 0\end{cases}$$

The dual of $(P'')$ is

$$(D'')\begin{cases}\max\ -b^Ty & \text{subject to}\\ -A^Ty + s = -c & \\ y,\ s \ge 0 & \end{cases}$$

If $\overline{x}$ is an optimal solution to $(P')$, then $\overline{x}$ is also optimal solution to $(P'')$. By the String Duality Theorem, there is an optimal solution $\overline{y},\ \overline{s}$ to $(D'')$ and $\text{OPT}(D'') = \text{OPT}(P'') = -\text{OPT}(P')$. Note that $\overline{y}$ is feasible for $(D')$ and

$$\text{OPT}(P') \le b^T\overline{y} = -\text{OPT}(D'') = \text{OPT}(P')$$

Hence $\text{OPT}(P') = \text{OPT}(D')$ □

**Another Ex**

$$(P)\begin{cases}\max\ 3x_1 - x_2 + x_3 & \text{subject to} & \\ 2x_1 + 2x_2 = 4 & & y_1\\ x_1 - 2x_2 + 2x_3 \le 3 & & y_2 \ge 0\\ x_1,\quad x_3 \ge 0 & & \end{cases}$$

Consider the implied inequality

$$(2y_1 + y_2)x_1 + (2y_1 - y_2)x_2 + 2y_2x_3 \le 4y_1 + 3y_2$$

We want

$$3x_1 - x_2 + x_3 \le (2y_1 + y_2)x_1 + (2y_1 - 2y_2)x_2 + 2y_2x_3$$

for all $x_1, x_2, x_3 \in \mathbb{R}$ with $x_1, x_3 \geq 0$

That is

$$\begin{cases} 2y_1 + y_2 \geq 3 \\ 2y_1 - 2y_2 = -1 \\ 2y_2 \geq 1 \end{cases}$$

The dual of (P) is

$$(D) \begin{cases} \min\ 4y_1 + 3y_2 & \text{subject to} & \\ 2y_1 + y_2 \geq 3 & & x_1 \geq 0 \\ 2y_1 - 2y_2 = -1 & & x_2 \\ 2y_2 \geq 1 & & x_3 \geq 0 \\ y_2 \geq 0 & & \end{cases}$$

## 2.6 Cheat Sheet

| (P) max | (D) min |
|---|---|
| $\leq$ constraint | non-negative |
| $\geq$ constraint | non-positive |
| $=$ constraint | free variable |
| non-negative variables | $\geq$ constraint |
| non-positive variables | $\leq$ constraint |
| free variable | $=$ constraint |

Note that we have variables on the left of the inequalities.

# 3 Geometry of Polyhedra

Recall

A <u>polyhedron</u> is a set of the form $\{x \in \mathbb{R}^n : Ax \geq b\}$. A <u>polytope</u> is a bounded polyhedron.

We'll prove:

**Theorem 3.1**

A set $P \subseteq \mathbb{R}^n$ is a polytope if and only if $P = \text{conv}(X)$ for some finite set $X \subseteq \mathbb{R}^n$

<u>Definition:</u> For sets $A, B \subseteq \mathbb{R}^n$, we let $A + B = \{a + b : a \in A, b \in B\}$

**Theorem 3.2**

A set $P \subseteq \mathbb{R}^n$ is a polyhedron *if and only if* there exist finite sets $X, D \subseteq \mathbb{R}^n$ such that

$$P = \text{conv}(X) + \text{cone}(D)$$


$D = \{d_1, d_2\}$

$P = \text{conv}(X) + \text{cone}(D)$

We start by proving that conv($X$) is a polytope:

**Lemma 3.3**

If $a_1, \ldots, a_k \in \mathbb{R}^n$, then

$$\mathrm{conv}(\{a_1, \ldots, a_k\}) = \{\lambda_1 a_1 + \ldots + \lambda_k a_k : \lambda \in \mathbb{R}^n, \lambda \geq 0, \lambda_1 + \ldots + \lambda_k = 1\}$$

**Theorem 3.4**

If $a_1, \ldots, a_k \in \mathbb{R}^n$, then $\mathrm{conv}(\{a_1, \ldots, a_k\})$ is a polytope.

**Proof**
Since $\mathrm{conv}(\{a_1, \ldots, a_k\})$ is bounded, it suffices to show that $\mathrm{conv}(\{a_1, \ldots, a_k\})$ is a polyhedron.
Let $P_0 = \left\{ \begin{pmatrix} x \\ \lambda \end{pmatrix} : x \in \mathbb{R}^n, \lambda \in \mathbb{R}^k, \lambda \geq 0, \lambda_1 + \ldots + \lambda_k = 1, x = \lambda_1 a_1 + \ldots + \lambda_k a_k \right\}$

By definition, $P_0$ is a polyhedron, and by Lemma 3.3, $\mathrm{conv}(\{a_1, \ldots, a_k\})$ is the projection of $P_0$ onto $x$. Then, by Theorem 1.1, $\mathrm{conv}(\{a_1, \ldots, a_k\})$ is a polyhedron. $\square$

## 3.1 Extreme Points


Let $S \subseteq \mathbb{R}^n$ be a convex set and $x \in S$, we call $x \in S$ an extreme point of $S$ if there are no two distinct points $x_1, x_2$ in $S$ such that

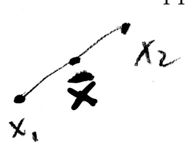
$$x \in \{\lambda x_1 + (1 - \lambda) x_2 : 0 < \lambda < 1\}$$

Equivalently, $S \setminus \{x\}$ is convex

**Theorem 3.5**

Let $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Let $\overline{x} \in P$, and let $A^= x \geq b^=$ be the subsystem for $\overline{x}$, then $\overline{x}$ is an extreme point if and only if $\mathrm{rank}(A^=) = n$

**Proof** Suppose that $\mathrm{rank}(A^=) < n$, and let $d \in \mathbb{R}^n$ be a non-zero vector such that $A^= d = 0$




Note that $A^=(\overline{x} + \lambda d) = A^= \overline{x} + \lambda A^= d = b^=$ for all $\lambda \in \mathbb{R}$. Then there exists $\varepsilon > 0$ such that $\overline{x} + \varepsilon d, \overline{x} - \varepsilon d \in P$, and hence $\overline{x}$ is not an extreme point.

Conversely, suppose that $\overline{x}$ is not an extreme point, then there exist distinct $x_1, x_2 \in P$ and $\lambda \in (0, 1)$ such that $\overline{x} = \lambda x_1 + (1 - \lambda) x_2$

Note that

$$\begin{aligned} b^= &= A^= \overline{x} \\ &= \lambda A^= x_1 + (1-\lambda) A^= x_2 \\ &\geq \lambda b^= + (1-\lambda) b^= \\ &= b^= \end{aligned}$$

Then, since $0 < \lambda < 1$

$$A^= x_1 = b^= \text{ and } A^= x_2 = b^=$$

Thus $A^=(x_2 - x_1) = 0$ and hence $\text{rank}(A^=) < n$ □

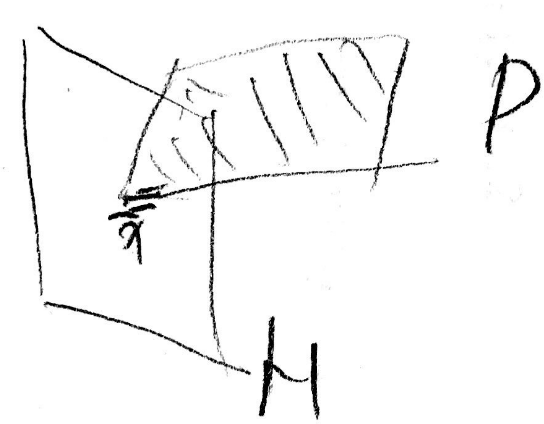
**Remark:** $\overline{x}$ is the unique solution to $A^= x = b^=$ if and only if $\text{rank}(A^=) = n$

**Corollary 3.6**

Polyhedra have only finitely many extreme points.

**Proof** Let $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Now consider an extreme point $\overline{x}$ and its associated equality subsystem $A^= x \geq b^=$. By Theorem 3.5, $\text{rank}(A^=) = n$. Therefor $\overline{x}$ is the solution to $A^= x = b^=$. There are only $2^m$ subsystems of $Ax \geq b$, so there are at most $2^m$ extreme points. □

## 3.2 Supporting Hyperplanes




A hyperplane of $\mathbb{R}^n$ is a set of the form

$$\{x \in \mathbb{R}^n : a^T x = a_0\}$$

where $a \in \mathbb{R}^n \setminus \{0\}$ and $a_0 \in \mathbb{R}$

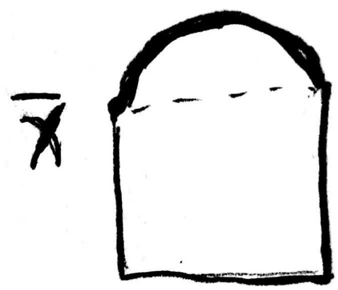
A supporting hyperplane for a set $S \subseteq \mathbb{R}^n$ is a hyperplane $H = \{x \in \mathbb{R}^n : a^T x = a_0\}$ such that

(i) $S$ is contained in either $\{x \in \mathbb{R}^n : a^T x \geq a_0\}$ or $\{x \in \mathbb{R}^n : a^T x \leq a_0\}$, and

(ii) $H \bigcap S \neq \varnothing$

Note that: if $H$ is a supporting hyperplane for a convex set $S \subseteq \mathbb{R}^n$ and $H \bigcap S = \{\overline{x}\}$, then $\overline{x}$ is an extreme point.

In general, the converse may not hold.




**Example**

$$S = \{x \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, -1 \leq x_2 \leq 0\} \cup \{x \in \mathbb{R}^2 : (x_1 - 1)^2 + x_2^2 \leq 1\}$$

Then $\overline{x} = 0$ is an extreme point, but there is no supporting hyperplane $H$ with $H \bigcup S = \{\overline{x}\}$

**Theorem 3.7**

If $\overline{x}$ is an extreme point of a polyhedron $P \subseteq \mathbb{R}^n$, then there is a supporting hyperplane such that $P \bigcap H = \{\overline{x}\}$.

**Proof** Suppose that $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ and let $\begin{cases} a_1^T x \geq b_1 \\ \vdots \\ a_k^T x \geq b_k \end{cases}$ be the equality subsystem for $\overline{x}$

Let $a = a_1 + \ldots + a_k, a_0 = b_1 + \ldots + b_k$, and $H = \{x \in \mathbb{R}^n : a^T x = a_0\}$

Note that $H$ is a supporting hyperplane for $P$.

Consider $\widetilde{x} \in P \bigcap H$

Thus

$$\begin{aligned} a_0 &= a^T \widetilde{x} \\ &= a_1^T \widetilde{x} + \ldots + a_k^T \widetilde{x} \\ &\geq b_1 + \ldots + b_k \\ &= a_0 \end{aligned}$$

Therefore $a_1^T \widetilde{x} = b_1, \ \ldots, \ a_k^T \widetilde{x} = b_k$

However, by Theorem 3.5, $\overline{x}$ is the unique solution to $\begin{cases} a_1^T x \geq b_1 \\ \vdots \\ a_k^T x \geq b_k \end{cases}$

So $H \bigcap P = \{\overline{x}\}$. $\square$

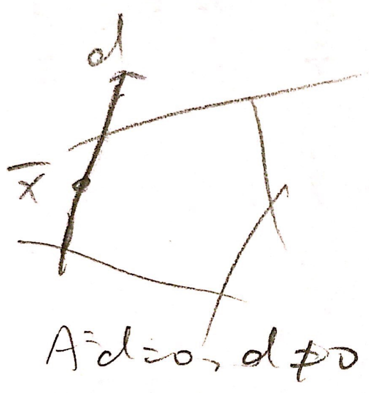
**Theorem 3.8**

Every polytope is the convex hull of its extreme points.

**Proof** Let $X$ be the set of extreme points of a polytope. $P = \{x \in \mathbb{R}^n : Ax \geq b\}$. Let $\overline{x} \in P$ and $A^= x \geq b^=$ be the equality subsystem for $\overline{x}$. Since $X \subseteq \text{conv}(X)$, we may assume $\overline{x} \notin X$ and, hence, $\text{rank}(A^=) < n$. However by Lemma 2.2, $\text{rank}(A) = n$ (Since $P$ is bounded). So $A \neq A^=$. We may assume that each point in $P$ that has more equality constraints than $\overline{x}$ is contained in $\text{conv}(X)$.
Since $\text{rank}(A^=) < n$, there is a non-zero vector $d$ such that $A^= d = 0$. Let

$$\lambda^+ = \max(\lambda \in \mathbb{R} : \overline{x} + \lambda d \in P)$$
$$\lambda^- = \min(\lambda \in \mathbb{R} : \overline{x} + \lambda d \in P)$$

Note that these exist since $P$ is closed and bounded
Since $A^= d = 0$, we have $A^=(\overline{x} + \lambda^- d) = b^=$ and $A^=(\overline{x} + \lambda^+ d) = b^=$. Therefore, by our choice $\lambda^+$ and $\lambda^-$, we have $\lambda^+ < 0 < \lambda^-$ and $\overline{x} + \lambda^+ d$ and $\overline{x} + \lambda^- d$ both have equality constraints that $\overline{x}$ does. By our choice of $\overline{x}$, $\overline{x} + \lambda^- d, \overline{x} + \lambda^+ d \in \text{conv}(X)$. Then since $\overline{x}$ is on the line segment between $\overline{x} + \lambda^+ d$ and $\overline{x} + \lambda^- d$, we have $\overline{x} \in \text{conv}(X)$ □




Note that Theorem 3.1 is implied by Theorem 3.4 and 3.7 and Corollary 3.6

### 3.2.1 Application (Helly's Theorem)

Recall

> **Theorem 1.3**
>
> Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$. If $Ax \geq b$ is infeasible, then there is an infeasible subsystem with at most $n + 1$ constraints.

**Equivalently** If $H_1, \ldots, H_m \subseteq \mathbb{R}^n$ are closed half-spaces with $H_1 \bigcap \ldots \bigcap H_m = \varnothing$, then there is a subcollection of at most $n + 1$ of these half-spaces that have empty intersection.

> **Corollary 3.9**
>
> If $P_1, \ldots, P_m \subseteq \mathbb{R}^n$, are polyhedra with $P_1 \bigcap \ldots \bigcap P_m = \varnothing$, then there is a subcollection of at most $n + 1$ of these polyhedra that have empty intersection.

**Proof** Each $P_i$ is itself or intersection of closed half-spaces. □

**Question** Does that hold if we allow infinitely many polyhedra?

Let $P_i = \{x \geq i\} \quad i = 1, 2, 3 \ldots$
$P_1 \bigcap P_2 \bigcap \ldots = \varnothing$, but each finite subcollection has non-empty intersection

**Theorem 3.10 (Helly's Theorem)**

If $S_1, \ldots, S_m \subseteq R^n$ are convex sets with $S_1 \bigcap \ldots \bigcap S_m = \varnothing$, then there is a subcollection of at most $n+1$ of these sets has empty intersection.

**Proof** Suppose otherwise. There is a set $X$ with $|X| \leq \binom{m}{n+1}$ such that each subcollection of $n+1$ of the sets contains an element of $X$ in its intersection. Let $P_i = \text{conv}(X \bigcap S_i)$. By Theorem 3.1, $P_1, \ldots, P_m$ are polyhedra. By construction, $P_1 \bigcap \ldots \bigcap P_m = \varnothing$, but the intersection of any $n+1$ $P_1, \ldots, P_m$ is non-empty. Contrary to Corollary 3.9. □

**Theorem 3.11**

If $X$ and $D$ are finite subsets of $\mathbb{R}^n$, then $\text{conv}(X) + \text{cone}(D)$ is a polyhedra.

**Proof** Exercise □

In the following results $D = \{x \in \mathbb{R}^n : Ax \geq b\}$ where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$

## 3.3 Pointed Polyhedra

**Pointed Polyhedra**

A polyhedra is <u>pointed</u> if it is nonempty and contains no line.

For a subspace $S \subseteq \mathbb{R}^n$ we define $S^\perp = \{y \in \mathbb{R}^n : y^T x = 0 \quad \text{for each } x \in S\}$

**Lemma 3.12**

Let $S = \{x \in \mathbb{R}^n : Ax = 0\}$ and let $P_0 = P \bigcap S^\perp$. If $P \neq \varnothing$, then $P_0$ is pointed and $P = P_0 + S$

**Proof** If $L = \{\widetilde{x} + \lambda \widetilde{d} : \lambda \in \mathbb{R}\}$ is a line in $P$, then, by Lemma 2.2, $\widetilde{d} \in S$.Hence $L$ is not contained in $P_0$. Hence $P_0$ is pointed. It remains to prove that we can write $x = z + d \quad z \in S^\perp$ and $d \in S$. By Lemma 2.2, $z \in P$, hence $z \in P_0$ and $X = P_0 + S$ □

**Lemma 3.13**

Let $X$ be the set of extreme points of $P$, and let $K = \{x \in \mathbb{R}^n : Ax \geq 0\}$. If $P$ is pointed, then $P = \text{conv}(X) + K$

**Proof** Exercise □

**Theorem 3.14**

A polyhedra is pointed if and only if it has an extreme point.

**Proof** Immediate by Lemma 3.13 □

## 3.4 Polyhedral Cones

**Exercise** Show that, if $P \subseteq R^n$ is both a polyhedra and a cone, then $P = \{x \in \mathbb{R}^n : Ax \geq 0\}$ where $A \in \mathbb{R}^{m \times n}$

**Theorem 3.15**

If $P \subseteq \mathbb{R}^n$ is a polyhedral cone, then $P = \text{cone}(D)$ for some finite set $D \subseteq P$

**Proof** Let $Q^1 = \{x \in P : -\mathbb{1} \leq x \leq \mathbb{1}\}$
So $Q$ is a polytope and $P = \text{cone}(Q)$. By Theorem 3.1, there is a finite set $D \subseteq R^n$ such that $Q = \text{conv}(D)$.
Now $P = \text{cone}(Q) = \text{cone}(\text{conv}(D)) = \text{cone}(D)$ since $\text{conv}(D) \subseteq \text{cone}(D)$ □

## 3.5 Proof of Theorem 3.2

The "if direction" was proved in Theorem 3.11. For the converse, consider a polyhedron $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Let $S = \{x \in \mathbb{R}^n : Ax = 0\}$ and $P_0 = P \bigcap S^\perp$. By Lemma 3.12, $P_0$ is pointed and $P = P_0 + S$. Let $d_1, \ldots, d_k$ be a basis for the subspace $S$ and let $D_0 = \{d_1, -d_1, \ldots, d_k, -d_k\}$. Then $S = \text{cone}(D_0)$ and hence $P = P_0 + \text{cone}(D_0)$

Since $P_0$ is a polyhedron, so we can write $P_0 = \{x \in \mathbb{R}^m : \widetilde{A}x \geq \widetilde{b}\}$ where $A \in \mathbb{R}^{\widetilde{m} \times n}$ and $\widetilde{b} \in \mathbb{R}^{\widetilde{m}}$. Let $X$ be the set of extreme points of $P_0$ and let $K = \{x \in \mathbb{R}^n : \widetilde{A}x \geq 0\}$. By Lemma 3.13, $P_0 = \text{conv}(X) + K$. By Theorem 3.15, $K = \text{cone}(D_1)$ for some finite set $D_1 \subseteq K$.

Therefore

$$\begin{aligned} P &= P_0 + \text{cone}(D_0) \\ &= \text{conv}(X) + K + \text{cone}(D_0) \\ &= \text{conv}X + \text{cone}(D_1) + \text{cone}(D_0) \\ &= \text{conv}(X) + \text{cone}(D_1 \bigcup D_0) \end{aligned}$$

□

[1] it's a box

# 4

# Algorithms for Linear Programming

Given $A \in \mathbb{Q}^{m\times n}, b \in \mathbb{Q}^m$, and $c \in \mathbb{Q}^n$, consider

$$(P)\begin{cases} \min c^T x & \text{subject to} \\ Ax \geq b \end{cases}$$

Let $P = \{x \in \mathbb{R}^n : Ax \geq b\}$

Feasibility Problem:
Find a feasible solution if it exists
Optimization Problem:
Given a feasible solution $\widetilde{x} \in \mathbb{Q}^n$, find a optimal solution if it exists.

The dual of (P) is

$$(D) \qquad \max(b^T y : A^T y = c, y \geq 0)$$

Note that (P) has an optimal solution if and only if the following system is feasible:

$$\begin{cases} Ax \geq b \\ A^T y = c \\ y \geq 0 \\ b^T y = c^T x \end{cases}$$

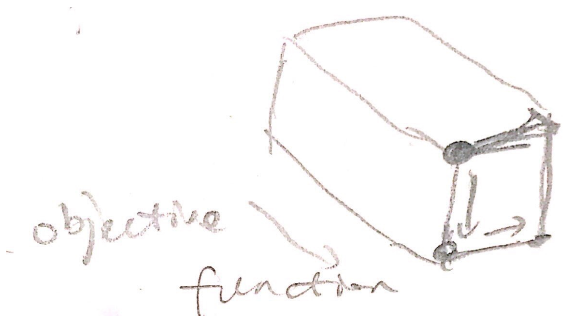
So the optimization problem reduces to the feasibility problem.

Consider the following auxiliary problem

$$(AP)\begin{cases} \min s_1 + \ldots + s_m & \text{subject to} \\ Ax + s \geq b \\ s \geq 0 \end{cases}$$

(AP) is feasible; take $\widetilde{x} = 0$ and $\widetilde{s}_j = \max(0, b_j)$ for $j \in \{1, \ldots, m\}$
Now (P) is feasible if and only if OPT(AP) $= 0$
This reduces the Feasibility Problem to the Optimization Problem. We will solve the Optimization Problem.

## 4.1 Simplex Method (Revised dual Simplex method and perturbation method)




**Idea:** Move from extreme point to extreme point around the boundary improving the objective value.

Finding an extreme point

### Problem 1

$P$ may not have an extreme point.
We are assuming that $P \neq \varnothing$
By Lemma 2.2 and Theorem 3.4, the following are equivalent

- $P$ has no extreme point,
- $P$ contains a line, and
- $\operatorname{rank}(A) \leq n$

Suppose that $\operatorname{rank}(A) < n$, and let $\mathbf{d}$ be a non-zero vector in $\mathbb{R}^n$ such that $Ad = 0$.
Let $\overline{x} \in P$ and consider the line $L = \{\overline{x} + \lambda d : \lambda \in \mathbb{R}\}$
By Lemma 2.2, $L \in P$.

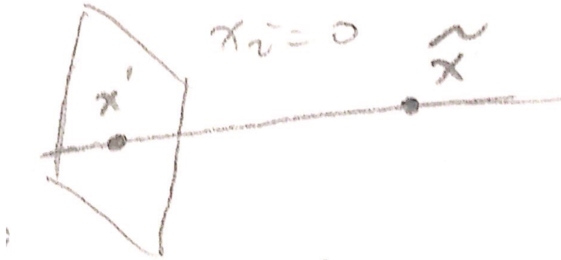
**Claim** If $c^T d \neq 0$, then (P) is unbounded.

**Proof** By replacing $d$ with $-d$ we may assume that $c^T d < 0$. Then by unboundedness theorem, (P) is unbounded. □

We may assume that $c^T d = 0$
Choose $i \in \{1, \ldots, \}$ such that $d_i \neq 0$

**Claim** For each $\widetilde{x} \in P$, there exists $x' \in P$ such that

$$c^T x' = c^T \widetilde{x} \qquad \text{and} \qquad x'_i = 0$$



**Proof** Let $\lambda = \frac{\widetilde{x}_i}{d_i}$, and let $x' = \widetilde{x} - \lambda d$. Since $c^T d = 0$, we have $c^T \widetilde{x} = c^T x'$. Moreover
$x'_i = \widetilde{x}_i - \dfrac{\widetilde{x}_i}{d_i} d_i = 0$ □

Let (P') be the problem obtained from (P) by setting $x_i = 0$. Now (P') has fewer variables than (P) and, by the claim, $\text{OPT}(P) = \text{OPT}(P')$.

Hence we'll assume that $\text{rank}(A) = n$ and hence that $P$ has an extreme point.

## Problem 2

Given $\widetilde{x} \in P$, find an extreme point of $P$.

### Algorithm

**Step1** Construct the equality subsystem $A^= x \geq b^=$ for $\widetilde{x}$. If $\text{rank}(A^=) = n$. STOP ($\widetilde{x}$ is an extreme point)

**Step2** Find a non-zero vector $d \in \mathbb{R}^n$ such that $A^= d = 0$. If $Ad \geq 0$, replace $d$ with $-d$.

**Step3** Let $\lambda^- = \max(\lambda \in \mathbb{R} : x + \lambda d \in P)$. Replace $\widetilde{x}$ with $\widetilde{x} + \lambda^- d$. Repeat from Step1.

**Exercise** Show that the algorithms works.

## Problem 3

Given an extreme point $\widetilde{x} \in P$, solve (P).

$\widetilde{x}$ is optimal if and only if there exists $y$ satisfies

$$(c = (A^=)^T y, y \geq 0),$$

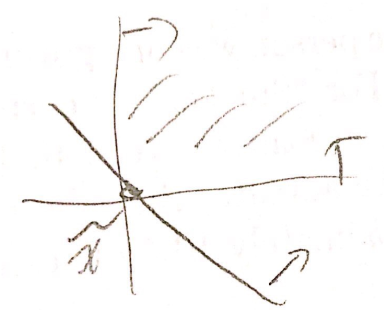
where $A^= x \geq b^=$ is the equality subsystem for $\widetilde{x}$

Let $A^= x \geq b^=$ be the equality subsystem for $\widetilde{x}$
Let $a_1^T, \ldots, a_m^T$ denote the rows
Let $(B, N)$ be the partition of $(1, \ldots, m)$ such that $a_i^T \widetilde{x} = b_i$ if and only if $i \in B$.

By Theorem 2.9, $\widetilde{x}$ is optimal if and only if there exists $y \in \mathbb{R}^n$ satisfying

$$(*) \qquad (c = (A^=)^T y, \quad y \geq 0)$$

**Remark** Since $\text{rank}(A^=) = n$, $\widetilde{x}$ is the unique solution to $A^= \widetilde{x} = b^=$. If $A^=$ has more than $n$ rows, then $A^= x = b^=$ is overdetermined. In this cases, we call $\widetilde{x}$ **degenerate**.



Assume that $\widetilde{x}$ is non-degenerate.

Thus $A^=$ is square and non-singular. Therefore there is a unique solution $\overline{y}$ to $(A^=)^T y = c$. By $(*)$, $\widetilde{x}$ is optimal if and only if $\overline{y} \geq 0$.

Suppose otherwise and choose $j \in B$ such that $\overline{y}_j < 0$

Define $e_j \in \mathbb{R}^n$ such that $e_j = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, where 1 denotes the $j^{\text{th}}$ row.

Let $\overline{d}$ denote the unique solution to $A^= d = e_j$

**Claim 1** $c^T \overline{d} < 0$

**Proof** $c^T \overline{d} = (\overline{y}^T A^=)\overline{d} = \overline{y}^T e_j = \overline{y}_i < 0$ □

**Claim 2** For sufficient small $\varepsilon > 0$, $\overline{x} + \varepsilon\, \overline{d} \in P$

**Proof** It suffices to prove that for each $i \in \{1, \ldots, m\}$, $a_i^T(\widetilde{x} + \varepsilon\, \overline{d}) \geq b_i$ for sufficiently small $\varepsilon > 0$
If $i \in N$, the result clearly holds since $a_i^T \widetilde{x} > b_i$
If $i \in B$, since $A^= \overline{d} = e_j$, we have

$$a_i^T(\widetilde{x} + \varepsilon \overline{d}) = b_i + \varepsilon a_i^T\ \overline{d} = \begin{cases} b_i & i \neq j \\ b_i + \varepsilon & i = j \end{cases}$$

□

- case 1: $A\overline{d} \geq 0$
  Then by the Unboundedness Theorem, (P) is unbounded
- case 2:
  There exists $i \in N$, such that $a_i^T \overline{d} < 0$.
  (Note that $a_i^T \overline{d} \geq 0$ for each $i \in B$ since $A^= \overline{d} = e_j$)

Choose $\lambda \in \mathbb{R}$ maximum such that $\overline{x} + \lambda d \in P$

**Claim 3** $\widetilde{x} + \lambda \overline{d}$ is an extreme point and $c^T(\widetilde{x} + \lambda \overline{d}) < c^T \widetilde{x}$

**Proof** By Claims 1 and 2, $c^T(\widetilde{x} + \lambda \overline{d}) < c^T \widetilde{x}$
Let $L = \{\overline{x} + \alpha \overline{d} : \alpha \in \mathbb{R}\}$
Note that

$$L = \{x \in \mathbb{R}^n : a_i^T = b_i, i \in B \setminus \{j\}\}.$$

By our choice of $\lambda$, there exists $\ell \in N$ such that

$$a_\ell^T(\widetilde{x} + \lambda \overline{d}) = b_\ell \qquad \text{and} \qquad a_\ell^T(\overline{x} + \alpha \overline{d}) < b_\ell$$

for each $\alpha > \lambda$
Hence $\widetilde{x} + \lambda \overline{d}$ is the unique solution to

$$(a_i^T x = b_i : i \in (B \setminus \{j\}) \cup \{\ell\})$$

Then $\widetilde{x} + \lambda \overline{d}$ is an extreme point □

### 4.1.1 Simplex Algorithm

Given an extreme point $\widetilde{x}$ to (P)

**Step 1** Check non-degeneracy
Let $B = \{i \in \{1, \dots, m\} : a_i^T \widetilde{x} = b_i\}$.
Let $N = \{1, \dots, m\} \setminus B$, and let $A^= x \geq b^=$ be the equality subsystem. If $|B| \neq n$, STOP ($\widetilde{x}$ is degenerate)

**Step 2** Test for optimality
Solve $(A^=)^T = c$ for $\widetilde{y}$. If $\widetilde{y} \geq 0$, STOP ($\widetilde{x}$ is optimal)

**Step 3** Choosing leaving constraint.
Choose $j \in B$, such that $\overline{y}_j < 0$

**Step 4** Check unboundedness
Solve $A^= d = e_j$ for $\widetilde{d}$ and let $z = A^= \widetilde{d}$. If $z \geq 0$, STOP ( $(P)$ is unbounded)

**Step 5** Choose entering constraint
Choose $i \in N$ with $z_i < 0$ minimizing $\dfrac{a_i^T \widetilde{x} - b_i}{-z_i}$

**Step 6** Update.
Let $\lambda = \dfrac{a_i^T \widetilde{x} - b_i}{-z_i}$, replace $\widetilde{x}$ with $\widetilde{x} + \lambda \widetilde{d}$, and repeat from Step 1.

---

**Remark** If there are no degenerate extreme points, then the algorithm will terminate, (Since there are at most $\binom{m}{n}$ extreme points) and will solve (P) correctly.

**Example**

$$\begin{cases} \min\ x_1 + x_2 & \text{subject to} \\ 2x_1 + x_2 \geq 4 & (1) \\ 2x_1 + 3x_2 \geq 6 & (2) \\ x_1 + 4x_2 \geq 4 & (3) \end{cases}$$

Consider a feasible solution $\widetilde{x} = \begin{bmatrix} 12/5 \\ 2/5 \end{bmatrix}$ The equality subsystem is

$$\begin{cases} 2x_1 + 3x_2 \geq 6 & (2) & y_2 \\ x_1 + 4x_2 \geq 4 & (3) & y_3 \end{cases}$$

So $\widetilde{x}$ is a non-degenerate extreme point. Solve $\begin{cases} 2y_2 + y_3 = 1 \\ 3y_2 + 4y_3 = 1 \end{cases}$. we get $\widetilde{y}_2 = \frac{3}{5}$ and $\widetilde{y}_3 = -\frac{1}{5}$

So the leaving constraint is (3). Solve $\begin{cases} 2d_1 + 3d_2 = 0 \\ d_1 + 4d_2 = 1 \end{cases}$ we get $d = \begin{bmatrix} -3/5 \\ 2/5 \end{bmatrix}$

Let $\widehat{x} = \underbrace{\begin{bmatrix} 12/5 \\ 2/5 \end{bmatrix}}_{\widetilde{x}} + \lambda \underbrace{\begin{bmatrix} -3/5 \\ 2/5 \end{bmatrix}}_{d}$

Choose $\lambda$ maximum subject to $2\widehat{x}_1 + \widehat{x}_2 \geq 4$ (1).
We get $\lambda = \frac{3}{2}$ and the entering constraint is (1)

The new extreme point is $\widehat{x} = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix}$ and the equality subsystem is $\begin{cases} 2x_1 + x_2 \geq 4 & (1) \\ 2x_1 + 3x_2 \geq 6 & (2) \end{cases}$. So $\widehat{x}$ is degenerate.

Solve $\begin{cases} 2y_1 + 2y_2 = 1 \\ y_1 + 3y_2 = 1 \end{cases}$. We get $\widetilde{y}_1 = \widetilde{y}_2 = \frac{1}{4}$. Since $\widetilde{y}_1, \widetilde{y}_2 \geq 0$. Note that $\widetilde{y} = \begin{bmatrix} 1/4 \\ 1/4 \\ 0 \end{bmatrix}$ is an optimal solution for the dual of (P).

### 4.1.2 Perturbation Method (avoiding degeneracy)

$$\begin{array}{ll} (P) & \min(c^T x : Ax \geq b) \\ (P') & \min(c^T x : Ax \geq b') \end{array} \qquad \text{for } b' = \begin{bmatrix} b_1 - \varepsilon \\ b_2 - \varepsilon^2 \\ \vdots \\ b_m - \varepsilon^m \end{bmatrix}$$

and $\varepsilon$ is a variable that we think of as a small positive real number.
For polynomials, $p(\varepsilon)$ and $q(\varepsilon)$, we write $p(\varepsilon) < q(\varepsilon)$ if $p(\varepsilon') < q(\varepsilon')$ for all sufficiently small $\varepsilon' > 0$

**Example** $1 + \varepsilon + 1000\varepsilon^2 < 1 + 2\varepsilon$

**Claim** (P') has not degenerate points

**Proof** Consider an extreme points $\widetilde{x}$ of (P'). Let $X = \{i \in \{1, \ldots, m\} : a_i^T \widetilde{x} = b_i'\}$. If $\widetilde{x}$ is degenerate, then the vectors $\{a_i : i \in X\}$ are linearly independent. So there is a non-zero $\lambda \in \mathbb{R}^X$ such that $\sum_{i \in X} \lambda_i a_i = 0$. Thus

$$0 = \sum_{i \in X} \lambda_i a_i^T \widetilde{x} = \sum_{i \in X} b_i' = \sum_{i \in X} \lambda_i (b_i - \varepsilon^i).$$

However, since $\lambda \neq 0$, $\sum_{i \in X} \lambda_i (b_i - \varepsilon^i)$ is a non-zero polynomial in $\varepsilon$ - contradiction □

**Remarks**

(1) Since (P') is non-degenerate, the Simplex Method will solve (P') correctly.

(2) There is some computational overhead in applying this method, but we can switch between (P) and (P') easily, so we need only use $(P^*)$ when we are at a degenerate solution for (P)

(3) the equality subsystem has $m'$ constraints, we need only perturb $m' - n$ of them.

# 5 Integer Programming

**Definition** An integer program is problem of the form:

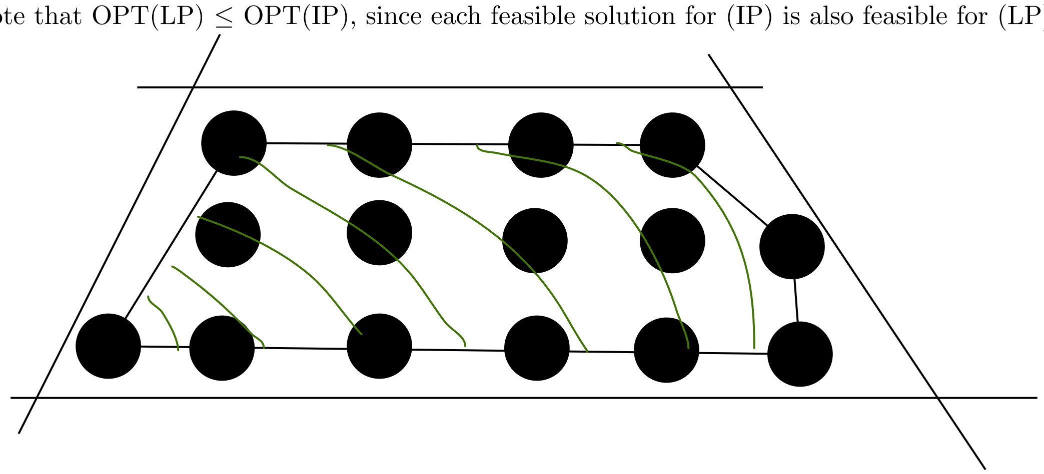
(IP) $\min(c^T x : Ax \geq b, x \in \mathbb{Z}^n)$,

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$

The linear programming relaxation of (IP) is

(LP) $\min(c^T x : Ax \geq b)$

Note that OPT(LP) $\leq$ OPT(IP), since each feasible solution for (IP) is also feasible for (LP).



Let $Z$ be the set of feasible solutions for (IP)
Then $\operatorname{conv}(Z) \subseteq \{x \in \mathbb{R}^n : Ax \geq b\}$.
Equality is unusual, even when $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$.

**Definition** A polyhedron $P \subseteq \mathbb{R}^n$ is integral is $P = \operatorname{conv}(P \cap \mathbb{Z}^n)$.

Note that if $P$ is integral, then $\min(c^T x : x \in P \cap \mathbb{Z}^n) = \min(c^T x : x \in P)$

**Lemma 5.1**

A polytope $P \subseteq \mathbb{R}^n$ is integral if and only if its points are integer valued.

**Proof** This form Theorem 3.8 $\square$

## 5.1 Totally Unimodular Matrices

A Matrix is totally unimodular (TU) if each of its square submatrices has determinant $0 \smile \pm 1$. (In particular, the entries are $0, \pm 1$)

Let $A \in \{0, \pm 1\}^{m \times n}$ be TU and $b \in \mathbb{Z}^m$, then the extreme points of $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ are integer valued.

**Proof** Let $\widetilde{x}$ be an extreme point of $P$, then there is a subsystem $A'x \geq b'$ of $Ax \geq b$ such that $A'\widetilde{x} = b'$ and $A'$ is square non-singular. Since $A$ is TU, $\det A = \pm 1$. By Cramer's Rule, each entry of $(A')^{-1}$ is integer valued. Hence each entry of $\widetilde{x} = (A')^{-1} b'$ is integer valued. □

Let $A \in \{0, \pm 1\}^{m \times n}$ be TU, then

(1) $A^T$ is TU

(2) $\begin{bmatrix} I & , & A \end{bmatrix}$ is TU

(3) If $A'$ is obtained from $A$ by scaling a row or column by $-1$, then $A'$ is TU

(4) $\begin{bmatrix} A & , & -A \end{bmatrix}$ is TU.

**Exercises** Let $A \in \{0, \pm 1\}^{m \times n}$ be TU and let $b \in \mathbb{Z}^m$:

- Show that the extreme points of $\{x \in \mathbb{R}^n : Ax \geq b, x \geq 0\}$ are integer valued. $([I, \ A]^T)$
- Show that the extreme points of $\{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ are integer valued. $([A, \ -A]^T)$

**Lemma 5.3**

If $A \in \{0, \pm 1\}^{m \times n}$ is TU, $b \in \mathbb{Z}^m$, and $l, u \in \mathbb{Z}^n$, then $P = \{x \in \mathbb{R}^n : Ax \geq b, \ l \leq x \leq u\}$ is integral.

**Proof** Let $A' = \begin{bmatrix} A \\ I \\ -I \end{bmatrix}$ and $b' = \begin{bmatrix} b \\ l \\ u \end{bmatrix}$. Note that $P = \{x \in \mathbb{R}^n : A'x \geq b'\}$. By the constructions (1) - (4) above, $A'$ is TU. By Lemma 5.2, each extreme point of $P$ is integer valued. Moreover, $P$ is polytope since $l \leq x \leq u$ for each $x \in P$. So by Lemma 5.1, $P$ is integral. □

**Theorem 5.4**

If $A \in \{0, \pm 1\}^{m \times n}$ is TU and $b \in \mathbb{Z}^m$, then $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ is integral.

**Proof** Exercise. (Hint: use Lemma 5.3) □

**Lemma 5.5**

Let $A \in \{0, \pm 1\}^{m \times n}$. If each column of $A$ has at most one 1 and at most one -1, then $A$ is TU.

$$A = \begin{bmatrix} \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots \\ \vdots & 0 & \vdots \\ \vdots & 0 & \vdots \\ \vdots & -1 & \vdots \\ \vdots & 0 & \vdots \\ \vdots & 1 & \vdots \\ \vdots & 0 & \vdots \\ \vdots & \vdots & \vdots \\ \cdots & 0 & \cdots \end{bmatrix}$$

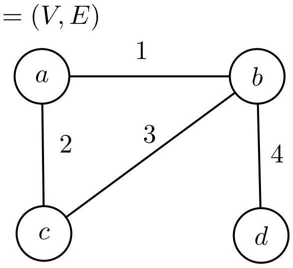
$$\left(\begin{array}{c|ccc} \pm1 & & & \\ \hline 0 & \iddots & \iddots & \iddots \\ 0 & \iddots & \iddots & \iddots \\ \vdots & \iddots & \iddots & \iddots \\ 0 & \iddots & \iddots & \iddots \end{array}\right)$$

**Proof** Suppose otherwise and consider a counterexample $A \in \{0, \pm 1\}^{m \times n}$ with $(m+n)$ minimum.

Clearly $m = n$ and $\det \notin \{0, \pm 1\}$. Since we have a minimum counterexample, each contains both a 1 and -1. But then the rows of $A$ sum to zero. Hence $\det(A) = 0$, a contradiction. $\square$

## 5.2 Incidence Matrix of a Graph

$G = (V, E)$




$$\begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

1, 2, 3, 4 are edges, and *abcd* are vertices. Note that

(i) The column-sums are all 2.

(ii) The row-sum for row $v \in V$ is the number of neighbours of $v$ and is denoted $\deg(v)$.

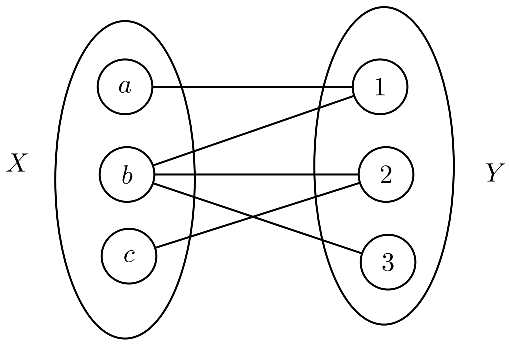
The incidence matrix need not to be TU.

**Example**

$$A' = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{array}$$

$\det(A') = -2$

### 5.2.1 Bipartite Graphs




A graph $G = (V, E)$ is bipartite with bipartition $(X, Y)$. If $(X, Y)$ is a partition of $V$, then each edge has an end in $X$ and an end in $Y$.

**Theorem 5.6**

The incidence matrix of a bipartite graph is TU.

$$\begin{array}{c} X \left\{ \vphantom{\begin{array}{c} 1 \\ 1 \end{array}} \right. \\ Y \left\{ \vphantom{\begin{array}{c} 1 \\ 1 \end{array}} \right. \end{array} \left[ \begin{array}{ccc} 1 & & \\ & & \\ \hdashline & & \\ 1 & & \end{array} \right]$$

**Proof** Let $(X, Y)$ be a partition of a graph $G = (V, E)$ and let $A$ be the incidence matrix. Let $A'$ be obtained by the rows indexed by $Y$ by $-1$. By Lemma 5.5, $A'$ is TU, and hence, $A$ is TU. $\square$

Let $A$ be the incidence matrix of graph $G = (V, E)$. Define

- matching $M(G) = \{x \in \mathbb{R}^E : Ax \leq 1, x \geq 0\}$, and
- perfect matching $PM(G) = \{x \in \mathbb{R}^E : Ax = 1, x \geq 0\}$.

For $x \in \mathbb{R}^E$, let $\mathit{Support}(x) = \{e \in E : x_e \neq 0\}$.

Note that

(1) For $x \in \{0, 1\}^E$, $x \in M(G)$ if and only if $\mathit{Support}(x)$ is a matching, and

(2) For $x \in \{0, 1\}^E$, $x \in PM(G)$ if and only if $\mathit{Support}(x)$ is a perfect matching.

Let $\mathcal{M}(G) = M(G) \cap \{0, 1\}^E$, and $\mathcal{PM}(G) = PM(G) \cap \{0, 1\}^E$.

**Theorem 5.7**

If $G$ is a bipartite graph, then

- $\text{conv}(\mathcal{M}(G)) = M(G)$, and
- $\text{conv}(\mathcal{PM}(G)) = PM(G)$.

**Proof** See above $\square$

### 5.2.2 Regular

A graph $G$ is $r$-regular if each of its vertices has degree $r$.

**Theorem 5.8**

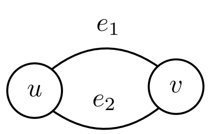
For each $r \geq 1$, if $G$ is an $r$-regular bipartite graph, then $G$ has a perfect matching.

**Proof** Let

$$\widetilde{x} = \left[\frac{1}{r}, \dots, \frac{1}{r}\right]^T$$

Hence $A\widetilde{x} = 1$, and $\widetilde{x} \geq 0$. Then by Theorem 5.7, $\widetilde{x} \in \text{conv}(\mathcal{PM}(G))$, then $PM(G) \neq \varnothing$. $\square$

### 5.2.3 Multigraph

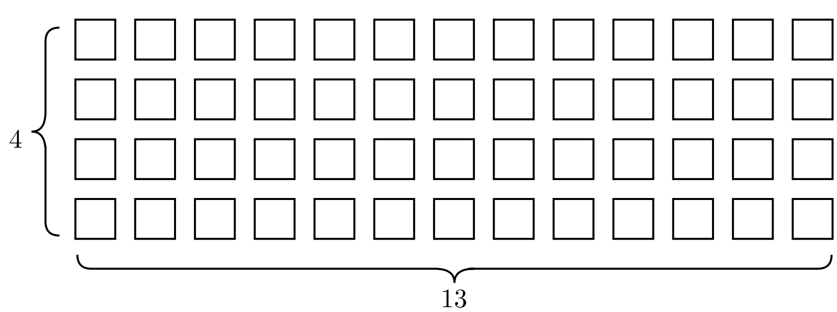



A multigraph is a graph which we allow parallel edges.

**Theorem 5.9**

For each $r \geq 1$, if $G$ is an $r$-regular bipartite multigraph, then $G$ has a perfect matching.

**Proof** Same as for Theorem 5.8. $\square$



**Exercise** Show that if we arrange a deck of cards in a rectangle with 4 rows and 13 columns, then rearranging the cards within each column, we can get each row containing the cards

ace, two, three, four, five, six, seven, eight, nine, ten, jack, queen, king

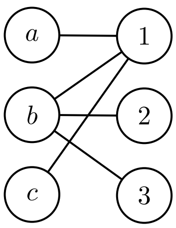
in same order.

**Exercise** Show that if each row and column sum of a matrix $A \in \mathbb{R}^{m\times n}$ is zero, then there is a matrix $B \in \mathbb{Z}^{m\times n}$ such that

(1) each row and column of $B$ sums to zero, and

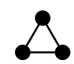
(2) $\lfloor a_{ij} \rfloor \le b_{ij} \le \lceil a_{ij} \rceil$ for each $i \in \{1,\dots,m\}$ and $j \in \{1,\dots,n\}$.

### 5.2.4 Covers




$C = \{1,\ b\}$

$C \subseteq V$ is a cover if $G - C$ has no edges.

Note that if $C$ is a cover and $M$ is a matching, $|M| \le |C|$. Equality is not always attained. 

### 5.2.5 Kőnig's Theorem

In a bipartite graph, the maximum size of a matching is equal to the minimum size of a cover.

**Proof** Let $A$ be the incidence matrix of a bipartite graph $G$.
Consider

$$(P) \quad \max\left(\sum_{e\in E} x_e : Ax \le 1, x \ge 0\right)$$

and its dual

$$(D) \quad \min\left(\sum_{v\in V} y_v : A^T y \ge 1, y \ge 0\right)$$

Note that $(P)$ is feasible $(x = 0)$ and $(D)$ is feasible $(y = 1)$. Hence $(P)$ and $(D)$ both have optimal solutions and $\mathrm{OPT}(P) = \mathrm{OPT}(D)$.

Moreover, since $A$ is TU, the feasible regions of both $(P)$ and $(D)$ are integral. Hence $(P)$ and $(D)$ have optimal solutions, $\widetilde{x}$ and $\widetilde{y}$ say, that are both integer valued.
Note that $\widetilde{x} \in \{0,1\}^E$ and $\widetilde{y} \in \{0,1\}^V$. Let $M = \mathit{Support}(\widetilde{x})$ and $C = \mathit{Support}(\widetilde{y})$.
Note that $M$ is a matching and $C$ is a cover.
Moreover, since $\mathrm{OPT}(P) = \mathrm{OPT}(D)$,

$$|M| = \sum_{e\in E} \widetilde{x}_e = \sum_{v\in V} \widetilde{y}_v = |C|$$

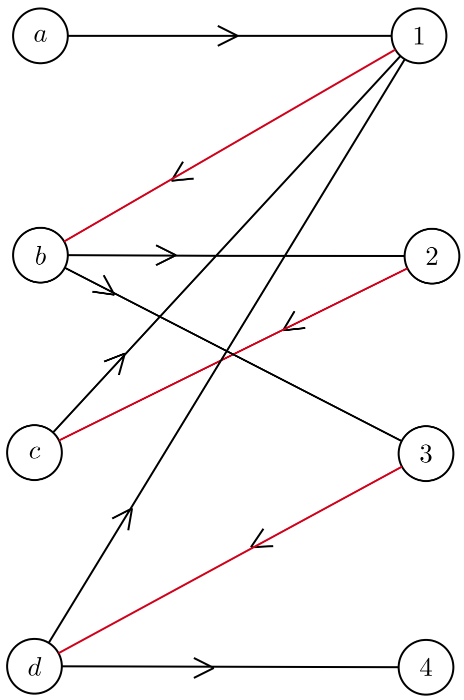
as required. $\square$

### 5.2.6 Finding a maximum matching

Let $G = (V, E)$ be a bipartite graph with bipartition $(X, Y)$ and $M$ be a matching.

**Problem** Find a larger matching if possible.

**Example**




$M = \{b1,\ c2,\ d3\}$

$a$ and 4 are $M$-exposed

$P = (a,\ 1,\ b,\ 3,\ d,\ 4)$

$M' = \{b1,\ c2,\ d3\}\ \Delta\ E(P)\ = \{a1,\ b3,\ c2,\ d4\}$

**Claim** $M$ is a maximum matching in $G$ if and only if there is no directed path form an $M$-exposed vertex in $X$ to an $M$-exposed vertex in $Y$.

**Proof** Exercise. □

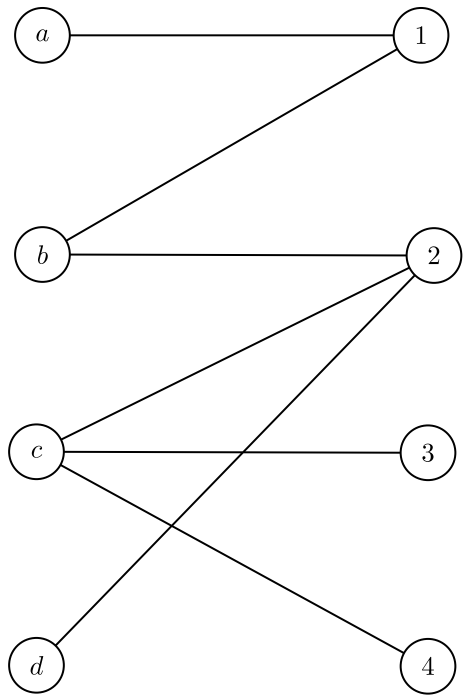
This an efficient algorithm for finding a maximum matching in a bipartite graph.

**Problem** How would you find a minimum cover in bipartite graph?

Note that if $v \in V$ is a minimum cover if and only if the matching number of $G$ decreases when we delete $v$. By repeating deleting vertices we can find a minimum cover.

### 5.2.7 Perfect Matchings

**Example**

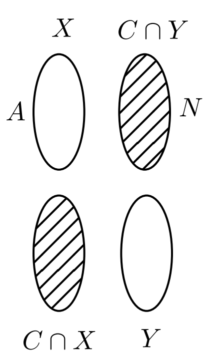



$N(\{a,b,d\}) = \{1,2\}$, so
$G$ has no perfect matching.

Here $N(X)$ denotes the set of vertices in $V \setminus X$ that have a neighbour in $X$

### 5.2.8 Hall's Theorem

A bipartite graph $G$ with bipartition $(X,Y)$ has a perfect matching if and only if $|X| = |Y|$ and $|N(A)| \geq |A|$ for each $A \subseteq X$.




**Proof** The conditions are clearly necessary. Suppose that $G$ has no perfect matching and that $|X| = |Y|$.

By Kőnig's Theorem, $G$ has a cover $C$ with $|C| < |X|$. Let $A = X \setminus C$ and $N = C \setminus X$. Since $C$ is a cover, $N(A) \leq N$. Moreover, since $|X| = |Y|$,

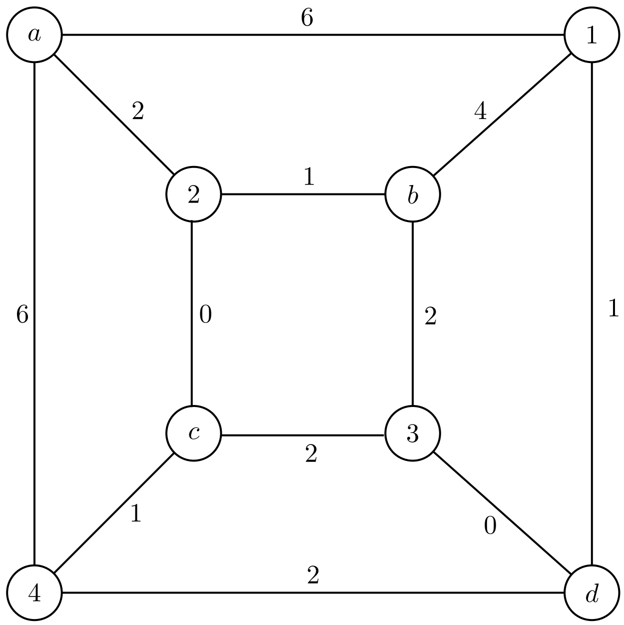
$$\begin{aligned} |A| &= |X| - |C| + |N| \\ &> |N| \\ &\geq |N(A)| \end{aligned}$$

□

## 5.3 Minimum Cost Perfect Matching in Bipartite Graphs

**Instance** A bipartite graph $G = (V,E)$ and $c \in \mathbb{Q}^E$.

**Problem** Find a perfect matching $M$ minimizing $c(M)$.
Here, $c(M) = \sum_{e \in M} c(e)$.




$\widetilde{M} = \{a2,\ b3,\ c4,\ d1\}$

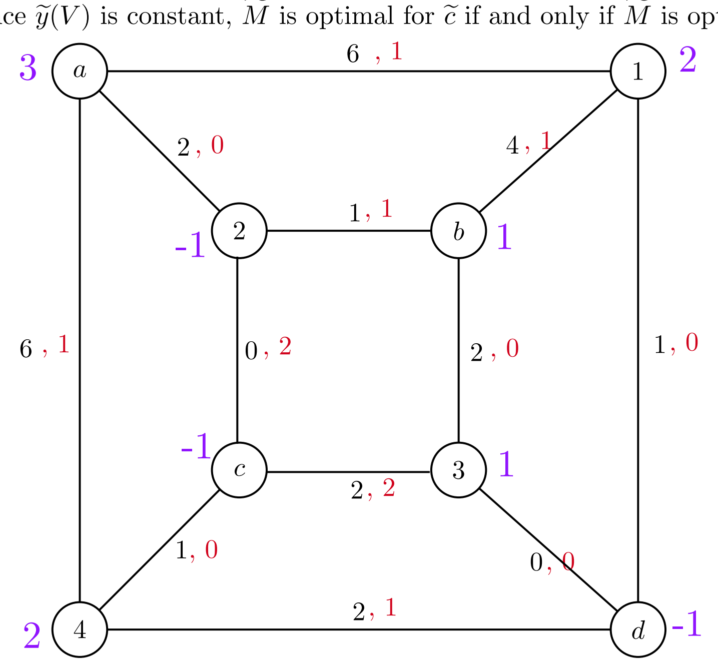
**Claim** $\widetilde{M}$ is optimal.

**Idea** Suppose $c'(e) = \begin{cases} c(e) + \alpha : & e \text{ incident with } a \\ c(e) : & \text{otherwise} \end{cases}$

Then for any perfect matching $M$, $c'(M) = c(M) + \alpha$. For $\widetilde{y} \in \mathbb{R}^V$, we define the reduced cost of $e = uv \in E$ to $\widetilde{c}_e = c_e - \widetilde{y}_u - \widetilde{y}_v$. Then for any perfect matching $M$,

$$\widetilde{c}(M) = c(M) - \widetilde{y}(V)$$

Since $\widetilde{y}(V)$ is constant, $\widetilde{M}$ is optimal for $\widetilde{c}$ if and only if $\widetilde{M}$ is optimal for $c$.




Vertex labels: $\widetilde{y}_v$
Edge labels: $c_e$ ,$\widetilde{c}_e$

Note that $\widetilde{c} \geq 0$ and $\widetilde{c}(\widetilde{M}) = 0$, so $\widetilde{M}$ is optimal for $\widetilde{c}$, and hence for $c$.

**Theorem 5.10**

If $M$ is a minimum cost perfect matching, then there exist $\widetilde{y} \in \mathbb{R}^V$ such that $\widetilde{c} \geq 0$ and $\widetilde{c}(M) = 0$.

**Proof** Consider the linear program:

$$(P) \qquad \min \left( c^T x : Ax = 1, x \geq 0 \right)$$

and its dual

$$(D) \qquad \max \left( y(V) : A^T y \leq c \right)$$

where $A$ is the incidence matrix of $G$.
Since $A$ is TU, $\mathrm{OPT}(P) = c(M)$. Let $\widetilde{y}$ be the optimal solution for $(D)$. Note that

$$\widetilde{c} = c - A^T y \geq 0$$

By strong Duality Theorem,

$$\widetilde{y}(V) = \mathrm{OPT}(P) = c(M)$$

So $\widetilde{c}(M) = c(M) - \widetilde{y}(V) = 0$. $\square$

### 5.3.1 Minimum cost perfect matching algorithm

We call $\widetilde{y} \in \mathbb{R}^V$ feasible if $\widetilde{c} \geq 0$. The equality subgraph $G^{=}(\widetilde{y})$, has vertex set $V$ and edge set $E^{=}(\widetilde{y}) = \{e \in E : \widetilde{c}_e = 0\}$. Thus if $M$ is a perfect matching of $G^{=}(\widetilde{y})$, then $M$ is optimal.

**Example** To be completed...

# 6

# Convex Optimization

To be completed...