# CO327 (2022 Spring) condensed notes on the theory part of LP and LIP

Instructor: Andersen Ang
Homepage: angms.science/CO327.html

Saturday $25^{\text{th}}$ June, 2022

The note is for quick review of useful mathematical definitions and concepts for the purpose of this course. We will not go to deep into theory and we may be not extremely precise here.

## Contents

**1 Standard form and canonical form** **2**
1.1 Type of LP . . . 3

**2 Equivalence and conversion between forms** **3**
2.1 Solving system of linear inequality: Fourier-Motzkin elimination . . . 4

**3 Basic linear geometry** **4**
3.1 Hyperplane and Half-space . . . 4
3.2 Polyhedron . . . 5
3.3 Fundamental Theorem of LP . . . 6

**4 Basic properties of LP** **7**
4.1 Active and passive constraint . . . 7
4.2 c at optimal point will be inside the cone formed by the active rows in A . . . 7

**5 Perturbation and sensitivity** **8**
5.1 Sensitivity of LP with perturbation on b . . . 8
5.2 Small perturbation . . . 8
5.3 Other cases not discussed . . . 9
5.4 Complementary slackness . . . 9

**6 Duality** **9**
6.1 Karush-Kuhn-Tucker condition . . . 10

**7 Example** **11**

**8 Linear integer programming** **12**
8.1 Unimodular matrix . . . 12
8.2 Fundamental Theorem of LP relaxation . . . 13
8.3 Integral polyhedron . . . 14
8.4 How to check a matrix is TU . . . 14

**9 Some graph problems** **15**
9.1 Min-Cost-Flow . . . 15
9.2 Shortest Path Problems . . . 16
9.3 Other graph problems . . . 17
9.4 Linear assignment problem . . . 17
9.5 Stable marriage problem . . . 17

# 1 Standard form and canonical form

**Definition 1.1: Standard form and Canonical form**

Given $\boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{c} \in \mathbb{R}^n$, find $\boldsymbol{x} \in \mathbb{R}^n$ by solving

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0}_n \end{aligned} \quad \text{(Standard form)} \qquad \begin{aligned} \max_{\boldsymbol{x}} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \end{aligned} \quad \text{(Canonical form)}$$

The primal objective value $p \coloneqq \boldsymbol{c}^\top \boldsymbol{x}$.

- Notation: $x$ denotes scalar and $\boldsymbol{x}$ denotes vector.
  The symbol $\geq$ is taken element-wise by convention.
- The function $\boldsymbol{c}^\top \boldsymbol{x}$ is called objective function and it is a linear function.
- In standard form, usually $\boldsymbol{A}$ is singular / non-invertible. Why: otherwise $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \implies \boldsymbol{x} = \boldsymbol{A}^\top \boldsymbol{b}$ is unique $\implies$ we solved the LP.
- In canonical form, the constraint $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ describes a "polyhedron membership" $\longrightarrow$ linear geometry.
- **Definition (Feasibility)** A point $\boldsymbol{x}$ is *feasible* if it satisfies the constraint.
- **Definition (Optimal point)** A point $\boldsymbol{x}^*$ is *optimal*, or called a *solution*, if $p^* \coloneqq \boldsymbol{c}^\top \boldsymbol{x}^* \geq p = \boldsymbol{c}^\top \boldsymbol{x}$ for all feasible $\boldsymbol{x}$.
- Trivial fact: a feasible point need not to be an optimal point.
- LP with $\boldsymbol{c} = \boldsymbol{0}$ is an instance of *Constraint Satisfaction Problem*.
- $<$ or $>$ not allowed in constraint because we want closedness (in the topological sense).

**Definition 1.2: Symmetric primal-dual form**

$$\begin{aligned} \max_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0}_n \end{aligned} \quad \text{(Symmetric primal form)} \qquad \begin{aligned} \min_{\boldsymbol{y} \in \mathbb{R}^m} \quad & \boldsymbol{b}^\top \boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{A}^\top \boldsymbol{y} \geq \boldsymbol{c} \\ & \boldsymbol{y} \geq \boldsymbol{0}_m \end{aligned} \quad \text{(Symmetric dual form)}$$

The dual objective value $d \coloneqq \boldsymbol{b}^\top \boldsymbol{y}$.

- Primal and dual are highly-related $\longrightarrow$ duality.
- Computational issue: depends on the size of $(m, n)$, it is easier to solve one of them.
  - If $m \gg n$, $\boldsymbol{A}$ is thin-and-tall, and we have lots of inequalities constraints but fewer variables on $\boldsymbol{x}$.
    $\implies \boldsymbol{A}^\top$ is short-and-fat, and we have a few inequalities constraint but many variables on $\boldsymbol{y}$.

### 1.1 Type of LP

- Infeasible LP
  The LP is unsolvable: there is no solution that satisfies all the constraints.
- Feasible LP
  The LP is solvable: there is (at least one) finite optimal sol. that satisfies all the constraints.
  - Solution is unique.
    Geometry: optimal sol. is located at a corner (vertex) of the feasible region.
  - Solution is not unique.
    Geometry: optimal sol. is located at an facet (face/edge) of the feasible region.
- Unbounded LP
  The LP is solvable: $p^* = +\infty$ and the optimal sol. is at $\infty$.

## 2 Equivalence and conversion between forms

**Theorem 2.1** (Informal)**.** *All LP can be converted to standard form (and canonical form). i.e.,*

$$\begin{array}{ll} \max\limits_{\boldsymbol{x}} & \boldsymbol{c}_S^\top \boldsymbol{x} \\ s.t. & \boldsymbol{A}_S \boldsymbol{x} = \boldsymbol{b}_S \\ & \boldsymbol{x} \geq \boldsymbol{0}_S \\ & \textit{Standard form} \end{array} \iff \begin{array}{ll} \max\limits_{\boldsymbol{x}} & \boldsymbol{c}^\top \boldsymbol{x} \\ s.t. & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \\ & \textit{Symmetric primal form} \end{array} \iff \begin{array}{ll} \max\limits_{\boldsymbol{x}} & \boldsymbol{c}_C^\top \boldsymbol{x} \\ s.t. & \boldsymbol{A}_C \boldsymbol{x} \leq \boldsymbol{b}_C \\ & \\ & \textit{Canonical form} \end{array}.$$

The matrices $\boldsymbol{A}, \boldsymbol{A}_{\mathrm{S}}, \boldsymbol{A}_{\mathrm{C}}$ in the above forms are in general different (same for $\boldsymbol{b}, \boldsymbol{c}$). Technically, solution $\boldsymbol{x}$ in these forms are not the same (because of different dimensions in $\boldsymbol{c}, \boldsymbol{c}_S, \boldsymbol{c}_C$) but they all solve the same problem.

**Example** Tutorial 1 slide 24.

**Conversion tricks** Tutorial 1 slide 19-31.

- $x_i \leq u_i$ is equivalent to $x_i + s_i = u_i$, $s_i \geq 0$. (Tutorial 1 slide 20)
- $x_i \geq l_i$ is equivalent to $x_i - s_i = l_i$, $s_i \geq 0$.
- $l_i \leq x_i \leq u_i$ is equivalent to $0 \leq s_i \leq u_i - l_i$ via $x_i = l_i + s_i, s_i \geq 0$ (Tutorial 1 slide 22)
- $x_i = b_i$ is equivalent to $x_i \leq b_i$ AND $x_i \geq b_i$.
- When converting to standard form, free variable $x_i$ is equivalent to $x_i = s_1 - s_2$ with $s_1 \geq 0, s_2 \geq 0$. (Tutorial 1 slide 23)
- See Tutorial 1 slide 24 for an example.

### 2.1 Solving system of linear inequality: Fourier-Motzkin elimination

**General problem**: determine whether a system of linear inequalities $\boldsymbol{Ax} \leq \boldsymbol{b}$ has a solution.

**Fourier-Motzkin elimination method**

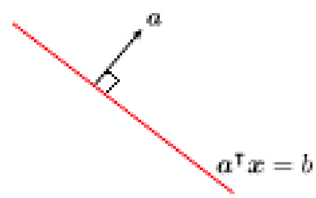
$$\boldsymbol{Ax} \leq \boldsymbol{b} \quad \iff \quad \boldsymbol{a}_i^\top \boldsymbol{x} \leq b_i \quad \iff \quad \sum_{j=1}^{n} a_{ij}x_j \leq b_i,\ i = 1, 2, \ldots, m$$

To eliminate $x_k$,

1. **"Normalization step"** Make the coefficient $a_{jk}$ to 1 or $-1$
2. **"Elimination step"** For each (new) inequalities that $x_k$ has positive coefficient and for each (new) inequalities that $x_k$ has negative coefficient, sum the two inequalities to get a new inequality , or combine two inequalities to form
$$f_1 \leq x \leq f_2$$
and hence we eliminate $x$ by considering $f_1 \leq f_2$

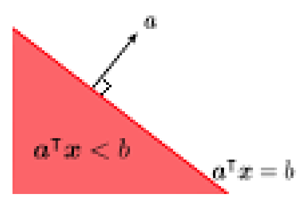
# 3 Basic linear geometry

### 3.1 Hyperplane and Half-space




**Definition 3.1** (Affine hyperplane)**.** *Given $\boldsymbol{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$, the linear equation $\boldsymbol{a}^\top \boldsymbol{x} = b$ is called a hyperplane. If $b \neq 0$, it is called affine; otherwise, it is called linear.*

- $\boldsymbol{a}$ controls the orientation of the plane and $b$ controls the its position.
- If $b \neq 0$, the hyperplane only has exactly one normal vector $\boldsymbol{a}$.
- If $b = 0$, the hyperplane has two normal vectors $\boldsymbol{a}$ and $-\boldsymbol{a}$.



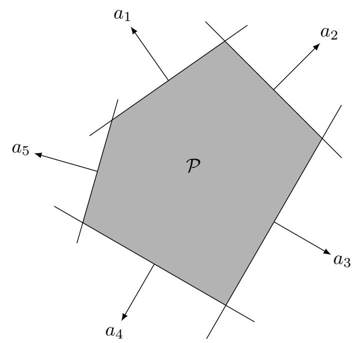

**Definition 3.2** (Half-space)**.** *Given $\boldsymbol{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$, the linear inequality $\boldsymbol{a}^\top \boldsymbol{x} > b$ is called a half-space.*

- The inequality can be strict $>$ or non-strict $\geq$.
- Half-space is not affine.
- the direction of $\boldsymbol{a}$ in $\boldsymbol{a}^\top \boldsymbol{x} = b$ and the direction of the halfspace need not agree with each other

**Implication of the expression of the primal objective function $\boldsymbol{c}^\top \boldsymbol{x}$**

- In $\boldsymbol{a}^\top \boldsymbol{x} = b$, the value $b$ controls the location of the plane: it "labels" the plane by the vatue $b$.
- Moving along the direction $\boldsymbol{a}$ increases the value $b$. This implies that moving $\boldsymbol{x}$ along the $\boldsymbol{c}$ direction in the primal objective function $\boldsymbol{c}^\top \boldsymbol{x}$ increases the value of $p$. So you want to move along the direction $\boldsymbol{c}$ as much as possible within the feasible region of the LP $\implies$ you want to move $\boldsymbol{x}$ to the extreme region of the feasible set.

## 3.2 Polyhedron




**Definition 3.3** ($\mathcal{H}$-representation of polyhedron)**.** *A polyhedron $\mathcal{P} \subset \mathbb{R}^n$ is a set of the form*

$$\mathcal{P}(\boldsymbol{A}, \boldsymbol{b}) \coloneqq \Big\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\Big\}.$$

*That is, it is a set defined by the system of linear inequalities $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$. Such representation of $\mathcal{P}$ is also known as the $\mathcal{H}$-representation ($\mathcal{H}$ stands for half-space).*

- As all LP can be written in canonical form (Theorem 2.1), so solving LP is to find the optimal point within a polyhedron.

**Type of polyhedron in LP**

- **Existence** of the optimal point in LP: $\mathcal{P}$ can be bounded or unbounded.
- **Uniqueness** of the optimal point in LP: $\mathcal{P}$ can be convex or non-convex.
- **Algorithm complexity**: $\mathcal{P}$ can be simple (few facets) or complicated (many facets).

**Definition 3.4** (Bounded)**.** *A set $\mathcal{X} \subset \mathbb{R}^n$ is bounded if there exists a constant $M > 0$ such that for all $\boldsymbol{x} \in \mathcal{X}$ we have $\|\boldsymbol{x}\| \leq M$*

**Definition 3.5** (Polytope)**.** *A polytope is a bounded polyhedron.*

**Definition 3.6** (Linear, conic and convex combination)**.** *Given a set of points $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^m$,*

| Combination | Expression | Condition on $c_i$ |
|---|---|---|
| Linear | $\sum c_i \boldsymbol{v}_i$ | No limitation |
| Conic | $\sum c_i \boldsymbol{v}_i,\ \forall\ c_i \geq 0$ | Nonnegative |
| Convex | $\sum c_i \boldsymbol{v}_i,\ \forall\ c_i \geq 0,\ \ \sum c_i = 1$ | Nonnegative and sum to 1 |

**Definition 3.7** (Line segment)**.** *Given two points $\boldsymbol{x}, \boldsymbol{y}$. A line $\ell(\boldsymbol{x}, \boldsymbol{y})$ is the set of all linear combination of $\boldsymbol{x}, \boldsymbol{y}$. Mathematically,*

$$\ell(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \{\boldsymbol{v} \mid \boldsymbol{v} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \lambda \in \mathbb{R}\}.$$

**Definition 3.8** (Internal line segment)**.** *Given two points $\boldsymbol{x}, \boldsymbol{y}$. An internal line $\ell(\boldsymbol{x}, \boldsymbol{y})$ is the set of all convex combination of $\boldsymbol{x}, \boldsymbol{y}$. The two ending points of $\ell(\boldsymbol{x}, \boldsymbol{y})$ is exactly $\boldsymbol{x}$ and $\boldsymbol{y}$. Mathematically,*

$$\ell(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \{\boldsymbol{v} \mid \boldsymbol{v} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \lambda \in [0, 1]\}.$$

**Definition 3.9** (Extreme point)**.** *Let $\mathcal{P}$ be a polyhedron. A vector $\boldsymbol{x} \in \mathcal{P}$ is an extreme point of $\mathcal{P}$ if there is no $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{P}$ such that $\boldsymbol{x} \in \ell(\boldsymbol{u}, \boldsymbol{v})$.*

**Definition 3.10** (Vertex)**.** *Let $\mathcal{P}$ be a polyhedron. A vector $\boldsymbol{x} \in \mathcal{P}$ is a vertex of $\mathcal{P}$ if there is a objective function $\boldsymbol{c}$ that $x$ is the unique optimal point maximizing over $\mathcal{P}$. Mathematically, $\boldsymbol{x} \in \mathcal{P}$ is a vertex of $\mathcal{P}$ if there exists some $\boldsymbol{c}$ such that $\boldsymbol{c}^\top \boldsymbol{x} > \boldsymbol{c}^\top \boldsymbol{u}$ for all $\boldsymbol{x} \neq \boldsymbol{u} \in \mathcal{P}$.*

**Fact** Let $\mathcal{P}$ be a polyhedron. The following are equivalent.

- $\boldsymbol{x}$ is a vertex.
- $\boldsymbol{x}$ is an extreme point.

- $\boldsymbol{x}$ is a basic feasible solution.

**Corollary 3.1.** *Any polytope (bounded polyhedron) has finitely many extreme points.*

**Definition 3.11** (Cone)**.** *A set $\mathcal{C} \subset \mathbb{R}^n$ is a cone if $\boldsymbol{x} \in \mathcal{C} \implies \lambda\boldsymbol{x} \in \mathcal{C}$ for all $\lambda \geq 0$.*

- The origin $\mathbf{0}$ must be a member of every cone.
- The origin is the only possible extreme point for a cone.
- An element of a cone is called a ray.

**Definition 3.12** (Conical hull)**.** *Given a set of vertices $V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. Thee conical hull of $V$, denoted as cone$(V)$, is defined as the set of all possible conic combination of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$.*

$$cone(V) = \left\{ \sum_{i=1}^{n} c_i \boldsymbol{v}_i \mid c_i \geq 0 \quad \forall i \right\}.$$

**Definition 3.13** (Convex hull)**.** *Given a set of vertices $V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. Thee convex hull of $V$, denoted as conv$(V)$, is defined as set of all possible convex combination of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$.*

$$conv(V) = \left\{ \sum_{i=1}^{n} c_i \boldsymbol{v}_i \mid \sum_{i=1}^{n} c_i = 1, c_i \geq 0 \quad \forall i \right\}.$$

**Definition 3.14** (Polyhedral cone)**.** *A $\mathcal{H}$-polyhedron $\mathcal{P}(\boldsymbol{A}, \boldsymbol{b})$ with $\boldsymbol{b} = \mathbf{0}$ in the form of $\mathcal{P} = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \geq \mathbf{0}\}$ is a polyhedral cone.*

## 3.3 Fundamental Theorem of LP

**Theorem 3.1** (Informal)**.** *If a bounded LP$(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c})$ has a solution $\boldsymbol{x}^*$, then the optimizer $\boldsymbol{x}^* \in$ bdry$(\mathcal{P})$, where $\mathcal{P}$ is the polytope described by $\mathcal{H}$-representation $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$.*

*Proof.* By contradiction. Assume optimizer $\boldsymbol{x}^* \notin \text{bdry}(\mathcal{P}) \implies \boldsymbol{x}^* \in \text{int}(\mathcal{P})$. This means $\exists\, \delta > 0$ such that $\boldsymbol{c}^\top(\boldsymbol{x}^* + \delta\boldsymbol{c}) = \boldsymbol{c}^\top\boldsymbol{x}^* + \underbrace{\delta\|\boldsymbol{c}\|_2^2}_{\geq 0} \geq \boldsymbol{c}^\top\boldsymbol{x}^*$. i.e., the point $\boldsymbol{x}^* + \delta\boldsymbol{c}$ gives a better optimal value than $\boldsymbol{x}^*$. Contradiction to the assumption that $\boldsymbol{x}^*$ is an optimal point. □

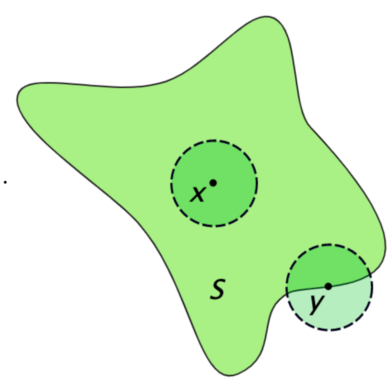
- int$(\mathcal{P})$ : interior of $\mathcal{P}$.

$$\boldsymbol{x} \in \text{int}(\mathcal{P}) \iff \exists \varepsilon \geq 0 \text{ s.t. } \mathbb{B}(\boldsymbol{x}; \varepsilon) \subset \mathcal{P},$$

- bdry$(\mathcal{P})$ : boundary of $\mathcal{P}$, which can be facet or extreme point (vertex).

$$\boldsymbol{x} \in \text{bdry}(\mathcal{P}) \iff \forall \varepsilon \geq 0, \exists (\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{B}(\boldsymbol{x}; \varepsilon) \text{ s.t. } \boldsymbol{u} \in \mathcal{P}, \boldsymbol{v} \notin \mathcal{P}$$

The example figure on the right: $\boldsymbol{x} \in \text{int}(\mathcal{S})$ and $\boldsymbol{y} \in \text{bdry}(\mathcal{S})$.



# 4 Basic properties of LP

## 4.1 Active and passive constraint

Consider canonical form: $\min\limits_{\boldsymbol{x}\boldsymbol{c}^\top\boldsymbol{x} \text{ s.t. } \boldsymbol{A}\boldsymbol{x}\leq\boldsymbol{b}}$.

**Definition 4.1: Active and passive constraint, active set**

For each row of the constraint $\boldsymbol{a}_i^\top\boldsymbol{x} \leq b_i$, we have either binding or non-binding situation:

- Binding / Active: $\boldsymbol{a}_i^\top\boldsymbol{x} = b_i$. The point $\boldsymbol{x}$ is on the hyperplane defined by $(\boldsymbol{a}_i, b_i)$.
- Non-binding / Passive: $\boldsymbol{a}_i^\top\boldsymbol{x} < b_i$ The point $\boldsymbol{x}$ is in the (strict) half-space defined by $(\boldsymbol{a}_i, b_i)$.

Active set $S(\boldsymbol{A}, \boldsymbol{x}, \boldsymbol{b}) \coloneqq \{i \mid \boldsymbol{a}_i^\top\boldsymbol{x} = b_i\}$.

Let $S$ be the active set of a feasible point $\boldsymbol{x}$.
Recall that optimal point for a LP is always on the boundary of $\mathcal{P}$, so $S(\boldsymbol{x}^*)$ is always nonempty and

$$\boldsymbol{A}\boldsymbol{x}^* \leq \boldsymbol{b} \implies \boldsymbol{A}_S\boldsymbol{x}^* = \boldsymbol{b}_S$$

where $\boldsymbol{A}_S$ is the matrix $\boldsymbol{A}$ only taking row with index inside $S$.

**Definition 4.2: Basic solution and basic feasible solution**

- If there are $N$ linearly independent constraints that are active at a point $\bar{\boldsymbol{x}} \in \mathbb{R}^N$, then $\bar{\boldsymbol{x}}$ is called a basic solution.
- For a basic solution $\bar{\boldsymbol{x}}$, if it is feasible, then it is called a basic feasible solution.

**Fact** At a basic feasible solution $\bar{\boldsymbol{x}}$, let the associated active set at $\bar{\boldsymbol{x}}$ for the constraint $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ be $S(\bar{\boldsymbol{x}})$. Then, the submatrix $\boldsymbol{A}_{S(\bar{\boldsymbol{x}})}$ is non-singular. This implies if we somehow know the active set $S$ of the minimizer $\boldsymbol{x}^*$, we get $\boldsymbol{x}_S^* = \boldsymbol{A}_S^{-1}\boldsymbol{b}_S$.

**Other words to describe active and passive**

- Active: $\boldsymbol{a}_i^\top\boldsymbol{x} = b_i$ is also called tight / binding.
- Passive: $\boldsymbol{a}_i^\top\boldsymbol{x} < b_i \iff \boldsymbol{a}_i^\top\boldsymbol{x} + s < b_i$, here $s \geq 0$ is call slack variable.
  Similarly, $\boldsymbol{a}_i^\top\boldsymbol{x} > b_i \iff \boldsymbol{a}_i^\top\boldsymbol{x} - s = b_i$, here $s \geq 0$ is call surplus variable.
  Both slack and surplus variables are nonnegative $s \geq 0$.

## 4.2 c at optimal point will be inside the cone formed by the active rows in A

Goal : you want to solve a LP$(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c})$.

1. We know that moving $\boldsymbol{x}$ along the direction $\boldsymbol{c}$ will increase the value of $p$ so we want to move along $\boldsymbol{c}$ as much as possible.
2. Moving $\boldsymbol{x}$ along $\boldsymbol{c}$ will ultimately hit the boundary of $\mathcal{P}$: by Theorem 3.1, the optimal point $\boldsymbol{x}^*$ is always on bdry$(\mathcal{P})$, either being an extreme point (vertex) or on a facet.
3. $\boldsymbol{x}^*$ on bdry$(\mathcal{P})$ will introduce an active set $\mathcal{S}(\boldsymbol{x}^*)$ and passive set for the constraints.

4. Considering the direction of the hyperplane in $\boldsymbol{A}_{\mathcal{S}(\boldsymbol{x}^*)}$, we have $\boldsymbol{c} \in (\boldsymbol{A}_{\mathcal{S}(\boldsymbol{x}^*)}^\top)$
5. Using conic combination: $\boldsymbol{c} \in (\boldsymbol{A}_{\mathcal{S}(\boldsymbol{x}^*)}^\top) \iff \exists \boldsymbol{y}, \boldsymbol{y} \geq \boldsymbol{0}$ such that $\boldsymbol{c} = \boldsymbol{A}^\top \boldsymbol{y}$, where $\boldsymbol{y}_{i \notin \mathcal{S}} = 0$ and $\boldsymbol{y}_{i \in \mathcal{S}} > 0$.
6. $\boldsymbol{c} = \boldsymbol{A}^\top \boldsymbol{y}$ with $\boldsymbol{y} \geq \boldsymbol{0}$ are actually the constraints of the dual problem of the original LP.

# 5 Perturbation and sensitivity

## 5.1 Sensitivity of LP with perturbation on b

Consider two LPs

$$(\mathcal{P}_0) \ : \ \begin{array}{cc} \min\limits_{\boldsymbol{x} \in \mathbb{R}^n} & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \end{array} \quad \text{and} \quad (\mathcal{P}_1) \ : \ \begin{array}{cc} \min\limits_{\boldsymbol{x} \in \mathbb{R}^n} & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} + \Delta \end{array}$$

where $\Delta \in \mathbb{R}^m$ is a perturbation vector that changes $\boldsymbol{b}$.

Suppose both $(\mathcal{P}_0)$ and $(\mathcal{P}_1)$ are solvable. Let $\boldsymbol{x}_0^*$ be a solution of $(\mathcal{P}_0)$ and let $\boldsymbol{x}_1^*$ be a solution of $(\mathcal{P}_1)$.

Q: What is the relationship between $\boldsymbol{x}_0^*$ and $\boldsymbol{x}_1^*$ regarding the perturbation $\Delta$?

Let $S_0$ be the active set of $\boldsymbol{x}_0^*$ for solving $(\mathcal{P}_0)$, so we have $\boldsymbol{A}_{S_0}\boldsymbol{x}_0^* = \boldsymbol{b}_{S_0}$.
Let $S_1$ be the active set of $\boldsymbol{x}_1*$ for solving $(\mathcal{P}_1)$, so we have $\boldsymbol{A}_{S_1}\boldsymbol{x}_1^* = (\boldsymbol{b} + \Delta)_{S_1} = \boldsymbol{b}_{S_1} + \Delta_{S_1}$.
Now there are two possibilities

- Case 1. The perturbation $\Delta$ is "small" such that it does not change the active set of the perturbed solution, i.e., $S_1 = S_0$.
- Case 2. The perturbation $\Delta$ is "large" such that it changes the active set of the perturbed solution, i.e., $S_1 \neq S_0$.

## 5.2 Small perturbation

Suppose $S_1 = S_0$. Now

$$\begin{array}{rcl} \boldsymbol{x}_0^* &=& \boldsymbol{A}_{S_0}^{-1}\boldsymbol{b}_{S_0} \\ \boldsymbol{x}_1^* &=& \boldsymbol{A}_{S_1}^{-1}(\boldsymbol{b}+\Delta)_{S_1} \\ & \overset{S_1=S_0}{=} & \boldsymbol{A}_{S_0}^{-1}(\boldsymbol{b}+\Delta)_{S_0} \\ &=& \boldsymbol{A}_{S_0}^{-1}\boldsymbol{b}_{S_0} + \boldsymbol{A}_{S_0}^{-1}\Delta_{S_0} \\ &=& \boldsymbol{x}_0^* + \boldsymbol{A}_{S_0}^{-1}\Delta_{S_0} \end{array} \quad \Longrightarrow \quad \begin{array}{rcl} p(\boldsymbol{x}_0^*) &=& \boldsymbol{c}^\top \boldsymbol{x}_0^* \\ p(\boldsymbol{x}_1^*) &=& \boldsymbol{c}^\top \boldsymbol{x}_1^* \\ &=& \boldsymbol{c}^\top(\boldsymbol{x}_0^* + \boldsymbol{A}_{S_0}^{-1}\Delta_{S_0}) \\ &=& \boldsymbol{c}^\top \boldsymbol{x}_0^* + \boldsymbol{c}^\top \boldsymbol{A}_{S_0}^{-1}\Delta_{S_0} \\ &=& p(\boldsymbol{x}_0^*) + \boldsymbol{c}^\top \boldsymbol{A}_{S_0}^{-1}\Delta_{S_0} \end{array}$$

Now $p(\boldsymbol{x}_0^*) - p(\boldsymbol{x}_1^*) = \boldsymbol{c}^\top \boldsymbol{A}_{S_0}^{-1}\Delta_{S_0}$. If we let $\boldsymbol{y}_{S_0} = \boldsymbol{A}_{S_0}^{-\top}\boldsymbol{c}$ then we have $p(\boldsymbol{x}_0^*) - p(\boldsymbol{x}_1^*) = \boldsymbol{y}_{S_0}^\top \Delta$, so

$$\frac{\Delta p}{\Delta} = \frac{\partial}{\partial \Delta}\Big(p(\boldsymbol{x}_0^*) - p(\boldsymbol{x}_1^*)\Big) = \frac{\partial}{\partial \Delta}\boldsymbol{y}_{S_0}^\top \Delta = \boldsymbol{y}_{S_0},$$

so $\boldsymbol{y}_{S_0}$ is telling us the amount of change in $p$ per the change $\Delta$. The vector $\boldsymbol{y}_{S_0}$ is often called *shadow price* in economics.

Note the the above analysis assume we know the active set $S_0(\boldsymbol{x}_0^*)$. In general we don't know the solution before we solve the problem and hence we don't know the associated active set. What we can do here is to remove the active set $S_0$ by adding zero in $\boldsymbol{y}$: we can set $y_{i \notin S_0} = 0$.

Recall from Section 4.2 that $\boldsymbol{c}$ at an optimal point $\boldsymbol{x}^*$ will be inside the cone($\boldsymbol{A}_{S(\boldsymbol{x}^*)}$), i.e., $\boldsymbol{c} = \boldsymbol{A}^\top \boldsymbol{y}$ for some $\boldsymbol{y} \geq \boldsymbol{0}$. Now we expand the vector $\boldsymbol{y}_{S_0}$ to $\boldsymbol{y}$ by letting $y_{i \notin S_0} = 0$ gives

$$\boldsymbol{y}_{S_0} = \boldsymbol{A}_{S_0}^{-\top}\boldsymbol{c} \iff \boldsymbol{y} = \boldsymbol{A}^{-\top}\boldsymbol{c} \iff \boldsymbol{c} = \boldsymbol{A}^\top \boldsymbol{y}.$$

### 5.3 Other cases not discussed

- Large perturbation on $\boldsymbol{b}$
- Change in $\boldsymbol{c}$
- Change in $\boldsymbol{A}$

### 5.4 Complementary slackness

- Primal constraint $i$ is tight (active), then dual variable $i$ is slack (passive)
- Primal constraint $i$ is slack (passive), then dual variable $i$ is tight (active)
- Dual constraint $i$ is tight (active), then primal variable $i$ is slack (passive)
- Dual constraint $i$ is slack (passive), then primal variable $i$ is tight (active)

## 6 Duality

**Definition 6.1: Symmetric primal-dual form**

$$P \;:\; \begin{array}{cl} \max\limits_{\boldsymbol{x}\in\mathbb{R}^n} & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \\ & \boldsymbol{x} \ge \boldsymbol{0}_n \end{array} \text{(Symmetric primal form)} \qquad D \;:\; \begin{array}{cl} \min\limits_{\boldsymbol{y}\in\mathbb{R}^m} & \boldsymbol{b}^\top \boldsymbol{y} \\ \text{s.t.} & \boldsymbol{A}^\top\boldsymbol{y} \ge \boldsymbol{c} \\ & \boldsymbol{y} \ge \boldsymbol{0}_m \end{array} \text{(Symmetric dual form)}$$

**Theorem 6.1** (Weak duality)**.** *For any $(\boldsymbol{u}\in\mathbb{R}^n, \boldsymbol{v}\in\mathbb{R}^m)$ that $\boldsymbol{u}$ is a sol. to $P$ and $\boldsymbol{v}$ is a sol. to $D$, we have*

$$p = \boldsymbol{c}^\top\boldsymbol{u} \le \boldsymbol{b}^\top\boldsymbol{v} = d.$$

- Dual feasible cost is always above primal feasible cost
- The theorem is true for any feasible $(\boldsymbol{u}, \boldsymbol{v})$, so it includes the case for $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ and we have $\boldsymbol{c}^\top\boldsymbol{x}^* \le \boldsymbol{b}^\top\boldsymbol{y}^*$.

*Proof.* As $\boldsymbol{A}^\top\boldsymbol{v} \ge \boldsymbol{c}$ in D and $\boldsymbol{u} \ge \boldsymbol{0}$ in P, we have $\boldsymbol{c}^\top\boldsymbol{u} \le (\boldsymbol{A}^\top\boldsymbol{v})\boldsymbol{u} = \boldsymbol{v}^\top\boldsymbol{A}\boldsymbol{u}$. Then by $\boldsymbol{A}\boldsymbol{u} \le \boldsymbol{b}$ in P and $\boldsymbol{v} \ge \boldsymbol{0}$ in D, we have $\boldsymbol{v}^\top\boldsymbol{A}\boldsymbol{u} \le \boldsymbol{v}^\top\boldsymbol{b}$. □

**Implications**

- We have it for free: weak duality always holds for feasible points.
- If $\boldsymbol{c}^\top\boldsymbol{x} = \boldsymbol{b}^\top\boldsymbol{y}$ then $(\boldsymbol{x}, \boldsymbol{y})$ are optimal.
- P unbounded implies D no solution: $+\infty = \boldsymbol{c}^\top\boldsymbol{x} \le \boldsymbol{b}^\top\boldsymbol{y}$
- By symmetry, D unbounded implies P no solution

**Theorem 6.2** (Strong duality)**.** *For any $(\boldsymbol{u}\in\mathbb{R}^n, \boldsymbol{v}\in\mathbb{R}^m)$ that $\boldsymbol{u}$ is a solution to $P$ and $\boldsymbol{v}$ is a solution to $D$, we have*

$$p = \boldsymbol{c}^\top\boldsymbol{u} = \boldsymbol{b}^\top\boldsymbol{v} = d.$$

*Proof.* Treat it as fact. Not the focus in this course. □

### 6.1 Karush-Kuhn-Tucker condition

The primal-dual problem pair and their slack form

$$P \;:\; \begin{array}{cl} \max\limits_{\boldsymbol{x}\in\mathbb{R}^n} & \boldsymbol{c}^\top\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \text{ (Symmetric primal form)} \\ & \boldsymbol{x} \ge \boldsymbol{0}_n \end{array} \qquad D \;:\; \begin{array}{cl} \min\limits_{\boldsymbol{y}\in\mathbb{R}^m} & \boldsymbol{b}^\top\boldsymbol{y} \\ \text{s.t.} & \boldsymbol{A}^\top\boldsymbol{y} \ge \boldsymbol{c} \text{ (Symmetric dual form)} \\ & \boldsymbol{y} \ge \boldsymbol{0}_m \end{array}$$

$$P' \;:\; \begin{array}{cl} \max\limits_{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{u}\in\mathbb{R}^n} & \boldsymbol{c}^\top\boldsymbol{x} + \boldsymbol{0}_n^\top\boldsymbol{u} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{u} = \boldsymbol{b} \\ & \boldsymbol{x} \ge \boldsymbol{0}_n,\, \boldsymbol{u} \ge \boldsymbol{0}_n \end{array} \qquad D' \;:\; \begin{array}{cl} \min\limits_{\boldsymbol{y}\in\mathbb{R}^m,\boldsymbol{v}\in\mathbb{R}^m} & \boldsymbol{b}^\top\boldsymbol{y} + \boldsymbol{0}_m^\top\boldsymbol{v} \\ \text{s.t.} & \boldsymbol{A}^\top\boldsymbol{y} - \boldsymbol{v} = \boldsymbol{c} \\ & \boldsymbol{y} \ge \boldsymbol{0}_m,\, \boldsymbol{v} \ge \boldsymbol{0}_m \end{array}$$

If $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{u}^*, \boldsymbol{v}^*)$ is the solution to $P'$ and $D'$, then

$$\begin{array}{ll} \boldsymbol{x}^* \ge \boldsymbol{0}_n, \underbrace{\boldsymbol{u}^* \ge \boldsymbol{0}_n, \boldsymbol{A}\boldsymbol{x}^* + \boldsymbol{u}^* = \boldsymbol{b}}_{\implies \boldsymbol{A}\boldsymbol{x}^* \le \boldsymbol{b}} & \text{Primal feasibility} \\ \boldsymbol{y}^* \ge \boldsymbol{0}_m, \underbrace{\boldsymbol{v}^* \ge \boldsymbol{0}_m, \boldsymbol{A}^\top\boldsymbol{y}^* - \boldsymbol{v}^* = \boldsymbol{c}}_{\implies \boldsymbol{A}^\top\boldsymbol{y}^* \ge \boldsymbol{c}} & \text{Dual feasibility} \\ x_i v_i = 0, u_i y_i = 0 & \text{Complementary slackness} \end{array}$$

**Other commonly seen notations of complementary slackness** The expression $x_i v_i = 0$ is equivalent to

- $\boldsymbol{x} \perp \boldsymbol{v}$ ($\boldsymbol{x}$ is perpendicular to $\boldsymbol{v}$)
- $\boldsymbol{x}^\top\boldsymbol{v} = 0$ ($\boldsymbol{x}$ and $\boldsymbol{v}$ have zero dot product)
- $\boldsymbol{x}$ and $\boldsymbol{v}$ have complementary zero pattern

# 7 Example

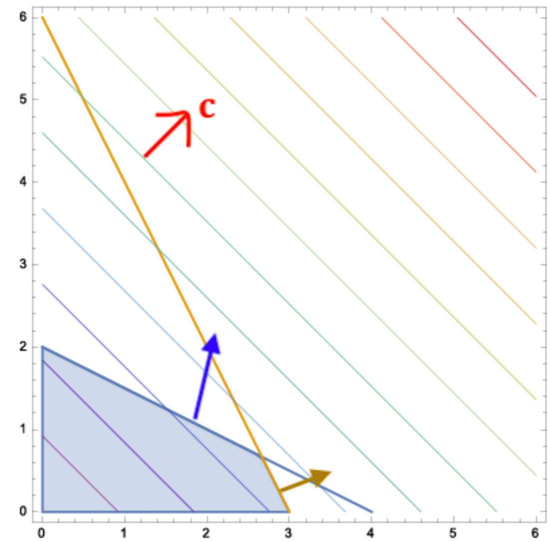
$$P_0 \ : \ \max_{x_1,x_2} \ x_1 + x_2 \ \text{ s.t. } \ x_1 + 2x_2 \leq 4, \ 2x_1 + x_2 \leq 6, \ x_1 \geq 0, x_2 \geq 0.$$

$$P_1 : \ \max_{\boldsymbol{x}\in\mathbb{R}^2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}^\top \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{s.t.} \quad \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 4 \\ 6 \\ 0 \\ 0 \end{bmatrix}$$

Solution of $P_1$i is $\boldsymbol{x}^* = \left[\frac{8}{3}, \ \frac{2}{3}\right] = \left[2.667, \ 0.667\right]$,

$p^* = \boldsymbol{c}^\top \boldsymbol{x}^* = \dfrac{10}{3} = 3.3333.$

Note that $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ and $u_1^* = u_2^* = 0.$




**The dual problem** We have

$$D_1 \ : \quad \begin{aligned} &\min_{\boldsymbol{y}\in\mathbb{R}^2} \begin{bmatrix} 4 \\ 6 \end{bmatrix}^\top \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &\text{s.t.} \quad \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \geq \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &\qquad y_1 \geq 0, y_2 \geq 0 \end{aligned}$$

Solution of $D_1$i is $\boldsymbol{y}^* = \left[\frac{1}{3}, \ \frac{1}{3}\right] = \left[0.3333, \ 0.3333\right]$,

and $d^* = \boldsymbol{b}^\top \boldsymbol{y}^* = \dfrac{10}{3} = 3.3333.$

Note $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} y_1^* \\ y_2^* \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v_1^* = v_2^* = 0.$

Furthermore,

- Weak and Strong duality: $p^* = d^* = 3.3333$
- $(\boldsymbol{x}^*, \boldsymbol{u}^*, \boldsymbol{y}^*, \boldsymbol{v}^*)$ are feasible
- Complementary slackness
  - $\boldsymbol{x}^* \geq \boldsymbol{0}_n$ is slack and $\boldsymbol{v}^* \geq \boldsymbol{0}$ is tight
  - $\boldsymbol{u}^* \geq \boldsymbol{0}_n$ is tight and $\boldsymbol{y}^* \geq \boldsymbol{0}$ is slack
- Active set of $\boldsymbol{x}^*$ is $\mathcal{S} = \{1, 2\}$, as only the first two rows of the constraints in $P_1$ is active.
- Geometry: At $\boldsymbol{x}^*$, $\boldsymbol{c} \in (\boldsymbol{A}_\mathcal{S})$, or equivalently, $\boldsymbol{c} = \boldsymbol{A}^\top \boldsymbol{y}^*$

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}}_{\boldsymbol{A}_\mathcal{S}^\top} \underbrace{\begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix}}_{\boldsymbol{y}^*} = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\boldsymbol{c}}$$

# 8 Linear integer programming

## 8.1 Unimodular matrix

**Definition 8.1: Linear Integer Program (LIP)**

Given $\boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{c} \in \mathbb{R}^n$, find $\boldsymbol{x} \in \mathbb{R}^n$ by solving

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{x} \in \mathbb{Z}^n \end{aligned}$$

- Integer constraint $\boldsymbol{x} \in \mathbb{Z}^n$ makes the problem NP-hard.
- Linear relaxation: instead of solving LIP, solve the relaxed LP by ignoring the interger constraint $\boldsymbol{x} \in \mathbb{Z}^n$.

**Definition 8.2: Unimodular matrix**

A matrix $\boldsymbol{M}$ is unimodular if $\boldsymbol{M}$ is a square integer matrix with $\det(\boldsymbol{M}) = \pm 1$.

**Facts 8.1.1** For linear equation $\boldsymbol{M}\boldsymbol{x} = \boldsymbol{b}$, if $\boldsymbol{M}$ is unimodular and $\boldsymbol{b} \in \mathbb{Z}^m$, then

- the equation has a solution
- the solution $\boldsymbol{x}^*$ is an integer vector

*Proof.* First $\det(\boldsymbol{M}) \neq 0$ so $\boldsymbol{M}^{-1}$ exists, then $\boldsymbol{x}^* = \boldsymbol{M}^{-1}\boldsymbol{b} = \underbrace{\frac{1}{\det(\boldsymbol{M})}}_{=\pm 1} \underbrace{\operatorname{adj}(\boldsymbol{M})\boldsymbol{b}}_{\in \mathbb{Z}^n} \in \mathbb{Z}^n$. □

**Facts 8.1.2** If $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{U}$, then

1. $\boldsymbol{A}\boldsymbol{B} \in \mathcal{U}$
2. $-\boldsymbol{A} \in \mathcal{U}$
3. $\boldsymbol{A}^{-1} \in \mathcal{U}$.

**Explanation / Proof** We have a fact: the determinant of product equals to the product of the determinants

$$\det(\boldsymbol{X}\boldsymbol{Y}) = \det(\boldsymbol{X}) \det(\boldsymbol{Y})$$

Then

1. $\det(\boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{A}) \det(\boldsymbol{B}) = (\pm 1) \times (\pm 1) = \pm 1$
2. $\det(-\boldsymbol{A}) = \det(-\boldsymbol{I}\boldsymbol{A}) = \det(-\boldsymbol{I}) \det(\boldsymbol{A}) = (-1) \times (\pm 1) = \mp 1 = \pm 1$
3. $1 = \det(\boldsymbol{I}) = \det(\boldsymbol{A}\boldsymbol{A}^{-1}) = \det(\boldsymbol{A}) \det(\boldsymbol{A}^{-1}) = (\pm 1) \det(\boldsymbol{A}^{-1})$

## 8.2 Fundamental Theorem of LP relaxation

- What's the big deal of unimodular matrix: a way to know when will LP relaxation works for solving LIP. If $\boldsymbol{A}_{\mathcal{S}} \in \mathcal{U}$, then the solution $\boldsymbol{x}^* = \boldsymbol{A}_{\mathcal{S}}^{-1}\boldsymbol{b}_{\mathcal{S}}$ solves LP and also the LIP.
- In general we don't know the active set $S(\boldsymbol{x}^*)$ before we solve the problem. So now the idea is that, instead of focusing on $\boldsymbol{A}_{\mathcal{S}}$ for a specific $\mathcal{S}$, we try to make make $\boldsymbol{A}_{\mathcal{S}} \in \mathcal{U}$ regardless of the possibilities of $S$, this lead to the idea of TUM.

**Definition 8.3: Total unimodular matrix (TUM)**

A matrix $\boldsymbol{M}$ is TU if all the square non-singular sub-matrices of $\boldsymbol{M}$ are in $\mathcal{U}$ or has determinant equal to 0.

**Example of TU matrix**

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \det(a_{11}) = 1, \quad \det(a_{22}) = -1, \quad \det(\boldsymbol{A}) = -1$$

$$\boldsymbol{B} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \ \det(b_{11}) = -1, \ \ \det(b_{21}) = \det(b_{22}) = \det(b_{13}) = 1, \ \det\left(\begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}\right) = -1, \ \det\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = -1$$

**Facts 8.2.1**

- $\boldsymbol{A}$ is TU, then $-\boldsymbol{A}$ is TU
- $\boldsymbol{A}$ is TU, then $\boldsymbol{A}^\top$ is TU
- $\boldsymbol{A}$ is TU, then $[\boldsymbol{A}\ \boldsymbol{I}], [\boldsymbol{A}\ \ -\boldsymbol{A}]$ are TU
- (Characterization of TUM) $\boldsymbol{A}$ is TU if and only if any one of the matrices $-\boldsymbol{A}, \boldsymbol{A}^\top, [\boldsymbol{A}\ \boldsymbol{I}], [\boldsymbol{A}\ \boldsymbol{A}]$ is TU.

**Theorem 8.1 (LP relaxation).** *Given $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{c} \in \mathbb{R}^n$, find $\boldsymbol{x} \in \mathbb{R}^n$ by solving*

$$\text{(LIP)} \quad \begin{array}{ll} \max\limits_{\boldsymbol{x}} & \boldsymbol{c}^\top \boldsymbol{x} \\ s.t. & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{array} \quad \xrightarrow{relax} \quad \text{(LP)} \quad \begin{array}{ll} \max\limits_{\boldsymbol{x}} & \boldsymbol{c}^\top \boldsymbol{x} \\ s.t. & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array}$$

*If $\boldsymbol{A}$ is TUM and $\boldsymbol{b} \in \mathbb{Z}^m$, then the LP has an integral solution.*

Proof by Cramer's rule: $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \iff \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} \iff x_i = \dfrac{\det(\boldsymbol{A}^i)}{\det(\boldsymbol{A})}$, $\boldsymbol{A}^i = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}, \boldsymbol{b}, \boldsymbol{a}_{i+1}, \ldots]$.

**Remarks**

- It is possible for LP with non-TUM $\boldsymbol{A}$ to have integral solution.
- Not all $0-1$ matrices are TUM. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ is not TUM because $\det = 2$
- If $\boldsymbol{A}$ is TUM and both $\boldsymbol{b}, \boldsymbol{c}$ are integral, then both primal and dual problems have integral solution. Why: $\boldsymbol{A}$ s TUM, then $\boldsymbol{A}^{-1}$ is TUM. For primal problem $\boldsymbol{b}$ is integral, for the dual problem, $\boldsymbol{c}$ is integral.

## 8.3 Integral polyhedron

**Definition 8.4: Integral polyhedron**

A polyhedron is integral if all extreme point is integral.

**Facts 8.3.1** If $\boldsymbol{A}$ is an $m \times n$ integral TUM, and vectors $\boldsymbol{b}, \boldsymbol{u}$ are integer vectors, then

$$\left\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\right\}$$
$$\left\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}\right\}$$
$$\left\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \boldsymbol{x} \geq 0\right\}$$
$$\left\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\right\}$$
$$\left\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, 0 \leq \boldsymbol{x} \geq \boldsymbol{u}\right\}$$

are all integral polyhedron.

**Facts 8.3.2** If $\boldsymbol{A}$ is an $m \times n$ integral matrix, the following are equivalent

- $\boldsymbol{A}$ is TUM
- The extreme points of $\left\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \boldsymbol{x} \geq 0\right\}$ are integral for all integral vectors $\boldsymbol{b}$
- All non-singular submatrix of $\boldsymbol{A}$ has an integral inverse.

**Facts 8.3.3**

- A LP problem with a TU matrix $\boldsymbol{A}$ yields an optimal solution in integers for any objective vector $\boldsymbol{c}$ and any integer vector $\boldsymbol{b}$.
- There are non-unimodular problems which yield integral optimal solution(s)
  - for any (possibly non-integral) vector $\boldsymbol{c}$ but only certain integer vector $\boldsymbol{b}$.
  - for any integer vector $\boldsymbol{b}$ but only certain (possibly non-integral) vector $\boldsymbol{c}$.

## 8.4 How to check a matrix is TU

Given a matrix $\boldsymbol{A}$, to test whether $\boldsymbol{A}$ is TU or not *by a efficient (easy) method* is a research problem.

- Brute force: by definition, check all the square non-singular sub-matrices of $\boldsymbol{A}$, see whether it has determinant equal to $-1, 0, +1$.
- If $\boldsymbol{A}$ has a million rows of millions columns, brute force is impossible.

**Camion's Characterization** Now we discuss one theorem to check a special class of matrices is TU.

**Definition ($(0, +1, -1)$ matrix)** A $(0, +1, -1)$ matrix only contains $0, +1$ or $-1$ in all its element.

**Definition (Eulerian matrix)** A matrix $\boldsymbol{A}$ is Eulerian if the sum of the elements in each row and each column is even.

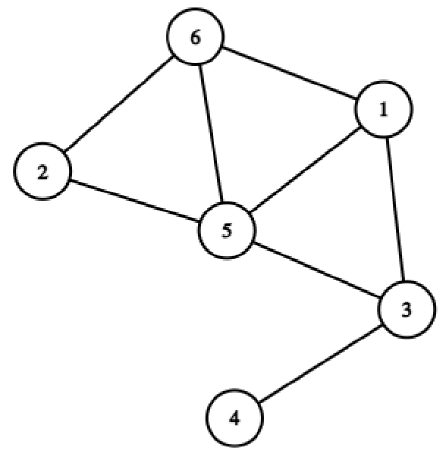
**Camion's theorem** A $(0, +1, -1)$ matrix $\boldsymbol{A}$ is TU if and only if the sum of the elements in each Eulerian square submatrix is a multiple of 4.

Ref: Camion P., "Matrices totalement unimodulaires et problemes combinatoires", Thèse, Université Libre de Bruxelles, Février, 1963

# 9 Some graph problems

**Terminology**

- Graph $G$ = a set of connected nodes.
- $V$: set of nodes
- $E$: set of edges. Each edge links two nodes.

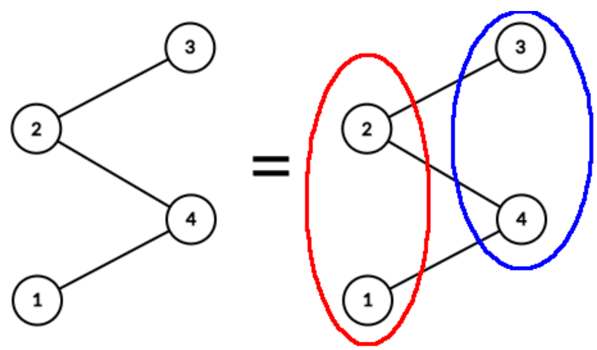



**Example** $V = \{1, 2, 3, 4, 5, 6\}$
$E = \{(1, 3), (1, 5), (1, 6), (2, 5), (2, 6), (3, 5), (3, 4), (5, 6)\}$

$$\text{Adjacency matrix} \quad \boldsymbol{A} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

**Example (Bipartite graph)** The "bi" means two and "partite" means part. A bipartite graph is a graph that all edges have an endpoint in each one of the two parts.
The adjacency matrix of a bipartite graph can always be written in block matrix form:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{B} \\ \boldsymbol{B}^\top & \boldsymbol{0} \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$




## 9.1 Min-Cost-Flow

In min-cost-flow problem,

- $G$ is a digraph (directed graph).
- each edge $(i, j) \in E$ has a cost $c_{ij}$ and a capacity constraint $u_{ij}$.
- There is one decision variable $x_{ij}$ per edge $(i, j) \in E$.
- $x_{ij}$: represents a flow of objects from $i$ to $j$.
- The cost of a flow $x_{ij}$ is $c_{ij}x_{ij}$.
- $s, k$: source and sink in the digraph
- Each node $j \in V \setminus \{s, d\}$ satisfies a flow constraint:

$$\sum_{k|(j,k)\in E} x_{jk} - \sum_{i|(i,j)\in E} x_{ij} = b_j$$

where $b_j$ is the amount of flow generated (if $b_{j>}0$) or consumed (if $b_j < 0$) by node $j$.

Given a digraph $G(V,E)$, the min-cost-flow problem is to find flows that minimize total cost subject to capacity and flow conservation. It can be written as a LIP:

$$\begin{aligned}
\max_{\boldsymbol{x}} \quad & \sum_{(i,j)\in E} c_{ij}x_{ij} \\
\text{s.t.} \quad & \sum_{k|(j,k)\in E} x_{jk} - \sum_{i|(i,j)\in E} x_{ij} = b_j \quad && \forall j \in V \\
& 0 \le x_{ij} \le u_{ij} && \forall (i,j) \in E \\
& \boldsymbol{x}_{ij} \in \mathbb{N} && \forall (i,j) \in E
\end{aligned} \tag{Min-Cost-Flow}$$

**Theorem 9.1.** *If all $b_j$ and $u_{i,j}$ are integer, then all $\bar{x}_{ij} \in \mathbb{N}$ for all basic feasible solution of Min-Cost-Flow $\bar{\boldsymbol{x}}$.*

*Sketch.* Let $\bar{\boldsymbol{x}}$ be a feasible sol. Now observe three facts:

1. (By flow conservation) If there is an incoming or outgoing edge at node $i$ for which the flow is fractional then there must be at least one additional incoming or outgoing edge at the same node for which the flow is fractional.

2. (By assumption of capacity being integral) If there is fractional flow along an edge, the capacity constraints for that edge cannot be tight (active/binding).

3. (By a fact in Section 3) A basic feasible solution can not be a convex combination of two other feasible solutions.

We prove by contradiction.

- Suppose that there is a basic feasible solution $\hat{\boldsymbol{x}}$ such that the flow $x_{ij}$ for some edge $(i,j)$ is not integer (i.e., it is a fractional value).

- By Fact 1, there must be a cycle of edges in the graph, each of which has fractional flow.

- Then, by Fact 2, a small flow of $\varepsilon$ can be added to or subtracted from the cycle, while retaining feasibility.

- Hence, $\hat{\boldsymbol{x}}$ is the convex combination of these two alternatives. So by Fact 3, $\hat{\boldsymbol{x}}$ is not a basic feasible solution.

□

**Implication of Min-Cost-Flow theorem** Basic feasible solutions are integer-valued $\implies$ if there is an optimal sol, there will be one that is integer-valued. Hence we can use LP to solve the LIP Min-Cost-Flow.

## 9.2 Shortest Path Problems

Given a graph $G(V,E)$.

- Each node represents a location (city).

- Each edge represents a road. Each edge is associated with a travel time $c_{ij}$.

- Task: we want to move from node $o$ (the origin) to node $d$ (the destination) in minimum time.

**Shortest path problem as Min-Cost-Flow** Suppose that for each $(i,j) \in E$, we take $x_{ij} \in \{0,1\}$ to be a binary variable to which we would assign a value of 1 if edge $(i,j)$ is to be part of the trip and 0 otherwise. Then, the trip with shortest travel time can be found by solving the a min-cost-flow problem with

- $b_o = 1$ (origin node only output)
- $b_d = -1$ (destination node only input)
- $b_i = 0$ for all $i \notin \{o,d\}$ (other nodes)
- $u_{ij} = 1$ for all $(i,j) \in E$
- $x_{ij} \in \{0,1\}$ as BIP constraint.

The Shortest path problem is a BIP, which is not linear. However, we can drop the BIP constraint and solve it as a Min-Cost-Flow using LP relaxation: as basic feasible solutions are integer-valued, hence if there exists an optimal sol (in this case if there is a path from node $o$ to node $d$), then there will be an integer-valued optimal solution. This means a binary-valued solution since each flow is constrained to be between 0 and 1.

## 9.3 Other graph problems

- Max-Flow
- Min-Cut
- Max-Cut

## 9.4 Linear assignment problem

Assignment matrix $\boldsymbol{X}$, a matrix:

$$x_{ij} = \begin{cases} 1 & i \text{ is assigned to } j \\ 0 & i \text{ is not assigned to } j \end{cases}$$

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \sum_{ij} c_{ij} x_{ij} && \text{minimize total allocation cost} \\ \text{s.t.} \quad & \sum_{j} x_{ij} = 1 && \text{each j is assigned once} \\ & \sum_{i} x_{ij} = 1 && \text{assign each i to one j} \end{aligned}$$

## 9.5 Stable marriage problem

Given $n$ men and $n$ women, where each person has ranked all members of the opposite sex in order of preference, marry the men and women together such that there are no two people of opposite sex who would both rather have each other than their current partners. When there are no such pairs of people, the set of marriages is called stable.

For more, see Vate, John H. Vande. “Linear programming bring marital bliss.” Operations Research Letters 8.3 (1989): 147-153