

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

Instructor: Andersen Ang

Homepage: angms.science/CO327.htmlThursday 31st March, 2022

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

Contents

1	Standard form and canonical form	2
1.1	Type of LP	2
2	Equivalence and conversion between forms	3
2.1	Solving system of linear inequality: Fourier-Motzkin elimination	3
3	Basic linear geometry	4
3.1	Hyperplane and Half-space	4
3.2	Polyhedron	4
3.3	Fundamental Theorem of LP	6
4	Basic properties of LP	6
4.1	Active and passive constraint	6
4.2	c at optimal point will be inside the cone formed by the active rows in A	7
5	Perturbation and sensitivity	7
5.1	Sensitivity of LP with perturbation on b	7
5.2	Small perturbation	8
5.3	Complementary slackness	8
6	Duality	8
7	Linear integer programming	9
7.1	Unimodular matrix	9
7.2	Fundamental Theorem of LP relaxation	9
8	Some graph problems	10
8.1	Min-Cost-Flow	10
8.2	Shortest Path Problems	12
8.3	Other graph problems	12

1 Standard form and canonical form

Definition 1.1: Standard form and Canonical form

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

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

The primal objective value $p := \mathbf{c}^\top \mathbf{x}$.

- Notation: x denotes scalar and \mathbf{x} denotes vector.
The symbol \geq is taken element-wise by convention.
- The function $\mathbf{c}^\top \mathbf{x}$ is called objective function and it is a linear function.
- In standard form, usually \mathbf{A} is singular / non-invertible. Why: otherwise $\mathbf{A}\mathbf{x} = \mathbf{b} \implies \mathbf{x} = \mathbf{A}^\top \mathbf{b}$ is unique \implies we solved LP.
- In canonical form, the constraint $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ describes a “polyhedron membership” \longrightarrow linear geometry.
- **Definition (Feasibility)** A point \mathbf{x} is *feasible* if it satisfies the constraint.
- **Definition (Optimal point)** A point \mathbf{x}^* is *optimal*, or called a *solution*, if $p^* := \mathbf{c}^\top \mathbf{x}^* \geq p = \mathbf{c}^\top \mathbf{x}$ for all feasible \mathbf{x} .
- Fact: a feasible point need not to be an optimal point.
- LP with $\mathbf{c} = \mathbf{0}$ 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

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

The dual objective value) $d := \mathbf{b}^\top \mathbf{y}$.

- 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.

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 ∞ .

2 Equivalence and conversion between forms

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

$$\begin{array}{ccc}
 \min_{\mathbf{x}} \mathbf{c}_S^\top \mathbf{x} & & \min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x} & & \min_{\mathbf{x}} \mathbf{c}_C^\top \mathbf{x} \\
 \text{s.t. } \mathbf{A}_S \mathbf{x} = \mathbf{b}_S & & \text{s.t. } \mathbf{A} \mathbf{x} \leq \mathbf{b} & & \text{s.t. } \mathbf{A}_C \mathbf{x} \leq \mathbf{b}_C \\
 \mathbf{x} \geq \mathbf{0}_S & \iff & \mathbf{x} \geq \mathbf{0} & \iff & \\
 \text{Standard form} & & \text{Symmetric primal form} & & \text{Canonical form}
 \end{array}$$

The matrices $\mathbf{A}, \mathbf{A}_S, \mathbf{A}_C$ in the above forms are in general different (same for \mathbf{b}, \mathbf{c}). Technically, solution \mathbf{x} in these forms are not the same (because of different dimensions in $\mathbf{c}, \mathbf{c}_S, \mathbf{c}_C$) 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 $0 \leq s_i \leq u_i - l_i$. (Tutorial 1 slide 22)
- $x_i = b_i$ is equivalent to $x_i \leq b_i$ AND $x_i \geq b_i$.
- 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 $\mathbf{A} \mathbf{x} \leq \mathbf{b}$ has a solution.

Fourier-Motzkin elimination method

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

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

3 Basic linear geometry

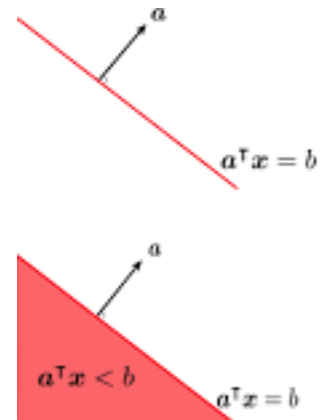
3.1 Hyperplane and Half-space

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

- \mathbf{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 \mathbf{a} .
- If $b = 0$, the hyperplane has two normal vectors \mathbf{a} and $-\mathbf{a}$.

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

- The inequality can be strict $>$ or non-strict \geq .
- Half-space is not affine.



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

- In $\mathbf{a}^\top \mathbf{x} = b$, the value b controls the location of the plane, it also “label” the plane by its value.
- Moving along the direction \mathbf{a} increases the value b . This implies moving \mathbf{x} along the \mathbf{c} in the primal objective function $\mathbf{c}^\top \mathbf{x}$ increase the value of p . So you want to move along the direction \mathbf{c} as much as possible in the LP problem.

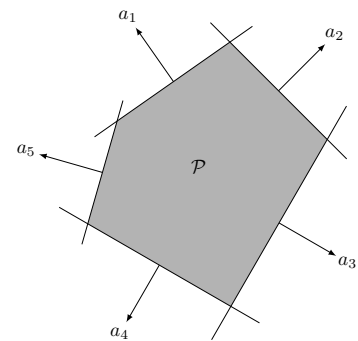
3.2 Polyhedron

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

$$\mathcal{P}(\mathbf{A}, \mathbf{b}) := \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b} \}.$$

That is, it is a set defined by the system of linear inequalities $\mathbf{A}\mathbf{x} \leq \mathbf{b}$. 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

- **Existence** of the optimal point: \mathcal{P} can be bounded or unbounded.
- **Uniqueness** of the optimal point: \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 $\mathbf{x} \in \mathcal{X}$ we have $\|\mathbf{x}\| \leq M$

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

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

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

Definition 3.7 (Internal line segment). Given two points \mathbf{x}, \mathbf{y} . A line $\ell(\mathbf{x}, \mathbf{y})$ is the set of all convex combination of \mathbf{x}, \mathbf{y} . The two ending points of $\ell(\mathbf{x}, \mathbf{y})$ is exactly \mathbf{x} and \mathbf{y} . Mathematically,

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

Definition 3.8 (Ray). Given two points \mathbf{x} . A ray $r(\mathbf{x})$ is the set of all conic combination of \mathbf{x} . Mathematically,

$$r(\mathbf{x}) := \{\mathbf{v} \mid \mathbf{v} = \lambda \mathbf{x}, \lambda \geq 0\}.$$

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

Definition 3.10 (Vertex). Let \mathcal{P} be a polyhedron. A vector $\mathbf{x} \in \mathcal{P}$ is a vertex of \mathcal{P} if there is a objective function \mathbf{c} that \mathbf{x} is the unique optimal point whesectionn maximizing over \mathcal{P} .

Mathematically, $\mathbf{x} \in \mathcal{P}$ is a vertex of \mathcal{P} if there exists some \mathbf{c} such that $\mathbf{c}^\top \mathbf{x} > \mathbf{c}^\top \mathbf{u}$ for all $\mathbf{x} \neq \mathbf{u} \in \mathcal{P}$.

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

- \mathbf{x} is a vertex.
- \mathbf{x} is an extreme point.
- \mathbf{x} is a basic feasible solution.

Corollary 3.1. Any polyhedron has finitely many extreme points

Definition 3.11 (Cone). A set $\mathcal{C} \subset \mathbb{R}^n$ is a cone if $\mathbf{x} \in \mathcal{C} \implies \lambda \mathbf{x} \in \mathcal{C}$ 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.

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

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

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

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

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

3.3 Fundamental Theorem of LP

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

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

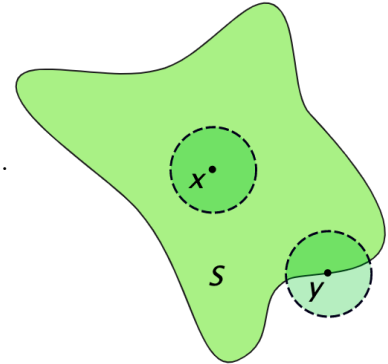
- $\text{int}(\mathcal{P})$: interior of \mathcal{P} .

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

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

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

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



4 Basic properties of LP

4.1 Active and passive constraint

Consider canonical form: $\min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x}$ s.t. $\mathbf{A}\mathbf{x} \leq \mathbf{b}$.

Definition 4.1: Active and passive constraint, active set

For each row of the constraint $\mathbf{a}_i^\top \mathbf{x} \leq b_i$,

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

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

Let S be the active set of a feasible point \mathbf{x} .

Recall that optimal point for a LP is always on the boundary of \mathcal{P} , so $S(\mathbf{x}^*)$ is always nonempty and

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

where \mathbf{A}_S is the matrix \mathbf{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 $\bar{\mathbf{x}} \in \mathbb{R}^N$, then $\bar{\mathbf{x}}$ is called a basic solution.
- For a basic solution $\bar{\mathbf{x}}$, if it is feasible, then it is called a basic feasible solution.

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

Other words to describe active and passive

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

4.2 \mathbf{c} at optimal point will be inside the cone formed by the active rows in \mathbf{A}

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

1. We know that moving \mathbf{x} along the direction \mathbf{c} will increase the value of p so we want to move along \mathbf{c} as much as possible.
2. Moving \mathbf{x} along \mathbf{c} will ultimately hit the boundary of \mathcal{P} : by Theorem 3.1, the optimal point \mathbf{x}^* is always on $\text{bdry}(\mathcal{P})$, either being an extreme point (vertex) or on a facet.
3. \mathbf{x}^* on $\text{bdry}(\mathcal{P})$ will introduce an active set $\mathcal{S}(\mathbf{x}^*)$ and passive set for the constraints.
4. As a result, $\mathbf{c} \in \text{cone}(\mathbf{A}_{\mathcal{S}(\mathbf{x}^*)}^\top)$
5. Using conic combination: $\mathbf{c} \in \text{cone}(\mathbf{A}_{\mathcal{S}(\mathbf{x}^*)}^\top) \iff \exists \mathbf{y}, \mathbf{y} \geq \mathbf{0}$ such that $\mathbf{c} = \mathbf{A}^\top \mathbf{y}$, where $\mathbf{y}_{i \notin \mathcal{S}} = 0$ and $\mathbf{y}_{i \in \mathcal{S}} > 0$.
6. $\mathbf{c} = \mathbf{A}^\top \mathbf{y}$ is related to duality.

5 Perturbation and sensitivity

5.1 Sensitivity of LP with perturbation on \mathbf{b}

Consider two LPs

$$(\mathcal{P}_0) : \begin{array}{ll} \min_{\mathbf{x} \in \mathbb{R}^n} & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{Ax} \leq \mathbf{b} \end{array} \quad \text{and} \quad (\mathcal{P}_1) : \begin{array}{ll} \min_{\mathbf{x} \in \mathbb{R}^n} & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{Ax} \leq (\mathbf{b} + \Delta) \end{array}$$

where $\Delta \in \mathbb{R}^m$ is a perturbation vector that changes \mathbf{b} a little bit. Suppose both (\mathcal{P}_0) and (\mathcal{P}_1) has a solution. Let \mathbf{x}_0^* be a solution of (\mathcal{P}_0) and let \mathbf{x}_1^* be a solution of (\mathcal{P}_1) .

Q: What is the relationship between \mathbf{x}_0^* and \mathbf{x}_1^* regarding the perturbation Δ ?

Let S_0 be the active set of \mathbf{x}_0^* for solving (\mathcal{P}_0) , so we have $\mathbf{A}_{S_0} \mathbf{x}_0^* = \mathbf{b}_{S_0}$.

Let S_1 be the active set of \mathbf{x}_1^* for solving (\mathcal{P}_1) , so we have $\mathbf{A}_{S_1} \mathbf{x}_1^* = (\mathbf{b} + \Delta)_{S_1} = \mathbf{b}_{S_1} + \Delta_{S_1}$.

Now there are two possibilities

- Case 1. The perturbation Δ 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 Δ 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

$$\begin{aligned} \mathbf{x}_0^* &= \mathbf{A}_{S_0}^{-1} \mathbf{b}_{S_0} & p(\mathbf{x}_0^*) &= \mathbf{c}^\top \mathbf{x}_0^* \\ \mathbf{x}_1^* &= \mathbf{A}_{S_1}^{-1} (\mathbf{b} + \Delta)_{S_1} & p(\mathbf{x}_1^*) &= \mathbf{c}^\top \mathbf{x}_1^* \\ &= \mathbf{A}_{S_0}^{-1} (\mathbf{b} + \Delta)_{S_0} & &= \mathbf{c}^\top (\mathbf{x}_0^* + \mathbf{A}_{S_0}^{-1} \Delta_{S_0}) \\ &= \mathbf{A}_{S_0}^{-1} \mathbf{b}_{S_0} + \mathbf{A}_{S_0}^{-1} \Delta_{S_0} & &= \mathbf{c}^\top \mathbf{x}_0^* + \mathbf{c}^\top \mathbf{A}_{S_0}^{-1} \Delta_{S_0} \\ &= \mathbf{x}_0^* + \mathbf{A}_{S_0}^{-1} \Delta_{S_0} & &= p(\mathbf{x}_0^*) + \mathbf{c}^\top \mathbf{A}_{S_0}^{-1} \Delta_{S_0} \end{aligned}$$

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

$$\frac{\Delta p}{\Delta} = \frac{\partial \mathbf{y}_{S_0}^\top \Delta}{\partial \Delta} = \mathbf{y}_{S_0},$$

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

Note the the above analysis assume we know the active set $S_0(\mathbf{x}_0^*)$. 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 \mathbf{y} : we can set $y_{i \notin S_0} = 0$.

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

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

5.3 Complementary slackness

See example in the class.

- 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

$$P : \begin{array}{ll} \max_{\mathbf{x} \in \mathbb{R}^n} & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{A} \mathbf{x} \leq \mathbf{b} \quad (\text{Symmetric primal form}) \\ & \mathbf{x} \geq \mathbf{0}_n \end{array} \quad D : \begin{array}{ll} \min_{\mathbf{x} \in \mathbb{R}^n} & \mathbf{b}^\top \mathbf{y} \\ \text{s.t.} & \mathbf{A}^\top \mathbf{y} \geq \mathbf{c} \quad (\text{Symmetric dual form}) \\ & \mathbf{y} \geq \mathbf{0}_m \end{array}$$

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

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

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

Proof. As $\mathbf{A}^\top \mathbf{v} \geq \mathbf{c}$ in D and $\mathbf{u} \geq \mathbf{0}$ in P , we have $\mathbf{c}^\top \mathbf{u} \leq (\mathbf{A}^\top \mathbf{v})^\top \mathbf{u} = \mathbf{v}^\top \mathbf{A} \mathbf{u}$. Then by $\mathbf{A} \mathbf{u} \leq \mathbf{b}$ in P and $\mathbf{v} \geq \mathbf{0}$ in D , we have $\mathbf{v}^\top \mathbf{A} \mathbf{u} \leq \mathbf{v}^\top \mathbf{b}$. \square

Implications

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

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

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

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

7 Linear integer programming

7.1 Unimodular matrix

Definition 7.1 (LIP). *Given $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^n$, find $\mathbf{x} \in \mathbb{R}^n$ by solving*

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

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

Definition 7.2. *A matrix \mathbf{M} is unimodular if \mathbf{M} is a square integer matrix with $\det(\mathbf{M}) = \pm 1$.*

Fact For linear equation $\mathbf{M}\mathbf{x} = \mathbf{b}$, if \mathbf{M} is unimodular and $\mathbf{b} \in \mathbb{Z}^m$, then

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

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

Fact If $\mathbf{A} \in \mathcal{U}, \mathbf{B} \in \mathcal{U}$ then, $\mathbf{AB} \in \mathcal{U}, -\mathbf{A} \in \mathcal{U}$ and $\mathbf{A}^{-1} \in \mathcal{U}$.

7.2 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. The solution $\mathbf{x}^* = \mathbf{A}_S^{-1}\mathbf{b}_S$ solves LP and also the LIP.
- In general we don't know the active set $S(\mathbf{x}^*)$ before we solve the problem. The idea is then make $\mathbf{A}_S \in \mathcal{U}$ regardless of the possibilities of S , this lead to TUM

Definition 7.3 (Total unimodular matrix (TUM)). *A matrix \mathbf{M} is TUM if all the square non-singular submatrices of \mathbf{M} are in \mathcal{U} .*

- \mathbf{A} is TUM, then $-\mathbf{A}$ is TUM
- \mathbf{A} is TUM, then \mathbf{A}^\top is TUM
- \mathbf{A} is TUM, then $[\mathbf{A} \ \mathbf{I}]$, $[\mathbf{A} \ -\mathbf{A}]$ are TUM

Theorem 7.1 (LP relaxation and TUM). Given $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^n$, find $\mathbf{x} \in \mathbb{R}^n$ by solving

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

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

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

Remarks

- It is possible for LP with non-TUM \mathbf{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 \mathbf{A} is TUM and both \mathbf{b}, \mathbf{c} are integral, then both primal and dual problems have integral solution.
Why: \mathbf{A} is TUM, then \mathbf{A}^{-1} is TUM. For primal problem \mathbf{b} is integral, for the dual problem, \mathbf{c} is integral.

8 Some graph problems

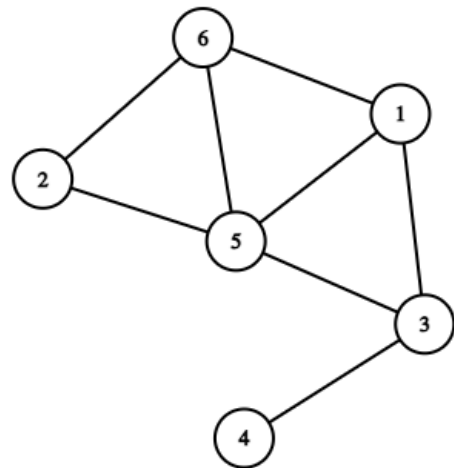
Terminology

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

$$V = \{1, 2, 3, 4, 5, 6\}$$

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

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



8.1 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:

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

Theorem 8.1. *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 $\bar{\mathbf{x}}$.*

Sketch. Let $\bar{\mathbf{x}}$ 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 $\hat{\mathbf{x}}$ 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 ε can be added to or subtracted from the cycle, while retaining feasibility.
- Hence, $\hat{\mathbf{x}}$ is the convex combination of these two alternatives. So by Fact 3, $\hat{\mathbf{x}}$ 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.

8.2 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.

8.3 Other graph problems

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