University of Waterloo

Fall 2012

CO 255 - Introduction to Optimization

Author:
David McLaughlin

Instructor:
Jim Geelen

These notes are presented without any guaranty of any kind. They might contain material not seen in the course and/or omit material seen in the course. These notes might also contain typos and errors.

Last updated: November 23, 2012

Contents

1 Duality **1**
1.1 Complementary Slackess . . . 2
1.2 Standard Inequality Form . . . 2
1.3 Standard Equality Form . . . 2
1.4 Basic Solutions . . . 3
1.5 Simplex method . . . 4
1.6 Optimality . . . 4
1.7 Unboundedness . . . 5
1.8 Perturbation Method . . . 5

2 Midterm Review **6**
2.1 Seperating Hyperplane Theorems . . . 6
2.2 Polyhedral Theory . . . 7
2.3 Applications . . . 7
2.4 Linear Programming . . . 7
2.5 Duality . . . 8
2.6 Application of Duality . . . 8

3 Integer Programming **9**
3.1 Totally Unimodular Matrices . . . 9
3.2 Min-Cost Perfect Matching . . . 10
3.3 Directed Graphs . . . 12

4 Complexity Theory **12**
4.1 Formalism . . . 13

5 Nonlinear Optimization **14**
5.1 Compactness . . . 14
5.2 Certifying Optimality . . . 14

1 Duality

Consider the LP

$$(P)\begin{cases} \text{maximize} & c^t x \\ \text{subject to} & Ax \le b \end{cases}$$

If $y \in \mathbb{R}^m$ and $y \ge 0$ then

$$y^t Ax \le y^t b$$

is a valid inequality for (P).

If $y^t A = c^t$, then

$$c^t \le y^t b = b^t y$$

The dual of (P) is

$$(D)\begin{cases} \text{minimize} & b^t y \\ \text{subject to} & A^t y = c, y \ge 0 \end{cases}$$

Weak Duality Theorem :

If $x \in \mathbb{R}^n$ is feasible for (P) and $y \in \mathbb{R}^m$ is feasible for (D), then $c^t x \le b^t y$.

Proof.

$$c^t x = (y^t A)x = y^t(Ax) \le y^t b = b^t y.$$

□

Corollary If (P) is unbounded, then (D) is infeasible.

Proof. Contrapositive is obvious. □

Corollary If (D) is unbounded, then (P) is infeasible.

Corollary If $\tilde{x}$ is feasible for (P), $\tilde{y}$ is feasible for (D) and $c^t\tilde{x} = b^t\tilde{y}$, then $\tilde{x}$ is optimal for (P) and $\tilde{y}$ is optimal for (D).

Strong Duality Theorem :

If (P) has and optimal solution $\tilde{x}$, then (D) has an optimal solution $\tilde{y}$, and $c^t\tilde{x} = b^t\tilde{y}$

Proof. Consider the system

$$(1)\begin{cases} -c^t x + b^t y & \le & 0 \\ Ax & \le & b \\ -A^t y & \le & -c \end{cases}.$$

If $\tilde{x}, \tilde{y}$ satisfy (1), then $\tilde{x}$ is feasible for (P) and $c^t\tilde{x} \ge b^t\tilde{y}$. By the weak duality theorem, $c^t\tilde{x} = b^t\tilde{y}$. So $\tilde{x}$ is optimal for (P) and $\tilde{y}$ is optimal for (D) as required. So we may assume that (1) has no solution.

Claim: If (1) has no solution the there exist $\bar{x} \in \mathbb{R}^n$, $\bar{y} \in \mathbb{R}^m$, and $\bar{z} \in \mathbb{R}$ satisfying

$$(2)\begin{cases} -c^t\bar{x} + b^t\bar{y} & \le & 0 \\ A\bar{x} & \le & \bar{z}b \\ A^t\bar{y} & = & \bar{z}c \\ \bar{y} & \ge & 0 \\ \bar{z} & \ge & 0 \end{cases}$$

This claim holds by Farkas' Lemma.

Consider the solution $(\bar{x}, \bar{y}, \bar{z})$ to (2).

Case 1: $\bar{z} > 0$. We can scale $(\bar{x}, \bar{y}, \bar{z})$ so that $\bar{z} = 1$. Now $(\bar{x}, \bar{y})$ satisfies (1), a contradiction.

Case 2: $\bar{z} = 0$. Now $\bar{y}^t A = 0$ and $\bar{y} \ge 0$. Since (P) is feasible, $\bar{y}^t b \ge 0$. That is $b^t\bar{y} \ge 0$. Moreover $A\bar{x} \le 0$. However (P) is bounded, so $c^t\bar{x} \le 0$. So $-c^t\bar{x} + b^t\bar{y} \ge 0$, contradicting (2).

Note

		(D) infeasible	unbounded	optimal
(P)	infeasible	*yes(exercise)*	*yes*	*no(exercise)*
	unbounded	*yes*	*no*	*no*
	optimal	*no*	*no*	*yes*

Case 3: $\bar{z} < 0$. □

Consider the following LPs:

$$(P1)\begin{cases} \text{maximize} & c^t x \\ \text{subject to} & Ax \leq b \end{cases}$$

$$(P2)\begin{cases} \text{maximize} & c^t(x^1 - x^2) \\ \text{subject to} & A(x^1 - x^2) \leq b \\ & x^1, x^2 \geq 0 \end{cases}$$

$$(P3)\begin{cases} \text{maximize} & c^t(x^1 - x^2) \\ \text{subject to} & A(x^1 - x^2) + s = b \\ & x^1, x^2, s \geq 0 \end{cases}$$

they are all equivalent.

1.1 Complementary Slackess

Consider

$$(P) \quad \max(c^t x : Ax \leq b)$$

and its dual

$$(D) \quad \min(b^t x : A^t y = c, y \geq 0).$$

If $\bar{x}$ is feasible for (P) and $\bar{y}$ is feasible for (D), then

$$\begin{aligned} b^t\bar{y} - c^t\bar{x} &= \bar{y}b - y^t A\bar{x} \\ &= \bar{y}^t(b - A\bar{x}) \\ &= \sum_{i=1}^{m} \bar{y}_i(b_i - \sum_{j=1}^{n} A_{ij}\bar{x}_j) \end{aligned}$$

Now $\bar{y}_i(b_i - \sum_{j=1}^{n} A_{ij}\bar{x}_j) \geq 0$ and equality holds if and only if either $\bar{y}_i = 0$ or $\sum_{j=1}^{n} A_{ij}\bar{x}_j = b_i$.

Complementary Slackness Theorem :
Let $\bar{x}$ be feasible for (P) and $\bar{y}$ be feasible for (D). Then $c^t\bar{x} = b^t\bar{y}$ if and only if for each $i \in \{1, \dots, n\}$, either $\bar{y}_i = 0$ or $[A_{i1}, \dots, A_{in}]\bar{x} = b_i$.

Proof. See above ☐

1.2 Standard Inequality Form

Let $\bar{x}$ be feasible for

$$(PSI) \quad \max(c^t x : Ax \leq b, x \geq 0)$$

and $\bar{y}$ be feasible for

$$(DSI) \quad \min(b^t x : A^t y \leq c, y \geq 0).$$

Then $c^t\bar{x} = b^t\bar{y}$ if and only if

1. For each $i \in \{1, \dots, n\}$ either $\bar{y}_i = 0$ or $[A_{i1}, \dots, A_{in}]\bar{x} = b_i$ and
2. For each $j \in \{1, \dots, m\}$ either $\bar{x}_i = 0$ or $[A_{1j}, \dots, A_{mj}]\bar{y} = c_j$

Proof. Exersise ☐

1.3 Standard Equality Form

Let $\bar{x}$ be feasible for

$$(PSE) \quad \max(c^t x : Ax = b, x \geq 0)$$

and $\bar{y}$ be feasible for

$$(DSE) \quad \min(b^t x : A^t y \leq c).$$

Then $c^t\bar{x} = b^t\bar{y}$ if and only if For each $j \in \{1, \dots, m\}$ either $\bar{x}_i = 0$ or $[A_{1j}, \dots, A_{mj}]\bar{y} = c_j$

Proof. Rewrite (DSE) as $(DSE') = \max(-b^t y : -A^t y \leq -c)$ and then apply slackness theorem. ☐

1.4 Basic Solutions

Consider

$$(P)\begin{cases} \max & c^t x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{cases}$$

and its dual

$$(D)\begin{cases} \min & c^t x \\ \text{subject to} & A^t y \geq c \end{cases}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. We assume that $\text{rank}(A) = m$ (whitout loss of generality).

Notation: $A = [A_1, \dots, A_n]$ and for $B \subseteq \{1, \dots, n\}$, $A_B = [A_i : i \in B]$. We call B a basis if $|B| = m$ and rank $A_B = m$.

For a basis B,

1. there is a unique solution to $\begin{cases} Ax = b \\ x_j = 0 \text{ for all } j \notin B \end{cases}$, this is the basic soultion for B.
2. there is a unique $y \in \mathbb{R}^m$ satisfying $(A_B)^t y = c_B$, this is the basic dual solution.

If $\bar{x}$ is a the basic solution for B and $\bar{x} \geq 0$, then we call $\bar{x}$ a basic feasible solution.
If $\bar{y}$ is the basic dual solution for B and $A^t\bar{y} \geq c$, then we call $\bar{y}$ a basic dual feasible solution.

Optimality Theorem :
Let $\bar{x} \in \mathbb{R}^n$ be the basic solution for B and $\bar{y} \in \mathbb{R}^m$ be the basic dual solution for B. Then $c^t\bar{x} = b^t\bar{y}$. Moreover, if $\bar{x}$ is feasible for (P) and $\bar{y}$ is feasible for (D), then $\bar{x}$ is optimal for (P).

Proof.

$$\begin{aligned} b^t\bar{y} - c^t\bar{x} &= \bar{x}^t A^t \bar{y} - \bar{x}c && (1)\\ &= \bar{x}^t(A^t\bar{y} - c) && (2)\\ &= \bar{x}^t_B(A^t_B\bar{y} - c_B) && (3)\\ & && (4) \end{aligned}$$

□

Note that this proof works since $\bar{x}$ and $\bar{y}$ satisfy the complemetary slackness conditions.

Remarks:

1. $\bar{x} \in \mathbb{R}^n$ is an extreme point of (P) if and only if it is a basic feasible solution. (See assignment 2)
2. $\bar{y} \in \mathbb{R}^m$ is an extreme point of (D) if and only if it is a basic dual feasible solution. (See assigment 1, less obvious though)

Claim: A feasible solution for (P) is a basic feasible solution if and only if the columns of $[A_j : \bar{x}_j \neq 0]$ are linearly independent.

Proof. $(\Rightarrow)$ By definition
$(\Leftarrow)$ Any linearly independent set extends to a basis. □

1.5 Simplex method

$$(P)\begin{cases} \text{max} & c^t x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{cases}$$

rank$(A) = m$ and

$$(D)\begin{cases} \text{min} & b^t y \\ \text{subject to} & A^t y \geq c \end{cases}$$

Let $\bar{x}$ be a basic feasible solution for a basis B, let $\bar{y}$ be the basic dual solution for B, and let $\bar{v} = c^t\bar{x} = b^t\bar{y}$.
Recall: $(A_B)^t y = c_B$
Note that for any feasible x,

$$c^t x = c^t x - \bar{y}^t(Ax - b) = (c - A^t\bar{y})^t x + \bar{y}^t b = (c - A^t\bar{y})^t x + \bar{v}.$$

We can rewrite (P) as

$$(P)\begin{cases} \text{max} & \bar{c}^t x + \bar{v} \\ \text{subject to} & \bar{A}x = \bar{b} \\ & x \geq 0 \end{cases}$$

where

$$\bar{c} = c - A^t\bar{y},$$

$$\bar{A} = (A_B)^{-1}A, \text{ and}$$

$$\bar{b} = (A_B)^{-1}b.$$

Note that :

1. $\bar{A}_B = I$ so we may assume that the rows of $\bar{A}$ are indexed by the elements of B and that $\bar{b}$ is indexed by B
2. $\bar{x}_B = \bar{b}$
3. $\bar{c}_B = c_B - A_B^t\bar{y} = 0$
4. $\bar{y}$ is feasible for (D) if and only if $\bar{c} \leq 0$

1.6 Optimality

If $\bar{c} \leq 0$, then $\bar{x}$ is optimal for (P) and $\bar{y}$ is optimal for (D). (by (4)).

Suppose that $\bar{c}_j > 0$ for some j. (Note that $j \notin B$ by (2)). x_j is the entering variable.
Define $\bar{d} \in \mathbb{R}^n$ by

$$\bar{d}_i = \begin{cases} -\bar{a}_{ij} & : & i \in B \\ 1 & : & i = j \\ 0 & : & \text{otherwise} \end{cases}$$

Note that the unique solution to

$$\bar{A}x = \bar{b}$$

$$x_j = t$$

$$x_i = 0, i \notin B \cup \{j\}$$

is $\bar{x} + td$, which has objective value $\bar{v} + t\bar{y}$ (in (P)).

1.7 Unboundedness

If $\bar{d} \geq 0$, (P) is unbounded. $\{\bar{x} + t\bar{d} : t \geq 0\}$ is a feasible halfline and $\bar{c}^t \bar{d} = \bar{c}_j \geq 0$.

Update: Suppose that $\bar{d}$ has a negative entry. Choose $t = \max(\lambda \in \mathbb{R} : \bar{x} + \lambda \bar{d} \geq 0)$ and replace $\bar{x}$ with $\bar{x} + t\bar{d}$. By our choice of t, there exists $i \in B$ such that $\bar{x}_i = 0$ and $\bar{d}_i < 0$. $\bar{x}_i$ is the leading variable.

Now $\bar{d}_i = \bar{a}_{ij} \neq 0$ so $B - \{i\} + \{j\}$ is a basis. Replace B with $B - \{i\} + \{j\}$. Note that $\bar{x}$ is the basic solution for B. Now we repeat.

Since the basis has changed in only two elements, it is easy to update the problem (P').

Termination:

- There are $\leq \binom{n}{m}$ bases.
- At each iteration, the objective value does not decrease.
- There are examples where the simplex method cycles (that is, it revisits a basis).
- If the objective value does not increase in an iteration, then the solution $\bar{x}$ is basic for two distinct bases B_1 and B_2. So $|\text{support}(\bar{x})| < m$.

A basic solution, $\tilde{x}$ is nondegenerate if $|\text{support}(\tilde{x})| = m$.

(P) is nondegenerate if each of its basic solutions are nondegenerate.

Note: The simplex method always terminates for nondegenerate linear programs with at most $\binom{n}{m}$ iterations.

Conjecture Hirsch Conjecture (1957) :

The distance between any bertices in the 1-skeleton of (P) (in standard equality form) is $\leq m$. (Proven false in 2010)

Problems:

1. Is there a polynomial bound on the diameter of the 1-skeleton?
2. Is there a "pivoting rule" for the simplex method that gives a polynomial-time algorithm?

1.8 Perturbation Method

Idea: We carefully select the leaving variable in order to avoid cycling, this is achieved by perturbing b. Given

$$(P) \begin{cases} \text{max} & c^t x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{cases}$$

with $rank(A) = m$, consider

$$(P') \begin{cases} \text{max} & c^t x \\ \text{subject to} & Ax = b' \\ & x \geq 0 \end{cases}$$

where

$$b' = \begin{bmatrix} b_1 + \epsilon \\ b_2 + \epsilon^2 \\ \vdots \\ b_n + \epsilon^n \end{bmatrix}$$

where ϵ is a variavle that we think of as a small positive real number.

For polynomials $p(\epsilon)$ and $q(\epsilon)$ we write $p(\epsilon) < q(\epsilon)$ if the coefficient of the smallest degree term of $q(\epsilon) - p(\epsilon)$ is positive.

Claim: (P') is nondegenerate.

Proof. For a basis B, consider the basic solution $\bar{x}$. We have

$$\bar{x}_B = (A_B)^{-1} b'$$

Since each row of $(A_B)^{-1}$ is a nonzero real vector and the entries of b' is are plynomials with distinct degrees, each term of $\bar{x}_B$ is nonzero. □

Note that we can solve (P) using the simplex method since it is nondegenerate.
Another way to avoid cycling: Smallest subscript rule
Break ties when choosing entering and leaving variables by taking the one of minimum subscript.

Theorem (Bland) :
The smallest subscript rule avoids cycling.

Feasibility: Consider

$$(P) \begin{cases} \text{max} & c^t x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{cases}$$

We have algorithms for:

1. Given a feasible solution, find a basic feasible solution.
2. Given a basic feasible solution, solve (P).

How do you find a feasible solution?
We can scale so that $b \geq 0$. Consider the following "auxiliary problem":

$$(P') \begin{cases} \text{max} & -s_1 - s_2 - ... - s_m \\ \text{subject to} & Ax + s = b \\ & x \geq 0, s \geq 0 \end{cases}$$

Note that:

1. $x = 0$, $s = b$ is a basic feasible solution to (P'), so we can solve this using the simplex method.
2. (P') is also bounded by 0. So the simplex method will terminate with an optimal solution.
3. The objective value of $(\bar{x}, \bar{s})$ is 0 if and only if $\bar{x}$ is a feasible solution for (P).

Remark: If $(\bar{x}, 0)$ is a basic feasible solution for (P'), then $\bar{x}$ is a basic feasible solution for (P).
Farkas' Lemma: Exactly one of the followint has a solution

1. $(Ax = b, x \geq 0)$
2. $(A^t y \geq 0, b^t y < 0)$

The dual of (P') is

$$(D') \begin{cases} \text{min} & b^t y \\ \text{subject to} & A' y \geq 0, y \geq -1 \end{cases}$$

If (P) is infeasible and $\bar{y}$ is an optimal solution to (D'), then $b^t \bar{y} < 0$.
So $\bar{y}$ satisfies $(A^t \geq 0, b^t < 0)$
Note that this gives a constructive proof of the Farkas Lemma.

2 Midterm Review

For $z^1, ..., z^n \in \mathbb{R}^m$, define $\text{conv}(z^1, ..., z^n) = \{\lambda_1 z^1 + ... + \lambda_n z^n : \lambda \in \mathbb{R}^n, \lambda \geq 0, ||\lambda||_1 = 1\}$ and $\text{cone}(z^1, ..., z^n) = \{\lambda_1 z^1 + ... + \lambda_n z^n : \lambda \in \mathbb{R}^n, \lambda \geq 0\}$.

2.1 Seperating Hyperplane Theorems

1. If $b \notin \text{conv}(z^1, ..., z^n)$, then there is a hyperplane seperating b from $\text{conv}(z^1, ..., z^n)$.
2. If $b \notin \text{cone}(z^1, ..., z^n)$, then there is a hyperplane seperating b from $\text{cone}(z^1, ..., z^n)$.

2.2 Polyhedral Theory

Polyhedron: $\{x \in \mathbb{R}^n : Ax \geq b\}$
Polytop: bounded polyhedron
Polyhedral cone: $\{x \in \mathbb{R}^n : Ax \geq 0\}$

Lemma 1 For a polyhedron $P = \{x \in \mathbb{R}^n : Ax \geq b\}$, the following are equivalent:

1. P has no extreme point
2. P contains a line
3. $\text{rank}(A) < n$

Lemma 2 Charecterization of extreme points ... $\Longrightarrow$ there are only finitely many extreme points.

Theorem A :
$S \subseteq \mathbb{R}^n$ is a polytope if and only if it is the convex hull of a finite set of points in $\mathbb{R}^n$.

Theorem B :
If S is a polyhedral cone, then there is a finite subset $Z \subseteq \mathbb{R}$ such that $S = \text{cone}(Z)$. (The converse is also true, we just haven't proved it.)

For $S_1, S_2 \subseteq \mathbb{R}^n$, define

$$S_1 + S_2 = \{a + b : a \in S_1, b \in S_2\}.$$

Theorem C :
Let Z be the set of exreme points of $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. If P does not contain a line, then

$$P = \text{conv}(Z) + \{x \in \mathbb{R}^n : Ax \leq 0\}$$

B and C $\Longrightarrow$ There exist finite sets $Z, D \in \mathbb{R}^n$ such that

1. $P = \text{conv}(Z) + \text{cone}(D)$
2. $||d||_2 = 1$ for all $d \in D$.

If P does not contain a line, then there are unique minimal Z, D of this form.
Z is the set of extreme points. D is the set of extreme rays.
Which implies that every polyhedron that does not contain a line is generated by its extreme points and extreme rays.

2.3 Applications

Caratheodory's Theorem, Helly's Theorem

2.4 Linear Programming

Let

$$(P) \quad \begin{cases} \text{max} & c^t x \\ \text{subject to} & Ax \leq b \end{cases}$$

Fundamental Theorem :
(P) is either infeasible, unbounded, or has an optimal solution.

Infeasibility Theorem (Farkas' Lemma) :
(P) is infeasible if and only if there exists $y \in \mathbb{R}^m$ satisfying

$$A^t y = 0, \; b^t y < 0, \; y \geq 0.$$

Unboundedness Theorem :
(P) is unbounded if and only if

- (P) is feasible, and
- there exists $d \in \mathbb{R}^n$ satisfying $(Ad \leq 0, \; c^t d > 0)$.

2.5 Duality

The dual of (P) is

$$(D)\quad \begin{cases} \min & b^t y \\ \text{subject to} & A^t y = c \\ & y \geq 0 \end{cases}$$

Weak Dulity Theorem :
Uf $\bar{x}$ is feasible for (P) and $\bar{y}$ is feasible for (D) then $c^t\bar{x} \leq b^t\bar{y}$.

Ideally we would like $\bar{x}, \bar{y}$ with $c^t\bar{x} = b^t\bar{y}$.
That is, we want $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ satisfying:

$$(1)\quad \begin{cases} -c^t x & + & b^t y & \geq & 0 \\ Ax & & & \leq & b \\ & & -A^t y & = & -c \\ & & y & \geq & 0 \end{cases}$$

Suppose no such x, y exists. By the Farkas Lemma (assignment question), there exist $z \in \mathbb{R}$, $x \in \mathbb{R}$, and $y \in \mathbb{R}^m$ satisfying:

$$(2)\quad \begin{cases} -c^t x & + & b^t y & < & 0 \\ Ax & & & \leq & bz \\ & & A^t y & = & cz \\ & & y & \geq & 0 \\ & & z & \geq & 0 \end{cases}$$

Note that if $z \neq 0$, then by scaling the solution so that $z = 1$ gives us a solution to (1), a contradiction.
So we have that $z = 0$, so either

1. x satisfies $(c^t x > 0, Ax \leq 0)$, or
2. y satisfies $(b^t y < 0, A^t y = 0, y \geq 0)$

In case 1: (P) is infeasible or unbounded and (D) is infeasible.
In case 2: (P) is infeasible and (D) is infeasible or unbounded.
In either case, neither (P) nor (D) has an optimal solution.

Strong Duality Theorem :
(P) has an optimal solution if and only if (D) has an optimal solution. Moreover, if $\bar{x}$ is optimal for (P) and $\bar{y}$ is optimal for (D), then $c^t x = b^t y$.

2.6 Application of Duality

Theorem :
If $\bar{x}$ is an extreme point of the polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, then there is a half space H such that $P \cap H = \{\bar{x}\}$.

Proof. Since $\bar{x}$ is an extreme point, there exists a partition $(A'x \leq b', A''x \leq b'')$ of the inequalities $Ax \leq b$ such that : $A'\bar{x} = b'$, $\text{rank}(A') = n$, and A' is $n \times n$.
Let

$$c = (A')^t \mathbb{1}$$

$$\alpha = c^t\bar{x} = \mathbb{1}^t A'\bar{x} = \mathbb{1}^t b'$$

and

$$H = \{x \in \mathbb{R}^n : c^t x \geq \alpha\}$$

Now consider the LP:

$$(P)\quad \begin{cases} \max & c^t x \\ \text{subject to} & A'x \leq b' \\ & A''x \leq b'' \end{cases}$$

and its dual

$$(D)\quad \begin{cases} \min & (b')^t y + (b'')^t z \\ \text{subject to} & (A')^t y + (A'')^t z = c \\ & y, z \geq 0 \end{cases} .$$

Let $\bar{y} = \mathbb{1}$ and $\bar{z} = 0$. Now $\bar{x}$ is feasible for (P), $(\bar{y}, \bar{z})$ is feasible for (D), and $c^t\bar{x} = (b')^t y + (b'')^t z = \alpha$. So $\bar{x}$ is optimal for (P) and $(\bar{y}, \bar{z})$ is optimal for (D). Consider another optimal solution $\tilde{x}$ for (P). Nothe that $\bar{y} > 0$, so by the complementary slackness conditions, $A'\tilde{x} = b'$. However A' is invertible, so $\tilde{x} = \bar{x}$. Hence $\bar{x}$ is the unique optimal solution and $H \cap P = \{\bar{x}\}$. □

Exercise: Let $\bar{x}$ be an extreme point of $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, where $A \in \mathbb{Z}^{m\times n}$ and $b \in \mathbb{Z}^m$. Show that if $\bar{x} \notin \mathbb{Z}^n$, then there exists $c \in \mathbb{Z}^n$ such that $\bar{x}$ is an optimal solution to $\max(c^t x : x \in P)$ and $c^t x \notin \mathbb{Z}$.

Do you need A, b to be interger valued?

3 Integer Programming

An interger program is a problem of the form:

$$(IP)\quad \begin{cases} \max & c^t x \\ \text{subject to} & Ax \leq b \\ & x \in \mathbb{Z}^n \end{cases} .$$

The linear programming relaxation is:

$$(LP)\quad \begin{cases} \max & c^t x \\ \text{subject to} & Ax \leq b \end{cases} .$$

We denote by $\text{OPT}(IP)$ and $\text{OPT}(LP)$ to be the oprimal values of IP and LP respectively. If infeasible, we say the optimal value is $-\infty$ and if unbounded we say the optimal value is ∞.
Note that $\text{OPT}(IP) \leq \text{OPT}(LP)$.
If Z is the set of feasible solutions to (IP), then

$$\text{conv}(Z) \subseteq \{x \in \mathbb{R}^n : Ax \leq b\},$$

equality is "rare".

A polyhedron is integral if its extreme points are integral.

Lemma If $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ is integral, $\text{rank}(A) = n$, and (LP) has an optimal solution, then

$$\text{OPT}(IP) = \text{OPT}(LP).$$

Proof. $\text{OPT}(LP)$ is attained on at extreme point. □

3.1 Totally Unimodular Matrices

A matrix is totally unimodular (TU) if each of its square submatrices have deteminant 0, 1, or -1. Note that TU matrices can only have 0, 1, or -1 as entries.

Theorem :

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

Proof. Let $\bar{x} \in \mathbb{R}^n$ be an extreme point of P. Then, by assignment 1, there is a subsystem $A'x \leq b'$ of $Ax \leq b$ that $\bar{x}$ satisfies with equality and $\text{rank}(A') = n$. Now we have

$$\bar{x} = (A')^{-1} b'.$$

So by Crammer's rule, $(A')^{-1}$ is integral and hence so is $\bar{x}$. □

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

1. A^t is TU

2. $[I, A]$ is TU
3. If A' is obtained from A by scaling a row or column by -1, then A' is TU.
4. $[A, -A]$ is TU.

These imply that for $b \in \mathbb{Z}^m$, the following polyhedra are integral:

- $P_1 = \{x \in \mathbb{R} : Ax \leq b, x \geq 0\}$
- $P_2 = \{x \in \mathbb{R} : Ax = b, x \geq 0\}$
- $P_3 = \{x \in \mathbb{R} : Ax \leq b, l \leq x \leq u, \text{ where } l, u \in \mathbb{Z}^n\}$

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

Proof. Suppose otherwise and consider counterexample $A \in \{0, \pm 1\}^{m \times n}$ with $m + n$ minimum. Thus $m = n$ and $\det(A) \notin \{0, 1, -1\}$. By minimality, each column has two non-zero entries, a 1 and a -1. So the rows sum to zero and hence, $\det(A) = 0$, a contradiciton. □

——*Mising a lecture* ——
Recall that for any matching M and cover C, then $|M| \leq |C|$.

Theorem Konig'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 $G = (V, E)$. Consider

$$(P) \quad \begin{cases} \text{max} & \mathbb{1}^t x \\ \text{subject to} & Ax \leq \mathbb{1} \\ & x \geq 0 \end{cases}.$$

The dual is

$$(D) \quad \begin{cases} \text{min} & \mathbb{1}^t y \\ \text{subject to} & A^t y \geq \mathbb{1} \\ & y \geq 0 \end{cases}.$$

Both programs are feasible so they have an optimal solution. Let $\bar{x}$ and $\bar{y}$ be optimal extreme points for (P) and (D) respectively. Since A is TU, $\bar{x}$ and $\bar{y}$ are integral. Note that $\bar{x}$ and $\bar{y}$ must be $\{0, 1\}$-valued also (ask youself why). Let $\bar{M} = \text{support}(\bar{x})$ and $\bar{C} = \text{support}(\bar{y})$. Note that $\bar{C}$ is a cover and $\bar{M}$ is a matching. Moreover

$$\left|\bar{C}\right| = \mathbb{1}^t y = \mathbb{1}^t x = \left|\bar{M}\right|,$$

so we found a matching and a cover of the same size. □

3.2 Min-Cost Perfect Matching

Problem: Given a bipartite graph $G = (V, E)$ and $c \in \mathbb{R}^E$, find a perfect matching minimizing $\sum_{e \in M} c_e$.
We denote $\sum_{e \in M} c_e = c(M)$.
We will assume that G has a perfect matching.
Example: PICTURE!!!!!
Claim: $\tilde{M}$ is a mincost perfect matching.
Suppose

$$c'(e) = \{c(e) + 1 : e \text{ incident with } a, c(e) \text{ otherwise}\}$$

Then, for any perfect matching M,

$$c'(M) = c(M) + 1.$$

So finding a mincost perfect matching with respect to c' is the same as finding a mincost perfect matching with respect to c Let A be the incidence matrix of G and consider

$$(P) \quad \begin{cases} \text{min} & c^t x \\ \text{subject to} & Ax = \mathbb{1} \\ & x \geq 0 \end{cases}$$

and its dual

$$(D) \quad \begin{cases} \max & y(V) \\ \text{subject to} & A^t y \leq c \end{cases}.$$

Since A is TU, there is an optimal perfect matching with $c(\tilde{M}) = \text{OPT}(P)$ (it's feasible because we assumed G to have a perfect matching). Let $y \in \mathbb{R}^V$ and let $\bar{c} = c - A^t y$. We call these $\bar{c}$ reduced costs. Note that $\bar{c} \geq 0$ if and only if y is feasible for (D). Define $G^=(y)$ to be the subgraph of G with vertex set $V(G)$ and edge set $\{e \in E : \bar{c}_e = 0\}$.

Complementary slackness: if M is a perfect matching and y' is a feasible solution for (D) then $c(M) = y'(V)$ if and only if $M \subseteq E(G^=(y'))$.

Claim: If $\bar{y}$ is a feasible solution for (D) and M is a perfect matching of $G^=(\bar{y})$, then M is a mincost perfect matching.

Algorithm: Let (X, Y) be the bipartition of G and assume that $|X| = |Y|$ since otherwise G has no perfect matching.

Overview: Find a feasible $\bar{y}$ for (D). Let $\bar{y}_0 = 0$, for each $v \in Y$. Let $\bar{y}_v = \min(c_e : e = vw, w \in Y)$.

Step 1: If $G^=(\bar{y})$ has a perfect matching M, stop: output M.

Step 2: Find a feasible solution y' for (D) with $y'(V) > \bar{y}(V)$. Replace $\bar{y}$ with y' and repeat step 1.

Example: See picture Note that $G^=(\bar{Y})$ has no perfect matching since $N_{G^=(\bar{y})}(\{a, b\}) = \{4\}$.

Hall's theorem

Lemma If $\bar{y} \in \mathbb{R}^V$ is a feasible solution for (D), and $G^=(\bar{y})$ has a perfect matching M, then M is a min cost perfect matching.

Theorem Hall's Theorem :

Let G be a bipartite graph with bipartition (X, Y) where $|X| = |Y|$. Then G has a perfect matching if and only if $|N(Z)| \geq |Z|$ for each $z \in X$.

Assumption: We have an efficient algoithm for Hall's Theorem (That is, we can find either a perfect mathcing or a set $z \subseteq X$ with $|N(Z)| < |Z|$).

Let $\bar{y} \in \mathbb{R}^V$ be a feasible solution for (D) and suppose that $G^=(\bar{y})$ has no perfect matching. Then there exists $Z \subseteq X$ such that

$$\left|N_{G^=(\bar{y})}(Z)\right| < |Z|.$$

Let

$$y'_v = \begin{cases} \bar{y}_v + \epsilon & v \in Z \\ \bar{y}_v - \epsilon & v \in N_{G^=(\bar{y})}(Z) \\ \bar{y}_v & \text{otherwise} \end{cases}$$

Note that, for small $\epsilon > 0$, $\bar{y}$ is feasible and has objective value

$$y'(V) = \bar{y}(V) + \epsilon(|Z| - |N(Z)|) > \bar{y}(V).$$

How large can we make ϵ?

$$\epsilon = \min\left(c_e : e = uv \in E(G), u \in Z, v \in Y - N_{G^=(\bar{y})}(Z)\right).$$

This is well defined unless G has no perfect matching.

Remarks:

- if $c \in \mathbb{Z}^E$ and $\bar{y} \in \mathbb{Z}^V$, then we get $y' \in \mathbb{Z}^V$. (So we keep an integral dual solution.)
- There is a way to choose Z so that the number of iterations is at most $|V(G)|^2$ (independent of c). (seen in CO 450)
- The mins cost perfect matching problem can be solved in polynomial time, even for nonbipartite graphs. (CO 450) We need additional constraints for $X \subseteq V(G)$ odd.

3.3 Directed Graphs

A directed graph is a pair (V, E) where V is a finite set and E is a set of ordered pairs of distinc vertices. V is the vertex set and E is the edge set. For $e = uv \in E$, u is the tail and v is the head of e.

Incidence matrix for $D = \big(\{1,2,3\}, \{12, 23, 31, 13\}\big)$:

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

Since A has one 1 and one -1 in each column, A is TU (by previous lemma).

For $X \subseteq V(G)$, we define

$$\mathrm{IN}(X) = \{uv \in E(G) : u \notin X, v \in X\}$$

and

$$\mathrm{OUT}(X) = \{uv \in E(G) : u \in X, v \notin X\}.$$

Suppose that

$$Ax = b,$$

Then

$$x(\mathrm{IN}(\{v\})) - x(\mathrm{OUT}(\{v\})) = b_v$$

for each $v \in V$.

!!!!!!!!!!!!!! — missed a lecture — !!!!!!!!!!

Claim: If x is a feasible (s,t)-flow and (S,T) is an (s,t)-cut, then inflow$(t) \leq u(S,T)$

Theorem Max-Flow Min-Cut Theorem :

The maximum value of a feasible (s,t)-flow is equal to the minimum cpacity of an (S,T)-cut.

Proof. See other peoples notes. □

4 Complexity Theory

Decision problems

LP Feasibility Problem:
Instance: $A \in \mathbb{Q}^{m\times n}, b \in \mathbb{Q}^m$
Question: Does there exists $x \in \mathbb{Q}^n$ such that $Ax \leq b$.

IP Feasibility Problem:
Instance: $A \in \mathbb{Q}^{m\times n}, b \in \mathbb{Q}^m$
Question: Does there exists $x \in \mathbb{Z}^n$ such that $Ax \leq b$.

Bipartite Matching Problem:
Instance: G bipartite, $k \in \mathbb{Z}_+$
Question: Does G have a matching of size $\geq k$?

Clique Problem:
Instance: a graph G, $k \in \mathbb{Z}_+$
Question: Does G contain a set of k pairwise adjacent vertices?

A "decision problem" is a yes/no question on a countable set of instances. The "size" of and instence is the length of a binary encoding.

An algorithm is polynomial-time if its running times is bounded by a polynomial in the size of the input. $\mathcal{P}$ is the set of all decision problems that can be solved in polynomial time.

Nondeterministic Polynomial Time
Problems with "easy to certify" yes-instances.

Claim: LP feasibility is in NP.

Proof. Consider $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Suppose that $\bar{x} \in P$. Without loss of generality, we can assume that $\bar{x} \geq 0$. By adding some inequalities, we may assume that $P \subseteq \mathbb{R}^n_+$. Now we have extreme points. So there is a subsystem $A'x \leq b'$ such that A' is $n \times n$ and $\text{rank}(A') = n$, and $A'\tilde{x} = b'$. Now $\tilde{x}$ is the unique solution to $A'x = b'$ and we have that

$$size(\tilde{x}) \leq polynomial(size(A', b')).$$

□

Exercise: Show that IP is in NP.

Definition A decision problem P is in NP if there is a polynomial time algorithm A and a polynomial p such that

1. for each yes-instance I of P there is a certificate c such that $|c| \leq p(|I|)$ and A accepts (I, C).

2. For each no-instance I of P and any c with $|c| \leq p(|I|)$, A rejects (I, C).

Definition We say that P_1 reduces to P_2 if there is a plynomial time algorithm A such that for each instance I of P, A generates an instance I_2 of P_2 such that I_1 is a yes instance of P_1 if and only if I_2 is a yes instance of P_2.

Example Consider the clique problem on an instance $G = (V, E), k$. Construct an instance of IP feasibility,

$$P \begin{cases} \sum_{v \in V} x_v = k \\ x_u + x_v \leq 1 (u \sim v) \\ 0 \leq x \leq 1 \\ x \text{ integer} \end{cases} .$$

Definition A problem $P \in$NP is NP-complete if every problem in NP reduces to it.

Theorem Cook :

IP feasibility is NP-complete. (Cook used "3-SAT").

4.1 Formalism

Let $\mathcal{A} = \{1, 0, -\}$. Let $\mathcal{A}^*$ denote the set of finite words in $\mathcal{A}$. A problem is any subset of $\mathcal{A}^*$. Given $w, w_1, w_2 \in \mathcal{A}^*$, we say that w conatins w_1 if

$$w = \alpha w_1 \beta$$

for some $\alpha, \beta, \in \mathcal{A}^*$. We say that w' is obtained from w by replacing w_1 by w_2 if

$$w = \alpha w_1 \beta \text{ and } w' = \alpha w_2 \beta.$$

An algorithm is a sequence

$$(w_1, w_1'), \cdots, (w_k, w_k').$$

To run the algorithm on a word w, for each $i \in \{1, ..k\}$ we replace the first instance of w_i with w_i' and start over with $i = 1$. Otherwise, $i++$.

An algorithm solves a problem Π if Π is the set of instances on which the algorithm terminates. An algorithm is polynomial-time if there is a polynomial p such that for each instance I on which the algorithm terminates, the algorithm terminates in $p(size(I))$ steps.

Exercise: Write an algorithm for checking $a + b = c$ on given intergers a, b, and c.

5 Nonlinear Optimization

minimize $(f(x) : x \in S)$, $S \subseteq \mathbb{R}^n$, $f : S \to \mathbb{R}$. Recall that

- "in theory" we canreduce to the case that $f(x)$ is linear and S is convex.
- Small problems are nontrivial.

Note that $\inf(f(x) : x \in S)$ may not be attained ever if $S \neq \emptyset$ and $f(x)$ is bounded below.

5.1 Compactness

A set $S \subseteq \mathbb{R}^n$ is closed if for each convergent sequence in S converge in S.
stuff, see other notes

Theorem :
Let $S \subseteq \mathbb{R}^n$ be a closed set and let $\bar{x} \in S$ be on the boundery. Then there exists a nonzero $c \in \mathbb{R}^n$ such that $\bar{x}$ minimizes $(c^t : x \in S)$.

Proof Sketch. Take $z \notin S$ such that $\bar{x}$ is the nearest point in S to z. Let $c = \bar{x} - z$. Now continue as in the proof of the seperating Hyperplane Theorem. □

5.2 Certifying Optimality

How can we prove that $\bar{x}$ minimizes $(c^t x : x \in S)$? In fact, we can't prove it in general. Nonlinear programming is undecideble.

Theorem Cost Splitting Theorem (Sufficient condition for Optimality) :
Let $S_1, \dots, S_m \subseteq \mathbb{R}^n$, let $S = S_1 \cap S_2 \cap \dots \cap S_m$. Let $c \in \mathbb{R}^n$ and let $\bar{x} \in S$. If there exist $c_1, .., c_m \in \mathbb{R}^n$ such that $c_1 + \dots + c_m = c$ and such that $\bar{x}$ minimizes $(c_i^t x : x \in S_i)$ for all i, then $\bar{x}$ minimizes $(c^t x : x \in S)$.

Proof.

$$\begin{aligned}
c^t\bar{x} &\geq \min(c^t x : x \in S) \\
&= \min(c_1^t x + \dots + c_m^t x : x \in S) \\
&\geq \min(c_1^t x_1 + \dots + c_m^t x_m : x_i \in S) \\
&= \min(c_1^t x_1 : x_1 \in S) + \dots + \min(c_m^t x_m : x_m \in S) \\
&\geq \min(c_1^t x_1 : x_1 \in S_1) + \dots + \min(c_m^t x_m : x_m \in S_m) \\
&= c_1^t \bar{x} + \dots + c_m^t \bar{x} \\
&= c^t \bar{x}
\end{aligned}$$

□

Note that cost splitting is not always possible. Here is an example: Let

$$S_1 = \text{Ball}\left(\begin{bmatrix}0\\0\end{bmatrix}, 1\right), \; S_1 = \text{Ball}\left(\begin{bmatrix}2\\0\end{bmatrix}, 1\right), \text{ and } S = S_1 \cap S_2.$$

Cost splitting for linear Programming: it always works.

Lemma If $S_1, S_2 \subseteq \mathbb{R}^n$ are convex and $\text{int}(S_1 \cap S_2) \neq \emptyset$, then

$$\text{closure}(S_1 \cap S_2) = \text{closure}(S_1) \cap \text{closure}(S_2).$$

Tangent cone: For $S \subseteq \mathbb{R}^n$ and $\bar{x} \in S$,

$$T(\bar{x}, S) := \text{closure}(\text{cone}(\{x - \bar{x} : x \in S\})).$$

Remarks:

- If S is convex, then $T(x, S)$ is a closed convex cone.
- This definition is nonstandard, but agrees with the usual definition on convex sets.

Theorem :
Let $S_1, S_2 \subseteq \mathbb{R}^n$ be closed convex sets with their intersection with non-empty interior. Let $\bar{x} \in S_1 \cap S_2$. Then

$$T(\bar{x}, S_1 \cap S_2) = T(\bar{x}, S_1) \cap T(\bar{x}, S_2).$$

Proof. We can translate S_1 and S_2 so that $\bar{x} = 0$. Now since S_1 and S_2 are convex sets and $0 \in S_1 \cap S_2$, we have

$$\text{cone}(S_1 \cap S_2) = \text{cone}(S_1) \cap \text{cone}(S_2).$$

Since the interior of $S_1 \cap S_2$ is non-empty, the interior of their cones is non-empty as well and we can therefore apply lemma 1 to get

$$\text{closure}(\text{cone}(S_1 \cap S_2)) = \text{closure}(\text{cone}(S_1)) \cap \text{closure}(\text{cone}(S_2))$$

which is the same as

$$T(\bar{x}, S_1 \cap S_2) = T(\bar{x}, S_1) \cap T(\bar{x}, S_2).$$

□

Convex cones:

Theorem Separating Hyperplane Theorem for Cones :
Let $\emptyset \neq K \subseteq \mathbb{R}^n$ be a closed convex cone and $z \in \mathbb{R}^n$. If $z \notin K$, then there exists $c \in \mathbb{R}^n$ such that $c^t x \geq 0$ for all $x \in K$ and $c^t z < 0$.

Proof. By the seperating hyperplane theorem, there exists $c \in \mathbb{R}^n$ and $b \in \mathbb{R}$ such that $c^t x \geq b$ for all $c \in K$ and $c^t z < b$. We may assume that $b = \inf(c^t x : x \in K)$. Note that $0 \in K$ so $b \leq 0$. We may assume that $b < 0$ since otherwise, we would be done. There exists $\bar{x} \in K$ with $b \leq c^t \bar{x} < \frac{b}{2}$. However, K is a cone so $2\bar{x} \in K$ but $c^t 2\bar{x} < \frac{2b}{2} = b$. □

Duality for cones: For $S \subseteq \mathbb{R}^n$, define

$$S^* = \{c \in \mathbb{R}^n : c^t x \geq 0, x \in S\}$$

Remarks:

- If $0 \in S$, then S^* is the set of all $c \in \mathbb{R}^n$ such that 0 minimizes $(c^t : x \in S)$.
- If K is a cone, then K^* is called the dual of K.

Lemma For any $S \subseteq \mathbb{R}^n$, S^* is a closed convex cone.