CO 255 Notes: Introduction to Optimization (Advanced Level)

Johnew Zhang

October 24, 2012

Contents

1 Optimization

Given a set S (the feasible region) and a function $f : S \to \mathbb{R}$ (the objective function). Solve $max(f(x) : x \in S)$ or $min(f(x) : x \in S)$

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

1.1 Linear Programming

$$f(x) = c^T x$$

$$S(x) = \{x \in \mathbb{R}^n : Ax \leq b\}$$

$(C \in \mathbb{R}^n, b \in \mathbb{R}^m, A \in \mathbb{R}^{m \times n})$.

1.2 Integer Linear Programming Problems

$$f(x) = c^T x$$
$$S(x) = \{x \in \mathbb{Z}^n : Ax \leq b\}$$

1.3 Complex Optimization

$$min(f(x) : x \in S)$$

such that $S \subseteq \mathbb{R}^n$ and convex; also f is convex

Remark Consider an optimization problem $min(f(x) : x \in S)$. We can assume without much loss of generality that

1. $S \subseteq \mathbb{R}^n$
2. f is linear.
$$min(f(x) : x \in S) = min(z : z = f(x), x \in S)$$
3. S is convex (since for linear function, $min(f(x) : x \in S) = min(f(x) : x \in conv(S))$

1.4 Example

A two player game Given $A \in \mathbb{R}^{m \times n}$. Rose chooses a row i and Colin chooses a column j independently then Colin pays Rose $\$a_{ij}$

e.g.

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

If Rose chooses 2 she gets ≥ 1. If she chooses 1, she gets $\geq (-2)$. If she chooses the two rows with equal probabilities. She expects $\geq min(\frac{2+1}{2}, \frac{-2+5}{2}) = \frac{3}{2}$

Rose wants to solve

$$max_{p\in\mathbb{R}^n} min_{j\in\{1,\cdots n\}}(\sum_{i=1}^{m} p_i \cdot a_{ij})$$

subject to $p_1 + \cdots + p_n = 1$ such that $p_1, \cdots, p_n \geq 0$.

Equivalently, to maximize z such that $z \leq \sum_{i=1}^{m} p_i a_{ij}$, for $j \in \{1, \cdots, n\}$, $\sum_{i=1}^{m} p_i = 1$, $p_i \geq 0$.

Weighted bipartite matching Problem: Given n jobs, n workers and a "utility" a_{ij} for workers to complete job i. Find an assignment maximizing the total utility (i.e. the sum of the utilities).

Formulation:

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

subject to

$$\sum_{j=1}^{n} x_{ij} = n, \text{ for } i \in \{1, \cdots, n\}$$

$$\sum_{i=1}^{n} x_{ij} = n, \text{ for } j \in \{1, \cdots, n\}$$

$$x_{ij} \in \{0, 1\}, \text{ for } i, j \in \{1, \cdots, n\}$$

This is an integer linear programming formulation.

3D Matching Problem: given $a \in \mathbb{R}^{n\times n\times n}$, a_{ijk} is the utility of job i completed by worker j on machine k; find an "assignment" of maximum total utility.

Formulation

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

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

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

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

$$0 \leq x_{ijk} \leq 1 \text{ integer } , i, j, k \in \{1, \cdots, n\}$$

Remark: The 3D matching problem is NP-hard and, hence integer linear programming is NP-hard

Diophantine Equation Example

$$max \sin(\pi x)^2 + \sin(\pi y)^2 + \sin(\pi z)^2$$

$$x^3 + y^3 - z^3 = 0$$

$$x, y, z \geq 1$$

Note this condition has optimal value is equal to 0 if and only if there are non-negative integers x, y, z such that

$$x^3 + y^3 = z^3$$

Here we have a side notes for **diophantine equation**: it is an equation $p(x_1, \cdots, x_n) = 0$ where p is a polynomial with integer coefficients. Can we decode whether or not there exist $x_1, \cdots, x_n \in \mathbb{Z}$ such that $p(x_1, \cdots, x_n) = 0$? No, not even if we fix $n = 9$

Formulation

$$min \sum_{i=1}^{n} \sin(\pi x_i)^2$$

$$\text{subject to } P(x_1, \cdots, x_n) = 0$$

This has optimal value 0 off p has an integer root.

Distance feasibility Problem: Given $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$, how far is z from the feasible region $\{x \in \mathbb{R}^n : Ax \leq b\}$?

Formulation

$$min \sum_{i=1}^{n} (x_i - z_i)^2$$

$$Ax \leq b$$

This is a convex optimization problem

2 Feasibility Problem

2.1 Linear Algebra Review

Remark: Matrices do not have ordered rows and columns $A \in \mathbb{F}^{X \times Y}$, $\mathbb{F}$ is a field, X, Y are finite sets.

2.1.1 Fundamental Theorem of Linear Algebra

For $A \in \mathbb{F}^{m \times n}, b \in \mathbb{F}^m$, exactly one of the following holds:

1. There exists $x \in \mathbb{F}^n$ such that $Ax = b$
2. There exists $y \in \mathbb{F}^m$ such that $y^T A = 0$ and $y^T b = 1$ (that is we can take a linear combination of the equation to get $0 = 1$)

2.1.2 Solutions to linear systems

Let $A \in \mathbb{F}^{m \times n}$ with $rank(A) = m$ and let $b = \mathbb{F}^m$. Let A_j denote the jth column of A, and for $B \in \{1, \cdots, n\}$, let $A_B = [A_j : j \in B]$. We call B a basis if $|B| = m$ and A_B is non-singular.

Note that if B is a basis, then there is a unique solution to

1. $Ax = b$
2. $x_j = 0, j \notin B$.

We call this the **basic solution** for B.The **support** of $x \in \mathbb{F}^n$ is $supp(x) = \{i : x_i \neq 0\}$

Theorem 1. For $A \in \mathbb{F}^{m \times n}, b \in \mathbb{F}^n$, if $Ax = b$ has a solution, then it has a solution whose support has size $\leq rank(A)$.

Note that $Ax = b$ can be solved in $O(mn \operatorname{rank}(A))$ arithmetic operations. Is this efficient? What about the size of the solution? $\cdots$.

2.1.3 The size of a solution

For $\in \mathbb{Z}$, define $size(a) \leq [log(|a|+1)]+1 \leq log_2(|a|)+2, \forall a \geq 1$. $size(\frac{a}{b}) = size(a)+size(b)$. Let $A \in \mathbb{Z}^{n \times n}, b \in \mathbb{Z}^n$ and let L be the size of the largest entry in A or b. Suppose that A is nonsingular,

$$size(det(A)) \leq size(n!(2^L)^n) \leq 2 + log_2(n!(2^L)^n) \leq 2 + n(log_2 n + L)$$

Now consider

$$x = A^{-1}b$$

By Cramer's Rule, each entry of $det(A)A^{-1}$ a determinant of a sub matrix of A and, hence, has size $\leq 2 + n(log_2(n) + L)$ so each entry of x has size $\leq n + (L + 1)(2 + n(log_2 n + L))$ (this is polynomially bounded in the size of A, b)

3 Systems of Linear Inequalities

Theorem 2. (Farkas Lemma; Thm 2.7) Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Exactly one of the following hold

1. There exists $x \in \mathbb{R}^n$ such that $Ax \le b$
2. There exists $y \in \mathbb{R}^m$ such that $y \ge 0, y^tA = 0$, and $y^tb = -1$.

Easy part: (1) and (2) cannot both hold: if $Ax \le b$ and $y \ge 0$, then $y^tAx \le y^tb$. So, if $y^tA = 0$, then $0 \le y^tb$. It remains to prove that: if (1) does not hold then (2) does. Restatement: Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be linear function $i \in \{1, \cdots, m\}$. If $f_i(x) \le 0, i \in \{1, 2, \cdots, m\}$ has no solution, then there exists $\alpha \in \mathbb{R}^n_+$ such that $\sum_{i=1}^m \alpha_i f_i = 1$. (Here $\mathbb{R}_+ = \{z \in \mathbb{R} : z \ge 0\}$)

3.1 Variable Elimination (Fourier-Motzkin)

Rewrite the inequalities as

$$x_n \ge g_i(x_1, \cdots, x_{n-1}), i \in A_1$$

$$x_n \le g_i(x_1, \cdots, x_{n-1}), i \in A_2$$

$$0 \ge g_i(x_1, \cdots, x_{n-1}), i \in A_3$$

(A_1, A_2, A_3) partition $\{1, \cdots, m\}$ and $g_1, \cdots, g_m$ are defined implicitly.

Note that there is a solution if and only if there exist $x_1, \cdots, x_{n-1} \in \mathbb{R}$ satisfying the third condition above such that

$$\max_{i\in A_1}(g_i(x_1, \cdots, x_{n-1})) \le \min_{i\in A_2}(g_1(x_1, \cdots, x_{n-1}))$$

Equivalently,

$$g_i(x_1, \cdots, x_{n-1}) \le g_j(x_1, \cdots, x_{n-1}), i \in A_1, j \in A_2$$

$$0 \ge g_i(x_1, \cdots, x_{n-1}), i \in A_3$$

Note that this is a system and linear inequalities in $n-1$ variables.

Assume that Farkas's Lemma holds for systems with $n-1$ variables. (The result is trivial when $n = 0$)

Suppose that there is no solution in the above two inequalities. Then by the inductive assumption, there exist $\alpha \in \mathbb{R}_+^{A_1\times A_2}$ and $\beta \in \mathbb{R}_+^{A_3}$ such that

$$\sum_{i\in A_1}\sum_{j\in A_2} \alpha_{ij}(g_i - g_j) + \sum_{k\in A_3} \beta_k g_k = 1$$

(Note $A_1 \times A_2 = \{(ij) : i \in A_1, j \in A_2\}$)

For $i \in \{1, \cdots, m\}$ we define $r_i = \begin{cases} \sum_{j\in A_2} \alpha_{ij} & : i \in A_1 \\ \sum_{j\in A_1} \alpha_{ji} & : i \in A_2 \\ \beta_i & : i \in A_3 \end{cases}$

Now $f_i(x_1, \cdots, x_n) = \begin{cases} g_0(x_1, \cdots, x_m) - x_0 & : i \in A_1 \\ -g_1(x_1, \cdots, x_m) + x_n & : i \in A_2 \\ g_i(x_1, \cdots, x_m) & i \in A_3 \end{cases}$

Now $\alpha \geq 0$ and

$$\sum_{i=1}^{m} \alpha_i f_i = \sum_{i\in A_1} (\sum_{j\in A_2} \alpha_{ij})(g_1 - x_m) + \sum_{i\in A_2} (\sum_{j\in A_1} \alpha_{ij}(x_n - g_i)) + \sum_{k\in A_2} \beta_k g_k$$
$$= (\sum_{i\in A_1} \sum_{j\in A_2} \alpha_{ij}(g_i - g_j) + \sum_{k\in A_3} \beta_k g_k) - (\sum_{i\in A_1} \sum_{j\in A_2} \alpha_{ij})x_n + (\sum_{i\in A_1} \sum_{j\in A_2} \alpha_{ij})x_n = 1$$

Hence this proves Farkas' Lemma.

3.2 Other Forms of Farkas' Lemma

Theorem 3. Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Exactly one of the following hold:

1. There exists $x \in \mathbb{R}^n$ such that $Ax = b, x \geq 0$.
2. There exists $y \in \mathbb{R}^n$ such that $y^t A \geq 0, y^t b = -1$

Proof. $(A_x = b, x \geq 0)$ can be rewritten as $A_2 \leq b, -A_x \leq -b, -x \leq 0$.

Let $A' = \begin{pmatrix} A \\ -A \\ -I \end{pmatrix}$ and $b' = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$ so (1) is equivalent to (1') there exists $x \in \mathbb{R}^n$ such that $A'x \leq b'$. By the Farkas Lemma, this is equivalent to (2') there do not exist $y_1, y_2, y_3 \in \mathbb{R}^m$ such that

$$[y_1, y_2, y_3]^T A' = 0, [y_1, y_2, y_3]^T b' = -1, y_1, y_2, y_3 \geq 0$$

That is :

$$y_1^T A - y_2^T A - y_3^T = 0$$
$$y_1^T b - y_2^T b = -1$$
$$y_1, y_2, y_3 \geq 0$$

That is

$$(y_1 - y_2)^T A = y_3$$
$$(y_1 - y_2)^T b = -1$$
$$y_1, y_2, y_3 \geq 0$$

That is equivalent to: there exists $y \in \mathbb{R}^m$ such that

$$y^T A \geq 0$$

$$y^T b = -1$$

$\square$

Theorem 4. Let $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^n$. Exactly one of the following hold:

1. There exists $x \in \mathbb{R}^n$ such that $Ax \leq b, x \geq 0$
2. There exists $y \in \mathbb{R}^m$ such that $y^T A \geq 0, y^T b = -1$ and $y \geq 0$.

From Geometry prospective, suppose $A = [A_1, \cdots, A_n] \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$
Define

$$cone(A_1, \cdots, A_n) = \{A_1 x_1 + \cdots + A_n x_n : x \in \mathbb{R}^n, x \geq 0\}$$

Problem is $b \in cone(A_1, \cdots, A_n)$?

Equivalently: does $(Ax = b, x \geq 0)$ have a solution. by the above theorem, if $b \notin cone(A_1, \cdots, A_n)$, then there exists $\alpha \in \mathbb{R}^m$ such that $\alpha_T A \geq 0$ and $a^T b = -1$.

Equivalently, $\alpha_1^T \geq 0, \alpha^T A_2 \geq 0, \cdots, \alpha_T A_n \geq 0, \alpha^T b = -1$

Equivalently, $A_1, \cdots, A_n$ are contained in the "half-space" $\{x \in \mathbb{R}^m : \alpha^T x \geq 0\}$ but b is not.

Equivalently: $cone(A_1, \cdots, A_n) \subseteq \{x \in \mathbb{R}^m : \alpha^T x \geq 0\}$ but $b \notin \{x \in \mathbb{R}^m : \alpha^T x \geq 0\}$

Theorem 5. $b \notin cone(A_1, \cdots, A_n)$ off there is a hyperplane separating b from $cone(A_1, \cdots, A_n)$

3.3 Separating Hyperplane Theorem

Theorem 6. Let $S \subseteq \mathbb{R}^n$ be a closed convex set and $b \in \mathbb{R}^m$, If $b \notin S$ then there is a hyperplane separating b from S.

4 Linear Programming

A linear program (or LP) is a problem of the form

$$\max(c^T x : Ax \leq b)$$

$$\text{or } \min(c^T x : A_x \geq b)$$

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

Note: $\max(c^T x : Ax \leq b) = -\min(-c^T x : A_x \leq b)$

Example

$$\begin{cases} \text{maximize} & x_2 \\ \text{subject to} & \\ & x_1 + x_2 \leq 3 \\ & 4x_1 + x_2 \geq 4 \\ & x_1 + 2x_2 \leq 4 \\ & x_1, x_2 \geq 0 \end{cases}$$

Now x^* satisfies

$$4x_1 + x_2 = 4$$
$$x_1 + 2x_2 = 4$$

so $x^* = [\frac{4}{7}, \frac{12}{7}]^T$ and the optimal value is $\frac{12}{7}$.

Problem:How in general can we prove that a given solution is optimal? Equivalently, how can we generate upper bound on the optimal value?

Answer: Take linear combination of the constraints.

Example:

$$x_1 + x_2 \leq 3$$
$$4x_1 + x_2 \leq 4$$
$$x_1 + 2x_2 \leq 1$$

so $x_2 \leq 2$. Each feasible solution has objective value ≤ 2.

Note that to prove that x^* is optimal we should only use inequalities that x^* satisfies with equality.

$$4x_1 + x_2 \geq 4$$
$$x_1 + 2x_2 \leq 4$$

$(4\alpha + \beta)x_1 + (\alpha + 2\beta)x_2 \leq 4(\alpha + \beta), (\alpha \leq 0, \beta \geq 0)$

We want $4\alpha + \beta = 0$ and $\alpha + 2\beta = 1$ to get the object function x_2. Thus $\alpha = -\frac{1}{7}, \beta = \frac{4}{7}$, which gives $x_2 \leq \frac{12}{7}$

Hence x^* is optimal.

4.0.1 Duality

Remark: The problem of determining the best bound on the objective function via linear combination of constraints, is an LP.

$$\begin{aligned} x_1 + x_2 &\leq 3 & \qquad y_1 &\geq 0 \\ 4x_1 + x_2 &\geq 4 & y_2 &\leq 0 \\ x_1 + 2x_2 &\leq 4 & y_3 &\geq 0 \\ x_1, x_2 &\geq 0 & & \end{aligned}$$

Take the linear combination:

$$(y_1 + 4y_2 + y_3)x_1 + (y_1 + y_2 + 2y_3)x_2 \leq 3y_1 + 4y_2 + 4y_3$$

$$x_1, x_2 \geq 0$$

We want

$$0x_1 + 1x_2 \leq (y_1 + 4y_2 + y_3)x_1 + (y_1 + y_2 + 2y_3)x_2$$

so we want

$$y_1 + 4y_2 + y_3 \geq 0$$

$$\text{and } y_1 + y_2 + 2y_3 \geq 1$$

The dual of (P) is: $\begin{cases} \text{minimize} & 3y_1 + 4y_2 + 4y_3 \\ \text{subject to} & y_1 + 4y_2 + y_3 \geq 0 \\ & y_1 + y_2 + 2y_3 \geq 1 \\ & y_1 \geq 0, y_2 \leq 0, y_3 \geq 0 \end{cases}$

By construction: If x is feasible for (P) and y is feasible for (D) then

$$x_2 \leq 3y_1 + 4y_2 + 4y_3$$

LHS is the objective function for (P) and the RHS is the objective function for (D)

Note that for $x^* = [\frac{4}{7}, \frac{12}{7}]^T$ and $y^* = [0, \frac{1}{7}, \frac{4}{7}]$. We get equality!

4.0.2 Unboundedness

$$\begin{cases} \text{minimize} & x_1 + x_2 \\ \text{subject to} & -2x_1 + x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{cases}$$

Let $\hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $d = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Then $\hat{x} + \alpha d$ is feasible for all $\alpha \geq 0$ and has objective value $3 + 2\alpha$, so (P) is unbounded. Note that the "half-line" $\{\hat{x} + \alpha d : \alpha \geq 0\}$ is contained in the feasible region and $c^T d > 0$.

Theorem 7. ("Fundamental Theorem of LP") Every linear program either

1. is infeasible
2. is unbounded, or
3. has an optimal solution

Consider the problem: $(NLP) \begin{cases} \text{minimize} & \frac{1}{x} \\ \text{subject to} & x \geq 1 \end{cases}$

Consider an LP: $(P) \max(c^T x : Ax \leq b), A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$. Suppose that $\hat{x}$ is a feasible region with $c^T \hat{x} = r$.

Lemma 1. If the column A is a linear combination of the other columns, then either

1. (P) has a feasible solution $\tilde{x}$ with $c^T \tilde{x} = r$ and $\tilde{x}_1 = 0$, or
2. there exists $d \in \mathbb{R}^n$ such that $Ad = 0$ and $c^T d > 0$.

(Hence (P) is unbounded)

Proof. There exists $z \in \mathbb{R}^n$ such that $Az = 0$ and $z_1 = -1$. We may assume that $c_T z = 0$ since otherwise (2) holds with $d = z$ or $-z$. Let $\tilde{x} = \hat{x} + \hat{x}_1 \cdot z$. Then $\hat{x}_1 = 0, \tilde{x}$ is feasible and $c^T \tilde{x} = c^T \hat{x} = r$. □

Lemma 2. Let $A'x \leq b'$ be the subsystem of $Ax \leq b$ that $\hat{x}$ satisfies with equality. Then $\hat{x}$ is an extreme point of $\{x \in \mathbb{R}^n : Ax \leq b\}$ if and only if $rank(A') = n$.

Proof. Suppose that $rank(A') = n$ and $\hat{x} = \lambda x' + (1 - \lambda) x^2$ where $0 < \lambda < 1$ and x^1 and x^2 are feasible.

Since $A'x' \leq b', A'x^2 \leq b'$ and $A'\hat{x} = b'$, we have $A'x^1 = b'$ and $A'x^2 = b'$. Since $rank(A') = n$, $x^1 = x^2$. Therefore $\hat{x}$ is an extreme point.

Conversely suppose that $rank(A') < n$. Then there exists $d \in \mathbb{R}^n$ such that $A'd = 0$ and $d \neq 0$. For small $\epsilon > 0$, $\hat{x} - \epsilon d$ and $\hat{x} + \epsilon d \in \{x' \in \mathbb{R}^n : Ax \leq b\}$ so $\hat{x}$ is not an extreme point. □

Note that there are only finitely many extreme points of $\{x \in \mathbb{R}^n : Ax \leq b\}$ (for each subsystem $A'x \leq b'$ of $Ax \leq b$ with $rank(A') = n$ there is at most one solution to $A'x = b'$).

Geometry: Let $z_1, \cdots, z_k \in \mathbb{R}^n$. We say that x is a convex combination of $z_1, \cdots, z_k$ if there exist $t_1, \cdots, t_k \in \mathbb{R}$ such that

$$x = t_1 z_1 + \cdots + t_k z_k$$

$$t_1 + \cdots + t_k = 1$$

$$t \geq 0$$

We define the convex hull of $\{z_1, \cdots, z_k\}$ denoted $conv(z_1, \cdots, z_k)$ to be the set of all convex combination.

Claim: $conv(z_1, \cdots, z_k)$ is the smallest convex set that contain $z_1, \cdots, z_k$.

Theorem 8. Let $A \in \mathbb{R}^{m \times n}$ with $rank(A) = n$, and let $b \in \mathbb{R}^m$. Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ and $K = \{x \in \mathbb{R}^n : Ax \leq 0\}$ and let C be the convex hull of the extreme points of P. For each $x \in P$ there exist $z \in C$ and $d \in K$ such that $x = z + d$.

Example:

$$(p) \begin{cases} x_1 - x_2 & \leq 2 \\ -2x_1 + x_2 & \leq 1 \\ x_1, x_2 & \geq 0 \end{cases}$$

$$(k) \begin{cases} x_1 - x_2 & \leq 0 \\ -2x_1 + x_2 & \leq 0 \\ x_1, x_2 & \geq 0 \end{cases}$$

$$(4,3) \in P = (1,0) \in C + (3,3) \in K$$

Note: For each $z \in C$ and $d \in K$, $z + d \in P$.

Proof. Let $\hat{x} \in P$ and let $A'x \leq b'$ be the subsystem of $Ax \leq b$ that $\hat{x}$ satisfies with equality. We may assume that

1. If $\hat{x} \in P$ satisfies more of the constraints with equality than $\hat{x}$, then there exist $\tilde{z} \in C$ and $\tilde{d} \in K$ such that $\tilde{x} = \tilde{z} + \tilde{d}$
2. $\hat{x}$ is not an extreme point. (otherwise take $\hat{z} = \hat{x}$ and $\hat{d} = 0$)

By 2, $rank(A') < n$ (lemma 2) so there exists $d \in \mathbb{R}^n$ such that $A'd = 0$ and $d \neq 0$. Since $rank(A) = n, Ad \neq 0$. By possibly replacing d with $-d$, we may assume that some entry of Ad is negative.

Case 1: $Ad \leq 0$, (so $d \in K$). Choose $t_1 = \max(t \in \mathbb{R} : \hat{x} - td \in P)$ (since Ad has a negative entry, this is well defined).Let $x_1 = \hat{x} - t_1 d$. Now x_1 satisfies more of the inequalities $Ax \leq b$ with equality than $\hat{x}$ so by (1), there exists $z_1 \in C$ and $d_1 \in K$ such that $x_1 = z_1 + d_z$. Hence

$$\hat{x} = x_1 + t_1 d = z_1 + (d_1 + t_1 d)$$

Note that $z_1 \in C$ and $d_1 + t_1 d \in K$, as required.

Case 2: not case 1 (That is, Ad has both positive and negative entries) Let $t_1 = \max(t \in \mathbb{R} : \hat{x} - td \in P)$ and $t_2 = \max(t \in \mathbb{R} : \hat{x} + td \in P)$. Note that these are well defined and positive. Let $x^1 = \hat{x} - t_1 d$ and $x^2 = \hat{x} + t_2 d$. Note that x^1 and x^2 satisfy more constraints with equality than $\hat{x}$. So by (1), there exists $z^1, z^2 \in C$ and $d^1, d^2 \in K$ such that $x^1 = z^1 + d^1$ and $x^2 = z^2 + d^2$. Now,

$$\begin{aligned}\hat{x} &= \frac{t_2}{t_1+t_2}x^1 + \frac{t_1}{t_1+t_2}x^2 \\ &= \frac{t_2}{t_1+t_2}(z^1+d^1) + \frac{t_1}{t_1+t_2}(z^2+d^2) \\ &= (\frac{t_2}{t_1+t_2}z^1 + \frac{t_1}{t_1+t_2}z^2) + (\frac{t_2}{t_1+t_2}d^1 + \frac{t_1}{t_1+t_2}d^2)\end{aligned}$$

Since C and K are convex

$$(\frac{t_2}{t_1+t_2}z^1 + \frac{t_1}{t_1+t_2}z^2) \in C \text{ and } (\frac{t_2}{t_1+t_2}d^1 + \frac{t_1}{t_1+t_2}d^2) \in K$$

□

Corollary 1. Consider the LP

$$(p)\ \max\,(c^T x : Ax \leq b)$$

where $A \in \mathbb{R}^{m\times n}$ with $rank(A) = n$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$, Either

1. (P) is infeasible
2. There is an extreme point of $\{x \in \mathbb{R}^n : Ax \leq b\}$ that is optimal for (p), or
3. There is a feasible half line $\{x + \lambda d : \lambda \geq 0\}$ with $c^T d > 0$ (Hence (p) is unbounded)

Proof. Assume that (p) is feasible. Let γ be the maximum objective value of an extreme point of $\{x \in \mathbb{R}^n : Ax \leq b\}$. We may assume that there is a feasible solution $\hat{x}$ with

$$v^T \hat{x} > \gamma$$

By the theorem, we can write

$$\hat{x} = \hat{z} + \hat{d}$$

where $\hat{d} \in \{x \in \mathbb{R}^n : Ax \leq 0\}$ and $\hat{z}$ is in the convex hull of the extreme point of $\{x \in \mathbb{R}^n : Ax \leq b\}$. Note that $c^T \hat{z} \leq \gamma$. Hence $c^T \hat{d} > 0$ and $\hat{x} + \lambda\hat{d}$ is feasible for all $\lambda \geq 0$ so 3 is satisfied. □

Corollary 2. (Fundamental Theorem) Consider the LP

$$(p)\max\,(c^T x : Ax \leq b)$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$. If (p) is feasible and bounded, then (p) has an optimal solution.

Proof. By Lemma 1, we may assume that $rank(A) = n$. Then the theorem follows from corollary 2. $\square$

Corollary 3. (Unboundedness Theorem) Consider the LP:

$$(p) \max(c^T : Ax \leq b)$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$. Then (p) is a feasible half-line $\{\hat{x} + \alpha\hat{d} : \alpha \geq 0\}$ with $c^T\hat{d} > 0$

Proof. $(\leftarrow)$ easy

$(\rightarrow)$ By Lemma 1, we assume that $rank(A) = n$. Now the result is an immediate corollary of 1. $\square$

Polytopes

A set of the form $\{x \in \mathbb{R}^n : Ax \leq b\}$ is called a polyhedron. A bounded polyhedron is a polytope.

Corollary 4. Every polytope is the convex hull of its extreme points.

Proof. Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polytope. Since P is bounded if it does not contain a line. so $rank(A) = n$.

By the theorem, if P is not the convex hull of its extreme points, then there exists $\hat{x} \in P$ that can be written as $\hat{z} + \hat{d}$ where $\hat{z}$ is in the cones hull of the extreme points and $\hat{d} \in \{x \in \mathbb{R}^n : Ax \leq 0\}$ with $\hat{d} \neq 0$.

Then $\{\hat{x} + \alpha\hat{d} : \alpha \geq 0\}$ is contained in P- contradicting that P is bounded. $\square$

Corollary 5. For $z_1, \cdots, z_t \in \mathbb{R}^n$, $conv(z_1, \cdots, z_t)$ is a polytope.

We call an inequality $\alpha^T x \leq \beta$ valid for $conv(z_1, \cdots, z_t)$ if $\alpha^T z_i \leq \beta$ for each $i \in \{1, \cdots, t\}$.

Lemma 3. If $\hat{x} \in \mathbb{R}^n$ is not contained in $conv(z_1, \cdots, z_t)$, then there is a valid inequality such that $\alpha^T \hat{x} > \beta$. (That is, there is hyperplane that separating $\hat{x}$ from $z_1, \cdots, z_t$.

For example, $Q_0 = \{\begin{pmatrix} \alpha \\ \beta \end{pmatrix} : \alpha^T z_1 \leq \beta, \cdots, \alpha^T z_t \leq \beta\}$. Note that

1. This is a cone (since you can scale valid inequalities by non-negative numbers)

2. Q_0 is a polyhedron since it is defined by a finite set of linear inequalities.

Proof. Now define

$$A_1 = \{\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in Q_0 : -1 \leq \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \leq 1\}$$

Now Q_1 is a polytope, let $\begin{pmatrix} \alpha^1 \\ \beta^1 \end{pmatrix}, \cdots, \begin{pmatrix} \alpha^s \\ \beta^s \end{pmatrix}$ be the extreme points of Q_1. Let $P = \{x \in \mathbb{R}^n : (\alpha^1)^T x \leq \beta^1, \cdots, (\alpha^s)^T x \leq \beta^s\}$.

Claim: $P = conv(z_1, \cdots, z_t)$.

Note $z_1, \cdots, z_t \in P$ so $Conv(z_1, \cdots, z_t) \subseteq P$. Suppose that $P \neq conv(z_1, \cdots, z_t)$. Then there exists $x \in P - conv(z_1, \cdots, z_t)$.By separation theorem, there exists $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in Q_0$ such that

$$\alpha^T \tilde{x} > \beta$$

By solving we may assume that $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in Q_1$. By Corollary 4, there exist $\lambda_1, \cdots, \lambda_s \geq 0$ with $\lambda_1 + \cdots + \lambda_s = 1$ and $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \lambda_1 \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} + \cdots \lambda_s \begin{pmatrix} \alpha^s \\ \beta^s \end{pmatrix}$ Now,

$$\beta < \alpha^T \tilde{x} = \lambda_1 (\alpha^1)^T \tilde{x} + \cdots + \lambda_s (\alpha^s)^T \tilde{x} \leq \lambda_1 \beta^1 + \cdots + \lambda_s \beta^s = \beta$$

Contradiction. □

Corollary 6. A set $S \subseteq \mathbb{R}^n$ is a polytope if and only lit it is the convex hull of a finite set of points.

4.1 Caratheodory's Theorem

Let $S \subseteq \mathbb{R}^n$ be finite. Then any point in $conv(S)$ can be written as a convex combination of at most $n+1$ points in S.

Theorem 9. Let $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$. If the system $Ax \leq b$ is infeasible, then it contains an infeasible subsystem with at most $n+1$ inequalities.

Eventually, if $H_1, \cdots, H_m \subseteq \mathbb{R}^n$ are half spaces with empty intersection (that is, $H_1 \cap \cdots \cap H_m = \emptyset$), then some sub collection of at most $n+1$ of these half spaces has an empty intersection.

Corollary 7. If $P_1, \cdots, P_m \subseteq \mathbb{R}^n$ are polyhedra with empty intersection of $\leq n+1$ of these polyhedra has empty intersection.

Proof. Each of the polyhedra is the intersection of finitely many half-space. □

4.2 Helly's Theorem

If $S_1, \cdots, S_m \subseteq \mathbb{R}^n$ are convex sets with empty intersection then there is some sub collection of $\leq n+1$ of these sets has empty intersection.

Proof. We may assume that $m \geq n+1$, suppose that each sub collection of $n+1$ of the sets has nonempty intersection. Then there is a set $X \subseteq \mathbb{R}^n$ with $|x| \leq \binom{m}{n+1}$ so that each sub collection of $n+1$ of the sets contains an element of X. For $i \in \{1, \cdots, m\}$ define $P_i = conv(X \cap S_i)$. So $P_1, \cdots, P_m$ are polytopes by corollary 6.

By construction, every $n+1$ of these polytopes has nonempty intersection. So $P_1 \cap \cdots \cap P_m \neq \emptyset$ by corollary 7. Therefore $S_1 \cap \cdots \cap S_m \neq \emptyset$ □

4.3 Duality

Consider the LP $\begin{cases} \max & c^T x \\ \text{subject to} & Ax \leq b \end{cases}$

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

$$y^T Ax \leq y^T b$$

is a valid inequality for (P). If $y^T A = c^T$, then

$$c^T x \leq y^T b$$

. The dual of (p) is $\begin{cases} \min & b^T y \\ \text{subject to} & A^T y = c, y \geq 0 \end{cases}$

4.3.1 Weak Duality Theorem

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

Proof. $c^T x = (y^T A)x = y^T(Ax) \leq y^T b = b^T y$

□

Corollary 8. If (P) is unbounded, the (D) is infeasible.

Proof. Contrapositive is obvious. □

Corollary 9. If (D) is unbounded then (P) is infeasible.

Corollary 10. 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 (D) and $\tilde{y}$ is optimal for (D).

4.3.2 Strong Duality Theorem

If (P) has 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:

$$-c^T x + b^T \leq 0$$
$$Ax \leq b$$
$$-A^T y = -c$$
$$y \geq 0$$

If $\tilde{x}, \tilde{y}$ satisfies above, then $\tilde{x}$ is feasible for (P), $\tilde{y}$ is feasible for (D) and $c^T\tilde{x} \geq b^T\tilde{y}$ 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 the inequalities has no solution. □

Claims: If the inequalities has no solution then there exist $\bar{x} \in \mathbb{R}^n, \bar{y} \in \mathbb{R}^m$, and $\hat{z} \in \mathbb{R}$, satisfying

$$(2)\begin{cases} -c^T\bar{x} + b^T\bar{y} & < 0 \\ A\bar{x} & \leq \bar{z}b \\ A^T\bar{y} & = \bar{z}c \\ \bar{y} & \geq 0 \\ \bar{z} \geq 0 \end{cases}$$

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

Case 1: $\bar{z} \geq 0$. We can scale $(\bar{x}, \bar{y}, \bar{z})$ so that $\bar{z} = 1$. Now $(\bar{x}, \bar{y})$ satisfies the inequalities before (2). Contradiction.

Case 2: $\bar{z} = 0$. Now $\bar{y}^T A = 0$ and $\bar{y} \geq 0$. Since (P) is feasible $\bar{y}^T b \geq 0$. That is $b^T\bar{y} \geq 0$. Moreover, $A\bar{x} \leq 0$. However, (P) is bounded, so $c^T\bar{x} \leq 0$ so $-c^T\bar{x} + b^T\bar{y} \geq 0$ - contradiction (2).

	inf	UB	OPT
infeasible	Y	Y	X
unbounded	Y	X	X
optimal	X	X	Y

Consider the following LPs:

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

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

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

Claim: For any $\gamma\in\mathbb{R}$, the following are equivalent

1. (P1) has a feasible solution with objective value γ.
2. (P2) has a feasible solution with objective value γ.
3. (P3) has a feasible solution with objective value γ.

(P2) is in standard inequality form

$$(PSI)\begin{cases}\max & c^Tx\\ \text{subject to} & Ax\leq b\\ & x\geq 0\end{cases}$$

. (P3) is in standard equality form.

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

The dual of (PSI) is:

$$(DSI)\begin{cases}\min & b^Ty\\ \text{subject to} & A^Ty\geq C\\ & y\geq 0\end{cases}$$

The dual at (PSE) is

$$(DSE)\begin{cases}\min & b^Ty\\ \text{subject to} & A^Ty\geq C\end{cases}$$

Theorem 10. (Strong duality for standard inequality form): If (PSI) has an optimal solution $\bar{x}$, then (PSI) has an optimal solution $\bar{y}$ and $c^T\bar{x}=b^T\bar{y}$

Proof. Note that $\bar{x}$ is optimal for

$$(\tilde{P})\begin{cases}\max & c^Tx\\ \text{subject to} & \begin{pmatrix}A\\-I\end{pmatrix}x\leq\begin{pmatrix}b\\0\end{pmatrix}\end{cases}$$

The dual of $(\tilde{P})$ is

$$(\tilde{D}) \begin{cases} \min & b^T y \\ \text{subject to} & A^T y - s = c \\ & y, s \geq 0 \end{cases}$$

By the strong Duality Theorem, $(\tilde{D})$ has an optimal solution $(\bar{y}, \bar{s})$ and $c^T \bar{x} = b^T \bar{y}$. Note that, since $\bar{s} \geq 0$, $\bar{y}$ is feasible for (DSI). However $c^T \bar{x} = b^T \bar{y}$, so $\bar{y}$ is optimal for (DSI). □

Corollary 11. If (DSI) has an optimal solution, then (PSI) has an optimal solution $\bar{x}$ and $c^T \bar{x} = b^T \bar{y}$

(That is "the dual of (DSI) is (PSI)")

Proof. Note that $\bar{y}$ is optimal for

$$(P) \begin{cases} \max & -b^T y \\ \text{subject to} & -A^T y \leq c \\ & y \geq 0 \end{cases}$$

which is in standard inequality form. The dual of (P) is

$$(D) \begin{cases} \min & -c^T x \\ \text{subject to} & -Ax \geq -b \\ & x \geq 0 \end{cases}$$

By the theorem, (D) has an optimal solution $\bar{x}$ and $-c^T \bar{x} = -b^T \bar{y}$. Note that $\bar{x}$ is clearly optimal for (PSI). □

Theorem 11. (Strong duality for standard equality form) If (PSE) has an optimal solution, $\bar{x}$, then (DSE) has an optimal solution $\bar{y}$ and $c^T \bar{x} = b^T \bar{y}$.

4.4 Yet Other Theorem

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

The dual of (P) is

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

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

4.5 Complementary Slackness

Theorem 12. Complementary Slackness Theorem:

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

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

Let x be feasible for (P) and y be feasible for (D). Then $c^T x = b^T y$ if and only if for each $i \in \{1, \cdots, m\}$ either $y_i = 0$ or $|A_{i,1}, \cdots, A_{i,n}|x = b_i$

Proof. Consider (P) $\max(c^T : Ax \leq b)$ and its dual (D) $\min(b^T y : A^T y = c, y \geq 0\}$

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

$$b^T y - c^T x = y^T b - y^T ax = y^T(b - Ax) = \sum_{i=1}^{m} y_i(b_i - \sum_{j=1}^{n} A_{ij}x_j)$$

$$y_i \geq 0, (b_i - \sum_{j=1}^{n} A_{ij}x_j) \geq 0$$

so

$$y_i(b_i - \sum_{j=1}^{n} A_{ij}x_j) \geq 0$$

Equality holds if and only if either $y_i = 0$ or $\sum_{j=1}^{n} A_{ij}x_j = b_i$ □

4.5.1 Standard Inequality Form

Let x be feasible for (PSI) $\max(c^T x : Ax \leq b, x \geq 0)$ and y be feasible for (DSI) $\min(b^T y : A^T y \geq c, y \geq 0)$. Then $c^T = b^T y$ if and only if

1. For each $i \in \{1, \cdots, m\} |A_{i,1}, \cdots, A_{i,n}| x = b_i$ or $y_i = 0$; and
2. For each $j \in \{1, \cdots, n\} |A_{i,1}, \cdots, A_{i,n}| y = c_j$ or $x_j = 0$

4.5.2 Standard Equality Form

Let x be feasible for (PSE) $\max(c^T x : Ax = b, x \geq 0)$ and y be feasible for (DSE) $\min(b^T y : A^T y \geq c)$. Then $b^T y = c^T x$ if and only if for each $j \in \{1, \cdots, n\}$ either $|A_{1,j}, \cdots, A_{n,j}| y = c_j$ or $x_j = 0$.

Proof. Rewrite (DSE) as (DSE') $\max(-b^T y : -A^t y \leq -c)$ and apply the original complementary slackness theorem □

4.6 Optimality Theorem

Consider

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

and its dual

$$(P) \begin{cases} \max & b^T y \\ \text{subject to} & A^T y \geq c \end{cases}$$

where $A \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. We can assume that $rank(A) = m$ (without loss of generality)

4.6.1 Basic Solution

$A = |A_1, \cdots, A_n|$ and for $B \subseteq \{1, \cdots, n\}, A_b = |A_i : i \in B|$. We call B a basis if $|B| = m$ and $rank(A_B) = m$. For a basis B,

1. There is unique solution to $\begin{cases} Ax = b \\ x_j = 0, j \notin B \end{cases}$ This is a basic solution 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 x is a basic solution for B and $x \geq 0$, then we call x a basic feasible solution. If y is the basic dual solution for B and $A^T y \geq c$, then we call y a basic dual feasible solution.

Theorem 13. Optimality Theorem: Let $x \in \mathbb{R}^n$ be the basic solution for B and $y \in \mathbb{R}^m$ be the basic dual solution for B. Then $c^T x = b^T y$. Moreover, if x is feasible for (P) and y is feasible, then x is optimal for (P) and y is optimal for (D).

Remarks:

1. $x \in \mathbb{R}^n$ is an extreme point of $\{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ if and only if it is a basic feasible solution
2. $y \in \mathbb{R}^m$ is an extreme point of $\{y \in \mathbb{R}^m : A^T y \geq c\}$ if and only if it is a basic dual feasible solution.

Claim: A feasible solution for (P) us a basic feasible solution if and only if the columns of $|A_j : x_j \neq 0|$ are linearly independent.

Proof. From LHS, by definition.
From RHS, any linearly independent set extends to a basis. □

Proof. (Proof of optimality theorem)

$$b^T y - c^T x = x^T A^T y - x^T = x^T (A^T y - c) = x_B^T (A_B^T y - c_B)$$

□

Note: this proof works since x and y satisfy the complementary slackness conditions.

4.6.2 Finding a basic feasible solution

Input A feasible solution $\bar{x}$

Output A basic feasible solution

1. Step 1: If $[A_j : \bar{x}_j \neq 0]$ has independent columns, then STOP: Output $\bar{x}$
2. Step 2: Find $d \in \mathbb{R}^n$ such that
   (a) $Ad = 0$
   (b) $d_j = 0$ whenever $\bar{x}_j = 0$.
   (c) $d \neq 0$
3. Step 3: If also, replace d with $-d$. Let $\lambda = \max(t \in \mathbb{R} : \bar{x} - td \geq 0)$
   Replace $\bar{x}$ with $\bar{x} - \lambda d$. Repeat from Step 1.

Note that: $|support(\bar{x})|$ decreases with each iteration, so the algorithm terminates, and by the claim, the solution returned is basic.

4.7 Simplex Method

Goal: Given a basic feasible solution, solve (P). Example:

$$(P)(1)\begin{cases}\text{max} & 2x_1+3x_2\\ \text{subject to} & x_1+x_3-x_4=4\\ & x_2-x_3+2x_4=2\\ & -x_1+x_2+x_5=4\\ & x_1,\cdots,x_5\geq 0\end{cases}$$

Note that $B=\{1,2,5\}$ is a basis. For any feasible x,

$$2x_1+3x_2=2(4-x_3+x_4)+3(2+x_3-2x_4)=14+x_3-4x_4$$

(Here we are eliminating the basic variable from the objective function) so (P) is equivalent to

$$(P_1)(2)\begin{cases}\text{max} & 14+x_3-4x_4\\ \text{subject to} & x_1+x_2-x_4=4\\ & x_2-x_3+2x_4=2\\ & 2x_3-3x_4+x_5=6\\ & x_1,\cdots,x_5\geq 0\end{cases}$$

Note that (1) and (2) are equivalent linear system. Warning (P) and (P_1) have different duals. The basic solution is

$$\bar{x}=[4,2,0,0,6]^T$$

and has objective value $=14$. Note that x_3 has a positive coefficient in the objective function for (P_1). Set $x_3=t$ and $x_4=0$. Now solve for x_1,x_2,x_5.

$$\tilde{x}=[4,2,0,0,6]^T-t[-1,-1,-1,0,2]^T$$

which has objective value $=14+t$. Take $t=3$, we get $\tilde{x}=[1,5,3,0,0]^T$ with objective value 17. This is basis for $B=\{1,2,3\}$.

Eliminate the new basic variables from the objective function:

$$14+x_3-4x_4=14+\frac{1}{2}(6+3x_4-x_5)-4x_4=17-2.5x_4-0.5x_5 \quad (3)$$

For any non-negative x we get an adjective value ≤ 17 with respect to (3), there $\tilde{x}=[1,5,3,0,0]^T$ is an optimal solution.

4.8 Simplex Method

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

$rank(A) = m$

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

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

$$c^Tx = c^Tx - \bar{y}^T(Ax-b) = (c - A^T\bar{y})^Tx + \bar{y}^Tb = (c - A^T\bar{y})^Tx + \bar{\sigma}$$

we can rewrite (P) as

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

where

1. $\bar{c} = c - A^T\bar{y}$
2. $\bar{A} = (A_B)^{-1}A$, and
3. $\bar{b} = (A_B)^{-1}b$

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 $\bar{c} \leq 0$

Optimality: if $\bar{c} \leq 0$, then $\bar{x}$ is optimal for (P) and $\bar{y}$ is optimal for (D). (by (4)). Suppose that $\bar{c}_j \geq 0$ for some j. (Note that $j \notin B$ - by (2)). x_j is the entering variable.

Definition. *$\bar{d} \in \mathbb{R}$ by*

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

Note that the unique solution to

$$\begin{cases} \bar{A}_x = \bar{b} \\ x_j = t \\ x_i = 0, i \notin B \cup \{j\} \end{cases}$$

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

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

Update: Suppose that $\bar{d}$ has a negative entry. Choose $t = \max(\lambda \in \mathbb{R}^n : \bar{x} + \lambda\bar{d} \geq 0\}$ and replace $\bar{x}$ with $\bar{x} + t\bar{d}$. By our choice and t, there exists $i \in B$ such that $\bar{x}_i = 0$ and $\bar{d}_i < 0$. $\bar{x}_i$ is the leaving variables. 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 problems (P').

Termination

- There are $\leq \binom{n}{m}$ bases;
- at each iteration the objective value does not go down.
- 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. Hence $|support(\bar{x})| < m$. (recall $supports(\bar{x}) = \{i \in \{1, \cdots, n\} : \bar{x}_i \neq 0\}$).

A basic solution $\tilde{x}$ is non degenerate if $|support(\tilde{x})| = m$. (P) is non-degenerate if each of its basic solutions is non-degenerate. Note: The simplex method will terminate given any non-degenerate linear program (in $\leq \binom{n}{m}$ iterations)

Hirsch Conjecture (1957)

The distance between any two terraces in 1-skeleton of (P) is $\leq m$. (False, 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?

4.9 Perturbation Method

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

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

$rank(A)=m$ Consider

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

where $b'=\begin{pmatrix}b_1+\epsilon^1\\ b_2+\epsilon^2\\ \vdots\\ b_n+\epsilon^n\end{pmatrix}$ hence ϵ is a variable 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. For example, $1+\epsilon+100000\epsilon^2<1+2\epsilon$.

Claim: (P') is non degenerate.

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 non-zero real vector and the entries of b' are polynomials with distinctt degrees, each term of $\bar{x}_B$ is nonzero. □

Note that we can solve (P) using the Simplex Method since it is non-degenerate.

4.9.1 Another way to avoid cycling-Smallest Subscript Rule

Break ties when choosing entering and leaving variables by taking the one of minimum subscript.

Theorem 14. (Bland) The smallest subscript rule avoids cycling.

4.9.2 Feasibility

Consider

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

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 $v \geq 0$. Consider the following "auxiliary problem".

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

Note that:

1. $x = 0$, $s = b$ is a basic feasible solution to (P'), so we can solve this using the Simplex Method.

2. Since $s \geq 0, -s_1 - s_2 - \cdots - s_m \leq 0$, so (P') is bounded so the Simplex Method will terminate with an optimal solution $(\bar{x}, \bar{s})$.

3. if $\bar{s} = 0$, then $\bar{x}$ is feasible solution to (P).

4. If $\tilde{x}$ is feasible for (P), then $(\tilde{x}, 0)$ is an optimal solution for (P')

Hence, the optimal value for (P') is zero if and only if (P) has a feasible solution. Remark, if $(\bar{x}, 0)$ is a basic feasible solution for (P') thus $\bar{x}$ is a basic feasible solution for (P).

Farkars Lemma Exactly one of the following has a solution

1. $Ax = b, x \geq 0$

2. $A^T y \geq 0, b^T y < 0$

The dual of (P') is

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

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

4.10 Midterm Review

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

Separating Hyperplane Theorem (Farkas Lemma)

1. If $b \notin conv(z^1, \cdots, z^n)$, then there is a hyperplane separating b from $conv(z^1, \cdots, z^n)$.
2. Similar for $cone(z^1, \cdots, z^n)$.

Polyhedral Theory

Polyhedron: $\{x \in \mathbb{R}^n : Ax \leq b\}$. Polytope: bounded polyhedron. A polyhedral cone is $\{x \in \mathbb{R}^n : Ax \leq 0\}$.

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

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

Lemma 2 : Characterization of extreme point $\implies$ 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 $S \subseteq \mathbb{R}^n$ is a polyhedron cone, then there is a finite set $z \in \mathbb{R}^n$ such that $S = cone(z)$. The converse is also true.

For $S_1, S_2 \in \mathbb{R}^n$, define $S_1 + S_2 = \{a + b : a \in S_1, b \in S_2\}$.

Theorem C Let z be the set of extreme points of $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. If P does not contain a line then $P = conv(z) + \{x \in \mathbb{R}^n : Ax \leq 0\}$.

Theorem B and C implies there exist $Z, D \in \mathbb{R}^n$ finite such that

1. $P = conv(Z) + cone(D)$. (We used that P does not contain a line, it is easy to remove this condition.)
2. Note, we can scale so that $\|d\| = 1$ for each $d \in D$.

If (P) does not contain a line then there are unique minimal subsets $Z, D \in \mathbb{R}^n$ satisfying (1) and (2). Z is the set of extreme point. D is the set of extreme rays. $\implies$ "every polyhedron that does not contain a line is generated by its extreme points and its extreme rays."

Applications

Caratheodary's Theorem

Helly's Theorem

Linear Programming

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

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

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

Infeasibility Theorem (Karkas Lemma) (P) is infeasible if and only there exists $y \in \mathbb{R}^m$ satisfying $(A^T y = 0, b^T y < 0, y \ge 0)$.

Unboundedness Theorem (P) is unbounded if and only if (P) is feasible, and there exists $d \in \mathbb{R}^n$ satisfying $(Ad \le 0, c^T d > 0)$.

The dual of (P) is

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

Weak Duality Theorem: if $\bar{x}$ is feasible for (P) and $\bar{y}$ is feasible for (D) then $c^T\bar{x} \le b^T\bar{y}$. Ideally we could 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:

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

Suppose no such x, y exists.

By the Assignment questions, there exist $z \in \mathbb{R}, x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ satisfying:

$$-c^T x + b^T < 0$$

$$Ax \le bz$$

$$-A^T y = -cz$$

$$y \geq 0$$

$$z \geq 0$$

Claim: $z = 0$.

Proof. Otherwise we can scale to get $z = 1$, and then (x, z) satisfies (1) - contradiction. □

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 (P) 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

$$c^T \bar{x} = b^T \bar{y}$$

Application of duality

Theorem 15. 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: $A'\bar{x} = b', rank(A') = n$ and A' is $n \times n$. ($\bar{x}$ may satisfy some of $A''x \leq b''$ with equality)

Let $c = (A')^T 1$, $\alpha = c^T \bar{x} = 1^T A'x = 1^T b'$. $H = \{x \in \mathbb{R}^n : c^T x \geq \alpha\}$

Now consider the LP:

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

and its dual

$$(D) \begin{cases} \text{min} & (b')^T y + (b'')^T z \\ & (A')^T y + (A'')^T = z \\ & y, z \geq 0 \end{cases}$$

Let $\bar{y} = 1$ and $\bar{z} = 0$.

Now $\bar{x}$ is feasible for (P), $(\bar{y}, \bar{z})$ is feasible for (D) and $c^T\bar{x} = (b')^T\bar{y} + (b'')^T\bar{z} = \alpha$, so $\bar{x}$ is optimal for (P) and $(\bar{y}, \bar{z})$ is optimal for (D). Consider another optimal solution $\tilde{x}$ for (D). Note that $\bar{y} > 0$, so by the complementary slackness condition, $A'\tilde{x} = b'$. 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 $\bar{x} \notin \mathbb{Z}^n$, there exists $c \in \mathbb{Z}^n$ such that $\bar{x}$ is an optimal solution to $\max(c^T x : x \in P)$ and $c^T\bar{x} \notin \mathbb{Z}$.