CO250 course note

Chenxuan Wei

Sep 2021

Contents

1 Formulations

1.1 Overview

1. Important Special Cases

   - Abstract optimization problem (P):
     - Given: a set $A \in R^n$ and a function $f : A \to R$
     - Goal: find $x \in A$ that minimizes and maximizes f
   - Linear Programming (LP):
     - A is implicitly given by linear constraints
     - f is a linear function
   - Integer Programming (IP)
     - We focus on integer max/min
   - Non linear Programming (NLP)
     - A is non-linear constraints
     - f is a non-linear function

2. Typical Workflow

   - English language description of practical problem
   - Develop a mathematical model for the problem
   - feed the model and data into solver

3. Ingredients of a Math Model

   - Decision Variables: Capture unknown information
   - Constraints: Describe which assignments to variables are feasible
   - Objective function: A function of the variables that we would like to maximize/minimize

1.2 LP Models

1. Constrained Optimization
   In this course, we consider optimization problems of the following form:
   $min\{f(x) : g_i(x) \leq b_i, (1 \leq i \leq m), x \in R^n\}$
   where:

   - $n, m \in N$
   - $b_1, \ldots, b_m \in R$
   - $f, g_1, \ldots, g_m$ are functions with $R^n \to R$, and we assume all of them are affine

2. Definition: affine
   A function $f : R^n \to R$ is affine if $f(x) = \sum_{i=1}^{m} \alpha_i x_i + \beta$ for $\alpha_i \in R^n$, $\beta \in R$
   it is linear if $\beta = 0$

3. Definnition: Linear program
   The optimization problem
   $\min\{f(x) : g_i(x) \leq b_i, \forall 1 \leq i \leq m, x \in R^n\}$
   is called a linear program if f is affine and $g_1, \ldots, g_m$ is finite number of linear functions
   Note:

   - often write LPs more verbosely
   - often give non-negativity constraints separately
   - constrain maybe max or min
   - $x \leq \vec{0}$ shows all x_i are non-negative
   - must be $\leq$ or $\geq$ no $<$ and $>$
   - finite number of constrainsts

4. Multiperiod Models

   - Time is split into periods
   - We have to make a decision in each period
   - All decisions influence the final outcome.

1.3 IP models

1. Definition of Integer programming
    - Added integer constraints to LP
    - Difficult to solve
2. Binary variable
   usefu to modify logical constraints
3. Set
   Create a set for choose from finite values

1.4 Optimization of Graphs

1. Basic definition of graphs

   - vertices: V
   - edges: $E = \{uv : u, v \in V\}$
   - u and v are adjacent if $uv \in E$
   - u and v are endpoints of $uv \in E$
   - $e = v_1v_2$ incident to $u \in V$ if u is an endpoint of e

2. Path
   A s,t-path in $G = (V, E)$ is a sequence of edges
   $v_1v_2, v_2v_3, \ldots, v_{k-1}v_k$
   where $v_1 = s, v_k = t$
   and $v_i \neq v_j$ if $i \neq j$

3. Matching

   - Definition
     A set of edges $M \subseteq E$ is matching if
     $\forall e_1, e_2 \in M where e_1 = v_{11}v_{12}, e_2 = v_{21}v_{22}$ we must have $v_{11} \neq v_{12} \neq v_{21} \neq v_{22}$
   - Perfect matching
     A matching is perfect if the following happened
     Let $M = \{e_i = v_{i1}v_{i2}\}$, $\forall v \in V, v \in \{v_{ij}\}$
   - Definition for $\delta(v)$
     $\delta(v) = \{e \in E : e = uv, u \in V$
   - Another def for perfect matching Let $G = (V, E)$, $M \subseteq E$ is a perfect matching $\iff |M \cap \delta(v)| = 1, \forall v \in V$

1.5 Shortest Paths

1. Cut
   $\delta(S)$ is an s, t cut if $s \in S$ and $t \notin S$

2. Remark
   if $S \subseteq E$ contains at least one edge from every s, t cut, then S contains an s, t path.

1.6 Nonlinear Programs

1. Definition of NLP
   is in the from of

   - min/max f(x)
   - s.t.
     - $g_1(x) \leq 0$
     - $g_2(x) \leq 0$
     - ...
     - $g_m(x) \leq 0$
   - where
     - $x \in R^n$
     - $f : R^n \rightarrow R$
     - $g_i : R^n \rightarrow R$

2 Solving Linear Programs

2.1 Possible Outcomes

1. Definition of feasible solution
   An assignment of values to each of the variables is a feasible solution if all the constraints are satisfied

2. Definition of feasible
   if there is at least one feasible solution, infeasible otherwise

3. Defintion of optimal solution

   - For maximization problem, an optimal solution is a feasible solution that max the objective function
   - For minimization problem, an optimal solution is a feasible solution that min the objective function

4. Definition of bounded

   - For maximization problem is unbounded if for every value M, there exists a feasible solution with objective value greater than M
   - For minimization problemis unbounded if for every value M, there exists a feasible solution with objective value smaller than M

5. Fundamental Theorem of Linear Programming
   For any linear Program, exactly one of the following holds:

   - It has an optimal solution
   - It is infeasible
   - It is unbounded

2.2 Certificates

1. Farkas's lemma
   There is no solution to $Ax = b, x \geq 0 \iff \exists y \in R^n, y^\intercal A \geq 0^\intercal$ and $y^\intercal b < 0$

2. Proposition 2.2
   The linear program, $\max\{c^\intercal x : Ax = b, x \geq 0\}$ is unbounded if we can find $\vec{x}$ and r such that

   - $\vec{x} \geq 0$
   - $r \geq 0$
   - $A\vec{x} = b$
   - $Ar = 0$
   - $c^\intercal r > 0$

2.3 Standard Equality Forms

1. Definition of SEF

   - it is a max problem
   - for every variable x_j, we have $x_j \geq 0$
   - all other constraints are equality constraints

2. Definition of equivalanet LPs
   Linear programs (P) and (Q) are equivalanet if

   - (P) infeasible $\iff$ (Q) infeasible
   - (P) unbounded $\iff$ (Q) unbounded
   - can create a optimal solution of (P) from optimal solution of (Q)
   - can create a optimal solution of (Q) from optimal solution of (P)

3. Theorem
   Every LP is equivalanet to an LP in SEF

4. Convert to SEF

   - max
     $\times -1$ to objective function
   - change to equality
     add a new variable
     ex. $x_1 - x_2 \leq 7$ is equvilent to $x_1 - x_2 + s = 7, s \geq 0$
   - Free variables
     idea:change the variable to the difference of 2 non-negative variable
     ex. $x_3 = a - b, a \geq 0, b \geq 0$
     Then rewrite everything in terms of a and b

2.4 Basis

1. Notation
   Let B be a subset of column indices, then A_B is a column sub-matrix of A in dexed by set B
   A_j is the j column of A

2. Basis
   Let B be a subset of column indices, B is a basis if
   - A_B is a square matrix
   - A_B is non-singular

3. Theorem on week 5 lecture 1 slide 10
   Max number of independent columns = max number of independent rows

4. Basic Solutions
   Let $Ax = b$ be a equation
   - Basic variable
     Let B be a basis for A
     - if $j \in B$, then x_j is a basic variable
     - if $j \notin B$, then x_j is a non-basic variable
   - solution
     x is a basic solution for basis B if
     - $Ax = b$
     - $x_j = 0$ if $j \notin B$
   - Uniquess of basic solutions
     Let B be a basis for A, then $Ax = b$ has a uniqe basic solution for B
   - feasible
     A basic solution x of $Ax = b$ is feasible if $x \geq 0$

2.5 Canonical Forms

1. Definition of Canonical form
   Let $maxc^\intercal x : Ax = b, x \geq 0$ be a LP
   Let B be a basis for A
   Then this is a canonical form for B if
   - $A_B = I$ (P1)
   - $c_j = 0 \forall j \in B$ (P2)
2. Proposition
   For any basis B, there exists (P') in canonical form for B such that
   - (P) and (P') have the same feasible region
   - feasible solution have the same objective value for (P) and (P')
3. rewriting constrain (P1)
   Assume we have the LP: $maxc^\intercal x : Ax = b, x \geq 0$
   And a Basis B
   Then mutiply A_B^{-1} to both side of the constrain, we get $A_B^{-1}Ax = A_B^{-1}b$
   Then the new $A' = A_B^{-1}A$ will have $A_B = I$
4. rewrite the objective function (P2)
   Do the following steps
   - Assume $A \in M_{m\times n}$, then create $Y = (y_1, \ldots, y_m)^\intercal$, and get $Y^\intercal Ax = Yb$
   - Choose y_i so that $\forall j \in B, c_j = 0$
   - The new objective function is $(c^\intercal - Y^\intercal A)x + Y^\intercal b$
   - In order to find y_i, solve the equation $Y^\intercal A_B = c_B^\intercal$, means $Y = A_B^{-\intercal} c_B$
5. Proposition 2.4
   for a LP $max\{c^\intercal x, Ax = b, x \geq 0\}$
   the canonical form for basis B is
   $max\{(c^\intercal - y^\intercal A)x + y^\intercal b, A_B^{-1}Ax = A_B^{-1}b, x \geq 0\}$
   where y = $A_B^{-\intercal} c_B$

2.6 Formalizing the simplex

1. idea
   Let $k \notin B$, such $c_k > 0$, then make x_k as large as possible, keep all other $x_i, i \notin B$ at 0

2. Algorithm
   Let a LP be $max\{c^\intercal x, Ax = b, x \geq 0\}$, and input a feasible basis B

   - Write in Basis B's canonical form
   - Find a better Basis or get required outcome

3. Algorithm Start with $\max\{c^\intercal x, Ax = b, x \geq 0\}$ in canconical form with Basis B

   - Choose $k \notin B$ with $c_k \geq 0$ and this will be entering variable
   - pick $x_B = b - tA_k$ and find the possible t that allowed $x \geq 0$ still holds
   - the $x_i = 0$ after choose t will be leaving, and we get a new basis
   - get the canconical from for the new basis
   - if $c^\intercal \leq 0$, stops with optimal solution
   - if $A_{B'} \leq 0$, stops with unbounded

4. Bland's rule

   - if we have a choice for entering value, choose small one
   - if we have a choice for leaving value, choose small one

2.7 Find a Feasible Solution

1. Algorithm

   - Rearrange the equation such that RHS is non-negative
   - Construct an auxiliary problem
     For a LP $max(c^\intercal x, Ax = b, x \geq 0)$
     Assume $A \in M^{m \times n}$
     Construct $x_{n+1}, \ldots, x_{n+m}$ as variable, and find the optimal solution of the following LP by Simplex
     $max(z = \sum_{t=1}^{m} x_{n+t}, (A|I_m)x = b, x \geq 0)$
     if the optimal solution x has $x_B = 0, B = \{n+1, \ldots, n+m\}$, then the original LP have a feasible solurion, it not, then it is infeasible

2. Proposition

   - if $z = 0$ then it has solution.
   - if $z > 0$, then it has no slution

2.8 Half-Space and Convexity

1. Feasible region
   For an optimization problem, it is the set of all feasible solutions

2. Polyhedron
   it $\exists A \in M^{m\times n}$ and $b \in R^n$ where $P = \{x : Ax \leq b\}$, then P is a polyhedron

3. Definition of hyper and half
   Let $\alpha \neq 0$ be a vector and $\beta \in R$

   - $\{x : \alpha^\intercal x = \beta\}$ is a hyper plane
   - $\{x : \alpha^\intercal x \leq \beta\}$ is a half space

   A polyhedron is the intersection of finite set of halfspaces

4. Norm
   Let $a, b \in R^n$ then
   $a^\intercal b = ||a||||b||cos(\theta)$
   where$||..||$ is the norem and θ is the angle between a and b

5. Traslate
   Let $S, S' \subseteq R^n$ (hence they are lines)
   then S' is a translate of S is $\exists p \in R^n$ and
   $S' = \{s + p : s \in S\}$
   and it has the following propositions

   - if $H := \{x : a^\intercal x = b\}$ $H' := \{x : a^\intercal x = 0\}$. then they are translate
   - if $H := \{x : a^\intercal x \leq b\}$ $H' := \{x : a^\intercal x \leq 0\}$. then they are translate

6. Dimension
   The dimension of a hyperplane in R^n is n - 1

7. A polyHedron has no "dents" and no "holes"

8. Line
   Let $x_1, x_2 \in R^n$, the line through x_1 and x_2 is
   $L = \{x = \lambda x_1 + (1 - \lambda)x_2 : \lambda \in R\}$
   The line segment is $S = \{L : 0 \leq \lambda \leq 1\}$

9. Convex
   A set $S \subseteq R^n$ is convex if , $\forall x_1, x_2 \in S$ $\exists$ line segment between x_1 and x_2 in S

10. Proposition

    - Let H be a halfspace, then H is convex
    - If P is polyhedron, then P is convex
    - The intersection of 2 convex set is convex

2.9 Extreme Points

1. Properly contained
   A point $x \in R^n$ is properly contained in the line segment L if
   - $x \in L$
   - x is distinct from end points in L
2. Extreme Points
   Let S be a convex set and $x \in S$
   x is not an extreme point if $\exists L \subseteq S$, where L properly contains x
3. Tight
   Let $P = \{x : Ax \leq B\}$ be a polyhedron and let $x \in P$
   - A constraints is tight got x if it is satisisfied with equality and
   - the set of all tight constraints is denoted as $\overline{A}x \leq \bar{b}$
4. Theorem
   Let $P = \{x \in R^n : Ax \leq b\}$ be a polyhedron and let $x \in P$
   - if $rank(\overline{A}) = n$, then x is not an extreme points
   - if $rank(\overline{A}) < n$, it is not an extreme points

3 Duality Through Examples

3.1 Duality Throught Examples

1. Find an Intuitive Lower Bound
   Let the graph $G = \{V, E\}$ and for $U \subseteq V$
   $\lambda(U) = \{uv \in E : u \in U, v \notin U\}$ is called an s, t if $s \in U, t \notin U$

2. Assignment of width to cuts
   A feasiblewidth assignment $\{y_U : \lambda(U), s, t - cut\}$ is feasible if
   $\forall e \in E$, the total width of all cuts containing e is no more than c_e

3. Propostion
   Let y be a feasible width assignment, then any s, t pathc must have length at least
   $\sum(y_U, Us - tcuts)$

3.2 Weak Duality

1. Find the lower bound of the LP
   Assume we have LP $\min\{c^\intercal x, Ax \geq b, x \geq 0\}$
   then we want to find y such that $b^\intercal y$ will be the best lower bound by LP
   $\max\{b^\intercal y, A^\intercal y \leq c, y \geq 0\}$
   and that is the dual of primal LP

2. Week Duality Theorem
   if x is feasible for P and y is feasible for D, then $b^\intercal y \leq c^\intercal x$

3.3 Shortest Path Algotithm

1. Arcs - oreded pair
   A Arcs from u to v is denote as $\vec{uv}$ and draw it as arrow from u to v
   A directd path is a sequence of arcs

2. Slack
   Let y be a feasible dual solution, the slack of an edges $e \in E$ is defined as
   $slack_y(e) = c_e - \sum(y_U : \delta(U), s, t - cut, e \in \delta(U))$
   This is just the extra part of a c_e

3. Shortest Path, Algorithm 3.2

   - Input: Graph $G = (V, E)$, cost $c_e \geq 0, \forall e \in E$, $s, t \in V, s \neq t$
   - Output : Shortest Path
   - $y_w = 0$ for all cut, set $U = \{s\}$
   - While $t \notin U$:
     - Let $ab \in E$ be an edge in $\delta(U)$ of smallest slack and where $a \in U, b \notin U$
     - $y_U = slack(ab)$
     - $U = U \cup \{b\}$
     - change ab to $\vec{ab}$
   - we can easily see path through arcs

3.4 Correctness

1. Proposition
   Let y be a feasible dual solution, and P is a s,t -path, P is ta shortest path if

   - all edges on P are equality edges, and
   - every active cut $\delta(U)$ has exactly one edges of P

4 Duality Theory

4.1 Weak Duality

1. General case of dual LP,Primal-Dual Pair
   Note: P_{max} on the left, P_{min} on the right

   - $\leq$ constraint $\iff$ ≥ 0 variable
   - $=$ constraint $\iff$ free variable
   - $\geq$ constraint $\iff$ ≤ 0 variable
   - ≥ 0 variable $\iff$ $\geq$ constraint
   - free variable $\iff$ $=$ constraint
   - ≤ 9 variable $\iff$ $\leq$ constraint

2. Theorem
   Let P_{max} and P_{min} represent the above. if $\overline{x}$ and $\overline{y}$ are feasible for two LPs, then
   $c^{\intercal}\overline{x} \leq b^{\intercal}\overline{y}$
   if equality holds, $\overline{x}$ is optimal for P_{max}, $\overline{y}$ is optimal for P_{min}

4.2 Strong Duality

1. Strong Duality Theorem
   If P_{max} has an optimal solution $\overline{x}$, then P_{min} has an optimal solution $\overline{y}$ such that $c^{\mathsf{T}}x = b_{\mathsf{T}}y$

2. Feasibility Version
   Let P, D be primal-dual pair, if both are feasible, then both have optimal solutions

3. Relationship between P and D

DP	Optimal	unbounded	infeasible
Optimal	possible	impossible	impossible
Unbounded	impossible	impossible	possible
infeasible	impossible	possible	possible

4.3 Geometric Optimality

1. Complementary Slackness
   Let $\overline{x}$, $\overline{y}$ be feasible for P and D, they are optimal $\iff$
   - $y_i = 0$
   - i th constraint of P is tight for $\overline{x}$

5 Solving Integer Programs

5.1 Convex Hull

1. Convex hull
   Let $C \subseteq R^n$, the convex hull of C is the smallest convex set that contains C
   And $\forall C$ there is a unique convex hull

2. Meyer's theorem
   Let $P = \{x : Ax \leq b\}$ where A, b are rational
   then the convex hull of all integer points in P is a polyhedron

3. Theorem
   Let $IP = \max\ \{c^\intercal x : Ax \leq b, x \in Z\}$,
   the convex hull of all feasible solution of IP = $\{x : A'x \leq b\}$
   LP = $max\{c^\intercal x : A'x \leq b', xinteger\}$

   - IP is infeasible $\iff$ LP is infeasible
   - IP is boundede $\iff$ LP is unbounded
   - an optimal solution to IP is an optimal solution to LP
   - an extreme optimal solution to LP is an optimal solution to IP

5.2 Cutting Planes

1. Cutting Planes Algorithms
   Let IP be $\max\{c^\intercal x, Ax \leq b, x \in Z\}$

   - Let P = $\max\{c^\intercal x, Ax \leq b\}$
   - if P infeasible, stop
   - Let $\overline{x}$ be optimal solution to P
   - if $\overline{x}$ is integral, stop, we found a solution
   - Find a cutting plane (new constaint) such all other feasible solution for IP is still in, but $\overline{x}$ is not in
   - Add the new constraint to IP, repeat

2. Find cutting plane constrain
   Let a constrain be $f(x) \leq b$
   floor all non-integer consstrain and b
   new $f(x) \leq b'$ will be all integer constrain that satisified requirement above

6 Non-linear optimization

6.1 Convexity

1. Non-linear Pgoram
   min f(x)
   s.t.

2. $g_i(x) \leq 0$

3. Local optimum
   Let $\min\{f(x) : x \in S\}$ be a NLP
   $x \in S$ is a local optimum if $\exists \delta > 0$ such
   $\forall x' \in S, ||x' - x|| \leq \delta$, and we have $f(x) \leq f(x')$

4. Proposition
   If S is convex and x is local optimum, then x is optimal

5. Convex function
   $f : R^n \to R$ is convex if $\forall a, b \in R^n$
   $f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b)$

6. Proposition
   If g is convex function
   $S = \{x \in R^n, g(x) \leq b\}$ is convex set

7. epigraph
   Let f be a function
   $epi(f) = \{(y, x)^\intercal : y \geq f(x), x \in R^n\} \subseteq R^{n+1}$

8. Propostion
   f is convex $\iff$ epi(f) is convex

6.2 The KKT theorem

1. subgradient
   Let f be convex function
   $s \in R^n$ will be subgradient of f at $\overline{x}$ if
   $h(x) = f(\overline{x}) + s^\intercal(x - \overline{x}) \leq f(x)$ for all $x \in R^n$

2. Supproting
   Let $C \subseteq R^n$ be convex set and $\overline{x} \in C$
   The halfspace $F = \{x : s^\intercal x \leq B\}$ is supporting C at $\overline{x}$ if

   - $C \subseteq F$
   - $s^\intercal \overline{x} = \beta$, that is, $\overline{x}$ is on the boundary of F

3. Proposition
   Let

   - g be convex $g(\overline{x}) = 0$
   - s be subgradient of g at x
   - $C = \{x : g(x) \leq 0\}$
   - $F = \{x : h(x) = g(\overline{x}) + s^\intercal(x - \overline{x}) \leq 0\}$

   Then f is supporting halfspace of C at $\overline{x}$

4. Proposition
   Let $\min\{c^\intercal x, g_i(x) \leq 0\}$ be NLP
   All g_i are convex
   $\overline{x}$ is feasible
   all $g_i(\overline{x}) = 0$
   s_i be subgradient
   if $-c \in cone\{s_i : i \in I\}$, then $\overline{x}$ is optimal

5. Slater point if all $g_i(\overline{x}) < 0$, then $\overline{x}$ is a slater point

6. Karush-Kuhn-Kucker (KKT) theorem
   Let $\min\{c^\intercal x, g_i(x) \leq 0\}$ be NLP
   Suppose

   - g_i are convex
   - there is a slater point
   - $\overline{x}$ is feasible
   - I $= \{i : g_i(\overline{x}) = 0\}$
   - $\forall i \in I$, there is a gradient $\Delta g_i(\overline{x})$of g_i at $\overline{x}$

   Then $\overline{x}$ is optimal $\iff -c \in cone\{\Delta g_i(\overline{x}) : i \in I\}$