CO 250 Review Questions - Solutions

Introduction to Optimization

David Duan, 2019 Spring

Contents

I LP/IP/NLP Formulation

Liner Program

- *Define what is meant for an optimization problem to be a linear program.*
  A linear progrma is the problem of maximizing or minimizing an *affine* function subject to a *finite* number of *linear* constraints.
- *What is an affine function?*
  An affine function is a linear function with a translation.
- *Can a LP take strict inequalities as constraints?*
  No, all constraints must be $=$, $\leq$, or $\geq$.

LP Formulation

- *What is the key to a multi-period model?*
  For a multi-period model, we need balance constraints.
- *How do we minimize the average of something?*
  Let x_i denote the target variables for each i. We use $\min \sum x_i$ as objective.
- *How do we minimize the maximum of something?*
  We add a constraint $\forall i : y \geq x_i$ and use $\min y$ as objective.
- *How do we minimize the maximum difference of something?*
  We add a constraint $\forall i, j : y \geq |x_i - x_j|$ and use $\min y$ as objective.

Assignment Problem

- *Let x_{ij} denote whether employee $i \in I$ gets assigned the job $j \in J$. Constraint?*
  $\sum_{j \in J} x_{ij} = 1 \ (i \in I); \quad \sum_{i \in I} x_{ij} = 1 \ (j \in J); x_{ij} \in \{0, 1\}$.
- *Let c_{ij} denote the resource needed for employee i to finish job j. Objective function?*
  The objective function is $\min \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij}$.

Knapsack Problem

- *How do we express "at least one of the following two conditions has to be satisfied"?*
  We introduce a binary variable y. If $y = 1$, we want the first condition to be true; if $y = 0$ we want the second condition to be true.
- *For example, how do we express "a total of at least four crates of type 1 or type 2 is selected, or a total of at least four crates of type 5 or type 6 is selected"?*
  Let y be the decision variable. We can express this as $x_1 + x_2 \geq 4y$ and $x_5 + x_5 \geq 4(1 - y)$.

Minimum Cost Matching

- *Give the IP for minimum cost matching.*

$$
\begin{array}{lll}
\min & \sum(c_e x_e : e \in E) & \\
s.t. & \sum(x_e : e \in \delta(v)) \leq 1 & (v \in V) \\
& x_e \geq 0 & (e \in E) \\
& x_e \in \mathbb{Z} &
\end{array}
$$

- *Explain the main constraint.*
  The number of edges incident to v that are selected is $\sum(x_e : e \in \delta(v))$. Since we want a matching, the number must be at most 1.

Shortest Path

- *What is the definition of a s, t-cut?*
  A s, t-cut for a subset of vertices $U \subseteq V$ is $\delta(U) = \{uv \in E : u \in U, v \notin U\}$.
- *Describe the IP for shortest path problem.*
  The IP for shortest path problem is given by

$$
\begin{array}{lll}
\min & \sum(x_e c_e : e \in E) & \\
s.t. & \sum(x_e : e \in \delta(U)) \geq 1 & (U \subseteq V, s \in U, t \notin U) \\
& x_e \geq 0 & (e \in E) \\
& x_e \in \mathbb{Z} & (e \in E)
\end{array}
$$

- *Describe the constraint matrix for the IP. In particular, when does an entry equal 1?*
  Each row of the matrix corresponds to an s, t-cut and each column corresponds to an edge; $A_{U,e} = 1$ when e has exactly one end in U.

NLP Formulation

- *Describe how we represent a statement as a NLP. In particular,*
  - *What is speical about the NLP's feasibility?*
    The NLP is trivially feasible by setting all variables 0.
  - *What is special about the NLP's objective and feasible solutions?*
    Any feasible solution is non-negative as the objective is a sum of squares.
  - *What is special about the NLP's objective value?*
    The value of a solution is zero iff it satisfies (or violates) the statement. Thus, the statement is proved (or disproved) if the value 0 cannot be attained.

II Solving Linear Programs

Linear Programming

- *State the fundamental theorem of linear programming.*
  A linear program is either infeasible, unbounded, or feasible with optimal solutions.
- *How do we show that a LP is infeasible?*
  We use Farka's Lemma: Let $A \in \mathbb{R}^{m,n}$ and $b \in \mathbb{R}^m$. Exactly one of the two holds:
    1. Feasible: There exists $x \in \mathbb{R}^n$ such that $Ax = b$ and $x \geq \mathbf{0}$.
    2. Infeasible: There exists $y \in \mathbb{R}^m$ such that $y^T A \leq \mathbf{0}$ and $y^T b > \mathbf{0}$.

  We call $y \in \mathbb{R}^m$ above a certificate of infeasibility for the LP.
- *How do we show that a LP is unbounded?*
  We find $\bar{x} \geq \mathbf{0}$ and $r \geq \mathbf{0}$ such that $A\bar{x} = b$, $Ar = \mathbf{0}$, and $c^T r > 0$. Then, $t \to \infty$ implies $A(\bar{x} + tr) = b$ and $c(\bar{x} + tr) \to \infty$, i.e., the constraint is satisfied and the objective value is unbounded from above.
- *How do we show that a feasible solution $\bar{x}$ is optimal for a LP ($c \leq 0$, $z(x) = c^T x + m$)?*
  We prove by contradiction. For any feasible solution $x \geq \mathbf{0}$, $c^T x + m \leq m$. This is used in Simplex to show that we have arrived at the optimal solution for the given LP.

Standard Equality Form

- *When is a LP in standard equality form?*
  A LP is in SEF when it is (1) a maximization problem with (2) non-negative variables and (3) equality constraints.
- *Why do we want a LP to be in SEF?*
  The Simplex algorithm can solve any LP as long as it is in SEF.
- *State the three steps to rewrite a LP into SEF.*
  Replace $\min f$ with $\max -f$; use slack variables to replace inequalities by equalities; express any free variable as the difference between two non-negative integers.

Bases and Basic Solutions

- *What is a basis?*
  A set B is a basis for A when the columns of A_B are linearly independent.
- *What is a basic variable?*
  Given basis B for A, x_ℓ is a basic variable when $\ell \in B$.
- *What is a basic solution?*
  Given basis B for A, $x_k = 0$ when $k \notin B$.
- *What is a basic feasible solution?*
  A basic feasible solution is a basic solution with $x \geq \mathbf{0}$.
- *Does each basis have a unique basic solution?*

Yes. Observe $b = Ax = \sum_j A_j x_j = \sum_{j \in B} A_j x_j + \sum_{j \notin B} A_j x_j = \sum_{j \in B} A_j x_j + 0 = A_B x_B$. Since B is a basis, A_B is non-singular, so A_B^{-1} exists. Hence, $x_B = A_B^{-1} b$ is unique.

- *Does each basic solution correspond to a unique basis?*
  No. Each basic solution can correspond to many different bases.

Canonical Form

- *When is a LP in canonical form?*
  Given a basis B of A, a LP is in canonical form when $A_B = I$ and $c_j = 0$ for $j \in B$.
- *Why do we want a LP to be in canonical form?*
  Given a basis B of A, there exists a unique basic feasible solution $\bar{x}$ to the LP; having the LP in canonical form allows us to easily check whether $\bar{x}$ is optimal.
- *Derive the procedure to rewrite the constraint into canonical form.*
  Multiply both sides by A_B^{-1}: $Ax = b \implies A_B^{-1} Ax = A_B^{-1} b$.
- *Derive the procedure to rewrite the objective function into canonical form.*
  Manipulate constraint: For any $y \in \mathbb{R}^m$, we have $y^T Ax = y^T b \implies 0 = -y^T Ax + y^T b$.
  Manipulate objective: $z = c^T x \implies z = [c^T - y^T A]x + y^T b =: \bar{c}^T x + \bar{z}$.
  Choose y such that $\bar{c}_j = 0$ for $j \in B$: $\bar{c}_B = \mathbf{0} = c_B^T - y^T A_B \implies y = A_B^{-T} c_B$.
- *Given* $\max\{c^T x : Ax = b, x \geq \mathbf{0}\}$*, describe its equivalent in canonical form.*
  The key here is to recognize $c_B = \mathbf{0}$ and $A_B = I$. We rewrite the objective function as $z = c_B^T x_B + c_N^T x_N + \bar{z} = c_N^T x_N + \bar{z}$ and the constraint $A_B x_B + A_N x_N = x_B + A_N x_N = b$. The equivalent is thus $\max\{c_N^T x_N + \bar{z} : x_B + A_N x_N = b, x \geq \mathbf{0}\}$.

Simplex

- *Describe the starting point of each iteration.*
  Given a basic solution to start with, we find a corresponding basis and rewrite the LP into canonical form with respect to it.
- *What is the geometric intuition behind the Simplex algorithm?*
  Recall each basic solution corresponds to an extreme point to the feasible region of the LP. The Simplex algorithm moves from one extreme point to another, increasing the objective value until we find the optimal solution or conclude that the LP is unbounded.
- *How do we move from one basic solution to another one?*
  We move to another basic solution by modifying the basis, i.e., constructing a new basis with $B' = B \cup \{k\} \setminus \{l\}$ where $k \notin B$ and $l \in B$.
- *Describe the Simplex algorithm. In particular,*
  - *When do we conclude that the current basic solution is optimal?*
    After rewriting the LP into canonical form, if the objective function is $z = c^T x + m$ with $c \leq \mathbf{0}$ and $m \in \mathbb{R}$, our basic solution is optimal with objective value m.
  - *What happens when we have not reached the optimal solution?*
    Pick $k \notin B$ such that $c_k > 0$, set $x_k = t \geq 0$ and keep other non-basic variables $x_j = 0$. We get $b = Ax = x_k A_k + A_N x_N + A_B x_B = tA_k + 0 + x_B \implies x_B = b - tA_k$.

- *When do we conclude that the LP is unbounded?*
  We claim that if $A_k \leq 0$, the LP is unbounded. First, $x_B = b - tA_k \geq \mathbf{0}$. Next, as $t \to \infty$, $z = \sum_{j \in N} c_j x_j + m \leq c_k x_k + m = tc_k + m \to \infty$ as $c_k > 0$ (otherwise, the algorithm would stop at step 1).
- *What happens when the LP is bounded?*
  We choose $t \geq 0$ as large as possible while keeping all basic variables non-negative. Solving inequalities, we essentially are taking $t := \min\{b_i / A_{ik} : \forall i : A_{ik} \geq 0\}$.
- *How do we update the basis?*
  Let x_l be a basis variable forced to be 0. Then the new basis is $B' = B \cup \{k\} \setminus \{l\}$. Update the objective function to $z = \bar{c}^T x + \bar{z}$ and go back to step 1.

- *What is Bland's rule? Why do we need it?*
  Bland's rule states that, if we have a choice for element entering or leaving the basis, always pick the one with lowest index. If we use Bland's rule, then the Simplex algorithm always terminates.

Initialization

- *Describe the initialization procedure. In particular,*
  - *How do we preprocess the constraint?*
    Multiply the equations such that b becomes non-negative.
  - *How do we augment the LP?*
    Construct an auxiliary LP where $x_{n+1}, \ldots, x_{n+m}$ are auxiliary variables:
    $\min\{x_{n+1} + \cdots + x_{n+m} : (A \mid I)\ x = b, x \geq \mathbf{0}\}$.
  - *When do we conclude that the LP is infeasible?*
    Solve the auxiliary LP using Simplex. If $(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m})^T$ is the optimal solution to the auxiliary problem, then LP is feasible if and only if
    $x_{n+1} = \cdots = x_{n+m} = 0$.
- *State the SEF version of fundamental theorem of linear programming.*
  An LP in SEF is either infeasible, unbounded, or feasible with a basic optimal solution.

Convexity

- *What is the definition of a polyhedron?*
  $P \subseteq \mathbb{R}^n$ is a polyhedron if there exists $A \in M_{m,n}$ and $b \in \mathbb{R}^n$ such that $P = \{x : Ax \leq b\}$. In other words, it is the solution set to a system of linear inequalities.
- *What is the relationship between LPs and polyhedra?*
  The set of feasible solutions (i.e., feasible region) of an LP is a polyhedron.
- *What is the definition of a hyperplane and how is it related to LPs?*
  Let $a \neq \mathbf{0}$ and $\beta \in \mathbb{R}$. Then $\{x : a^T x = \beta\}$ is a hyperplane.
  A hyperplane is the solution set to a single linear equation.
- *What is the definition of a half-space and how is it related to LP?*
  Let $a \neq \mathbf{0}$ and $\beta \in \mathbb{R}$. Then $\{x : a^T x \leq \beta\}$ is a halfspace.

A halfspace is the solution set to a single linear inequality.

- *Let $x, y \in \mathbb{R}^n$. What is the expression of a line through x and y?*
  The line through x and y is defined as $L = \{\lambda x + (1 - \lambda)y : \lambda \in \mathbb{R}\}$.
- *Let $x, y \in \mathbb{R}^n$. What is the expression of a line segment through x and y?*
  The line segment through x and y is defined as $L = \{\lambda x + (1 - \lambda)y : \lambda \in \mathbb{R}, 0 \leq \lambda \leq 1\}$.
- *What is the definition of a convex set?*
  A set $S \subseteq \mathbb{R}^n$ is convex when the line segment between any $x, y \in S$ is contained in S.
- *Is a half-space always convex? How is this related to LP?*
  Yes. Since A, B convex $\implies A \cap B$ convex, P is a polyhedron implies P is convex. Thus, the feasible region of an LP is always a convex set.

Extreme Points

- *What is the definition of "properly-contained" points?*
  A point $x \in \mathbb{R}^n$ is properly contained in the line segment L if $x \in L$ and x is distinct from the endpoints of L.
- *What is the definition of extreme points?*
  A point $x \in \mathbb{R}^n$ is an extreme point when $x \in S$ and there does not exist any line segment $L \subseteq S$ where L properly contains x.
- *How many extreme points can a convex set have? Give examples for each case.*
  It can have zero (halfspace), finitely many (polyhedron), or infinitely many (circle).
- *What does it mean for a constraint to be tight?*
  A constraint is tight for x when it is satisfied with equality.
- *How can we tell whether $\bar{x}$ is an extreme point from the number of tight constraints?*
  Denote the set of tight constraints by $\bar{A}x \leq \bar{b}$. Then $\bar{x}$ is extreme iff $\text{rank}(\bar{A}) = n$.
- *What is the relationship between extreme points and basic feasible solutions?*
  $\bar{x}$ is an extreme point if and only if it is a basic feasible solution.

III The Shortest Path Algorithm

Cardinality Case

- *What is the definition of a* s, t*-cut?*
  A s, t-cut for a subset of vertices $U \subseteq V$ is $\delta(U) = \{uv \in E : u \in U, v \notin U\}$.
- *Name the four statements characterizing an* s, t*-cut.*
  1. If P is a s, t-path and $\delta(U)$ is a s, t-cut, then P contains an edge from $\delta(U)$.
  2. If $S \subseteq E$ contains an edge from every s, t-cut, then S contains a s, t-path.
  3. All s, t-cuts are pairwise disjoint, i.e., $\delta(U_i) \cap \delta(U_j) = \varnothing$ for $i \neq j$.
  4. All s, t-paths must contain an edge from each cut.

General Case

- *What is a width assignment?*
  A width assignment $\{y_U : \delta(U)\ s, t\textbf{-cut}\}$ assigns a integer (called width) to each s, t-cut.
- *When is a width assignment feasible?*
  A width assignment is feasible when the total width of all cuts containing e is no more than c_e for each $e \in E$. That is, $\forall e \in E : \sum(y_U : e \in \delta(U)\ s, t\textbf{-cut}) \leq c_e$.
- *What is the intuition behind width assignments?*
  Intuitively, a width assignment counts the number of cuts of the same type in the corresponding cardinarlity graph.

Weak Duality

- *What is the dual LP of* $\min\{c^T x : Ax \geq b, x \geq \mathbf{0}\}$*?*
  The dual is given by $\max\{b^T y : A^T y \leq c, y \geq \mathbf{0}\}$
- *State the Weak Duality theroem.*
  If x is feasible for (P) and y is feasible for (Q), then $b^T y \leq c^T x$.
- *What is the LP relaxation of an IP?*
  The LP relaxation of an IP is obtained by dropping the integrality restriction.
- *What is the (primal) shortest path LP?*
  The shortest path LP is given by

$$\begin{array}{ll} \max & \sum(x_e c_e : e \in E) \\ s.t. & \sum(x_e : e \in \delta(U)) \geq 1 \quad (U \subseteq V, s \in U, t \notin U) \\ & x \geq \mathbf{0} \end{array}$$

- *What is the dual of the shortest path LP?*
  The dual of the shortest path LP is given by

$$\begin{array}{ll} \max & \sum(y_U : \delta(U)\ s,t\text{-cut}) \\ s.t. & \sum(y_U : e \in \delta(U)) \leq c_e \quad (e \in E) \\ & y \geq \mathbf{0} \end{array}$$

Shortest Path

- *What is an arc?*
  An arc is an ordered pair of vertices.
- *What is a directed path?*
  A directed path is a sequence of arcs.
- *What is the definition of slack of an edge?*
  The slack of an edge $e \in E$ is defined as $\text{slack}_y(e) = c_e - \sum(y_U : e \in \delta(U)\ s,t\text{-cut})$.
- *Describe the shortest path algorithm.*
  Start with $y_U = 0$ for all s,t-cut $\delta(U)$ and set $U = \{s\}$. We want to increase y_U as much as we can while still maintaining feasibility. During each iteration, we find te minimal slack among all vertices reachable from U, set y_U to that value (for the current U), and add the new vertex to U. When the algorithm returns, we get a directed s,t-path along with a feasible width assignment.

Shortest Path Proof

- *What is an equality edge?*
  An equality edge is an edge with zero slack.
- *What is an active cut?*
  A cut $\delta(U)$ is active when $y_U > 0$.
- *State the shortest path characterization theorem.*
  Let y be a feasible dual solution and P an s,t-path. Then P is a shortest path if all edges on P are equalitiy edges, i.e., $c_e = \sum(y_U : e \in \delta(U))$ for all $e \in P$, and every active cut $\delta(U)$ has exactly one edge of P, i.e., $|P \cap \delta(U)| = 1$ for all $\delta(U)$ if $y_U > 0$.

IV Duality Theory

Weak Duality and Strong Duality

- *State the Weak Duality theorem.*
  Let ($\mathrm{P}_{\max}$) and ($\mathrm{P}_{\min}$) be a primal-dual pair. If x and y are feasible for the two LPs, then $c^T x \leq b^T y$. Moreover, if $c^T x = b^T y$, then x is optimal for ($\mathrm{P}_{\max}$) and y is optimal for ($\mathrm{P}_{\min}$).
- *State the Strong Duality theorem.*
  If ($\mathrm{P}_{\max}$) has an optimal solution $\bar{x}$, then ($\mathrm{P}_{\min}$) has an optimal solution $\bar{y}$ such that $c^T \bar{x} = b^T \bar{y}$.

(P)-(D) Conversion

- *Describe the procedure to derive (P_{min}) given (P_{max}).*

  Flip signs of constraints for signs of variables; keep signs of variables the same for the constraints.
- *Describe the procedure to derive (P_{max}) given (P_{min}).*

  Flip signs of variables for signs of constraints; keep signs of constraints the same for the signs of variables.
- *What is the dual LP (D) of (P) $:= \min\{c^T x : Ax = b, Bx \leq d, x \geq \mathbf{0}\}$, given $A \in \mathbb{R}^{m,n}$ and $B \in \mathbb{R}^{p,n}$?*
  There are two classes of primal constraints: the $=$ constraints give free dual variables y_1 and the $\leq$ constraints give dual variables $y_2 \leq \mathbf{0}$. The primal variables are non-negative, which gives $\leq$ in all dual constraints. For (D), the objective function is $(b^T, d^T)^T(y_1, y_2) = b^T y_1 + d^T y_2$ and the constriants are $(A, B)^T(y_1, y_2) = A^T y_1 + B^T y_2$. Hence, we get (D) $:= \max\{b^T y_1 + d^T y_2 : A^T y_1 + B^T y_2 \leq c, y_2 \leq \mathbf{0}\}$.

Possible Outcomes

- *State the 3 consequences of Weak Duality Theorem and Fundamental Theorem of LP.*
  1. If (P) is unbounded, then (D) is infeasible.
  2. If (D) is unbounded, then (P) is infeasible.
  3. If both are feasible, then both have optimal solutions.
- *Draw the table for nine possible outcomes*.

$(D)\backslash(P)$	Optimal	Unbounded	Infeasible
Optimal	Possible	Impossible	Impossible
Unbounded	Impossible	Impossible	Possible
Infeasible	Impossible	Possible	Possible

Complementary Slackness

- *Suppose (P_{max}) and (P_{min}) are a pair of primal and dual LPs with feasible solutions $\bar{x}$ and $\bar{y}$. How do we show the solutions are optimal?*
  The solutions are optimal if the complementary slackness conditions hold: for all variables x_j of ($\mathrm{P}_{\max}$), $\bar{x}_j = 0$ or the jth constriant of ($\mathrm{P}_{\min}$) is satisfied with equality for $\bar{y}$; for all variables y_i of ($\mathrm{P}_{\min}$), $\bar{y}_i = 0$ or the ith constriant of ($\mathrm{P}_{\max}$) is satisfied with equality for $\bar{x}$.

Cone of Tight Constraints

- *What is the definition of a cone of vectors?*
  Given $a^{(1)}, \ldots, a^{(k)} \in \mathbb{R}^n$, the cone generated by these vectors is given by
  $C = \{\lambda_1 a^{(1)} + \cdots + \lambda_k a^{(k)} : \lambda \geq \mathbf{0}\}$.
- *What is the definition of a tight constraint?*
  Let $J(\bar{x})$ be the row indices of A corresponding to the tight constraints of $Ax \leq b$ for $\bar{x}$.
  Then $i \in J(\bar{x}) \iff \mathrm{row}_i(A)\bar{x} = b_i$.
- *What is the definition of cone of tight constraints?*
  The cone of tight constraints for $\bar{x}$ is the cone C generated by the rows of A correspondings to the tight constraints, i.e., $C = \{\sum_{i \in J(\bar{x})} \lambda_i \mathrm{row}_i(A) : \lambda_i \geq \mathbf{0}\}$.
- *How do we use the cone of tight constraints to prove $\bar{x}$ is optimal?*
  The feasible solution $\bar{x}$ is optimal iff c is in the cone of tight constraints of x.

V Integer Programs

Convex Hull

- *What is the definition of convex hull?*
  Let $C \subseteq \mathbb{R}^n$. The convex hull of C is the smallest convex set that contains C.
- *State Meyer's theorem.*
  For $P = \{x : Ax \leq b\}$ where A, b are rational, the convex hull of all integer points in P is a polyhedron.
- *For IP* $\max\{c^T x : Ax \leq b, x \in \mathbb{Z}\}$ *and LP* $\max\{c^T x : A'x \leq b'\}$, *give the four propositions.*
  1. IP is infeasible iff LP is infeasible.
  2. IP is unbounded iff LP is unbounded.
  3. An optimal solution to IP is an optimal solution to LP.
  4. An extreme optimal solution to LP is an optimal solution to IP.

Cutting Plane

- *Suppose Simplex returns a non-integral optimal solution $\bar{x}$ for the LP relaxation of an IP. What is the definition of cutting plane for $\bar{x}$?*
  A cutting plane for $\bar{x}$ is a constraint $a^T x \leq \beta$ which is satisfied for all feasible solutions to the IP but not satisfied for $\bar{x}$.
- *Describe the cutting plane scheme.*
  Given IP $\max\{c^T x : Ax \leq b, x \in \mathbb{Z}\}$, let P denote its LP relaxation. If P is infeasible, stop as IP is also infeasible. Otherwise, let $\bar{x}$ be the optimal solution to P returned by Simplex. If $\bar{x}$ is integral, stop as $\bar{x}$ is also optimal for IP. Otherwise, find a cutting plane $a^T x \leq \beta$ for $\bar{x}$; add the constraint $a^T x \leq \beta$ to the system $Ax \leq b$ and repeat.
- *Describe how to find a cutting plane in general.*
  Suppose $\bar{x}$ is not an integral solution. Then b_i is fractional for some value i. Every feasible solution to the LP relaxation satisfies

  $$x_{r(i)} + \sum_{j \in N} A_{ij} x_j = b_i \implies x_{r(i)} + \sum_{j \in N} \lfloor A_{ij} \rfloor x_j \leq b_i \implies x_{r(i)} + \sum_{j \in N} \lfloor A_{ij} \rfloor x_j \leq \lfloor b_i \rfloor.$$

  but $\bar{x}$ does not satisfy the constraint above as

  $$\underbrace{x_{r(i)}}_{b_i} + \sum_{j \in N} \lfloor A_{ij} \rfloor \underbrace{x_j}_{=0} = b_i > \lfloor b_i \rfloor.$$

VI Nonlinear Programs

NLP

- *What is the general form of a NLP?*
  $\min\{c^T x : g_i(x) \leq 0 \ (i = 1, \ldots, k)\}$.
- *How do we generalize binary IPs with NLPs?*
  $x_j \in \{0,1\} \iff x_j(1 - x_j) = 0$.
- *How do we generalize general IPs with NLPs?*
  $x_j \in \mathbb{Z} \iff \sin(\pi x) = 0$.

Convex Sets and Convex Functions

- *What is a local optimum for $S \subseteq \mathbb{R}^n$?*
  $x \in S$ is a local optimum if $\forall x' \in S : \|x' - x\| \leq \delta \implies f(x) \leq f(x')$.
- *What's special about a local optimum in a convex set?*
  If S is convex and x is a local optimum, then x is optimal.
- *What is the definition of a convex function $f : \mathbb{R}^n \to \mathbb{R}$?*
  f is convex if for all $a, b \in \mathbb{R}^n$, $f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda) f(b)$ for all $\lambda \in [0, 1]$.
- *What is special about the set $S = \{x \in \mathbb{R}^n : g(x) \leq \beta\}$ given g is convex?*
  The set $S = \{x \in \mathbb{R}^n : g(x) \leq \beta\}$ is a convex set.
- *What is special about the feasible region of (P) given all constraint functions are convex?*
  If all constraint functions are convex, then the feasible region is convex.

Epigraph

- *What is the definition of the epigraph of a function $f : \mathbb{R}^n \to \mathbb{R}$?*
  The epigraph of a function contains all the points above the graph of the function, i.e., $epi(f) = \{(y, x)^T : y \geq f(x), x \in \mathbb{R}^n\} \subseteq \mathbb{R}^{n+1}$.
- *What is the relationship between the function and its epigraph?*
  f is convex if and only if $epi(f)$ is a convex set.

Subgradient

- *What is a subgradient of $f : \mathbb{R}^n \to \mathbb{R}$ at $\bar{x} \in \mathbb{R}^n$?*
  $s \in \mathbb{R}^n$ is a subgradient of f at $\bar{x}$ if $h(x) := f(\bar{x}) + s^T(x - \bar{x}) \leq f(x)$ for all $x \in \mathbb{R}^n$.
- *What are the three key points in the definition?*
  h is affine; $h(\bar{x}) = f(\bar{x})$; h is a lower bound for f.
- *What is the intuition behind a subgradient?*
  The subgradient generalizes the derivative to convex functions which are not necessarily differentiable. If the gradient $\nabla f(\bar{x})$ of f exists at $\bar{x}$, then it is a subgradient. Otherwise, the subgradient might not be unique.

Supporting Halfspace

- *Let $C \in \mathbb{R}^n$ be convex and $\bar{x} \in C$. When is $F = \{x : s^T x \leq \beta\}$ a supporting halfspace for C at $\bar{x}$?*
  Two conditions: (1) $C \subseteq F$ and (2) $s^T \bar{x} = \beta$, i.e., $\bar{x}$ is on the boundary of F.
- *State the proposition that proves F is a supporting halfspace of C at $\bar{x}$.*
  1. Let $g : \mathbb{R}^n \to \mathbb{R}$ be convex and let $\bar{x}$ where $g(\bar{x}) = 0$.
  2. Let s be a subgradient of g at $\bar{x}$.
  3. Let $C = \{x : g(x) \leq 0\}$.
  4. Let $F = \{x : h(x) := g(\bar{x}) + s^T(x - \bar{x}) \leq 0\}$.

  Then F is a supporting halfspace of C at $\bar{x}$.

LP Relaxation

- *What is the intuition behind finding optimal solution for NLP using its LP relaxation?*
  We replace the non-linear constraint $g(x) \leq 0$ with the linear constraint $h(x) := g(x) + s^T(x - \bar{x}) \leq 0$.
- *State the proposition that proves $\bar{x}$ is optimal for NLP* $\min\{c^T x : g_i(x) \leq 0, i = 1, \ldots k\}$.
  Suppose all constraints are convex. Let I denote the tight constraints for $\bar{x}$, i.e., $g_i(\bar{x}) = 0$ for $i \in I$. Let $s^{(i)}$ be a subgradient for g_i at $\bar{x}$ for all $i \in I$. If $-c \in \text{cone}\{s^{(i)} : i \in I\}$, then $\bar{x}$ is optimal.

KKT

- *What is a Slater point?*
  A feasible solution $\bar{x}$ is a slater point if every inequality $g_i(\bar{x}) < 0$ is satisfied strictly by $\bar{x}$.
- *State the KKT theorem.*
  Given NLP $\min\{c^T x : g_i(x) \leq 0, i \in J\}$, if
  1. g_i is convex for all $i \in J$,
  2. There exists a Slater point,
  3. $\bar{x}$ is a feasible solution,
  4. I is the set of indices i for which $g_i(\bar{x}) = 0$, and
  5. For all $i \in I$ there exists a gradient $\nabla g_i(\bar{x})$ of g_i at $\bar{x}$.

  Then $\bar{x}$ is optimal iff $-c \in \text{cone}\{\nabla g_i(\bar{x}) : i \in I\}$.

Appendix A Checklist

You must know how to do the following:

1. Find **certificates** to show a LP is infeasible (Farka's) or unbounded (find $\bar{x}$ and r).
2. **Rewrite** a LP into standard equality form.
3. **Rewrite** a LP into canonical form.
4. Given a basis or a basic feasible solution, perform **Simplex** to solve the LP.
5. Find a basic feasible solution for a LP by solving the **augmented** LP.
6. **Formulate** the shortest path IP and its dual given a graph.
7. Use the **shortest path algorithm** to find a shortest path along with the width assignment given a graph.
8. **Derive** the dual LP given a primal LP.
9. Use **complementary slackness** to prove that $\bar{x}$ and $\bar{y}$ are optimal for a primal-dual pair.
10. Use **cone of tight constraints** to prove that $\bar{x}$ is optimal.
11. Find a **cutting plane** given a non-integral optimal solution for the LP relaxation of an IP.
12. Use the **cutting plane scheme** to find an optimal solution for an IP.
13. Use the definition to prove claims regarding **convex sets**, **convex functions**, **subgradients**, and **supporting halfspaces**.
14. Use **subgradient theorem** (one direction of KKT) to prove $\bar{x}$ is optimal for a NLP.
15. Use **KKT** theorem to prove $\bar{x}$ is optimal for a NLP.