# Advanced Operations Research Techniques IE316

## Quiz 1 Review

Dr. Ted Ralphs

# Reading for The Quiz

- Material covered in detail in lecture.
  - 1.1, 1.4, 2.1-2.6, 3.1-3.3, 3.5
- Background material you should know.
  - 1.2, 1.3, 1.5
- Material you should read and be familiar with.
  - 1.6, 2.7, 3.7
- Optional material that may help your comprehension.
  - 2.8, 3.6

# How to Study for The Quiz

- You should have done all the reading and understand the theorems and proofs in the sections we covered in detail.
- Review and understand the homework solutions.
- Review the lecture slides and make sure you can answer the questions that are posed.
- Go through the review slides and make sure you understand the flow of ideas and the connections among them.
- Ask questions if something is not clear.

# Math Programming Models

- A mathematical model consists of:
  - Decision variables
  - Constraints
  - Objective Function
  - Parameters and Data

The general form of a *math programming model* is:

$$\begin{array}{ll} min \text{ or } max & f(x_1, \ldots, x_n) \\ s.t. & g_i(x_1, \ldots, x_n) \left\{ \begin{array}{c} \leq \\ = \\ \geq \end{array} \right\} b_i \end{array}$$

We might also want the values of the variables to belong to discrete set $X$.

# Linear Functions

- A *linear function* $f : \mathbf{R}^n \to \mathbf{R}$ is a weighted sum, written as

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} c_i x_i$$

  for given coefficients $c_1, \ldots, c_n$.

- We can write $x_1, \ldots, x_n$ and $c_1, \ldots, c_n$ as vectors $x, c \in \mathbf{R}^n$ to obtain:

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

- In this way, a linear function can be represented simply as a vector.

# Linear Programs

- A linear program is a math program that can be written in the following form:

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{s.t.} \quad & a_i^T x \geq b_i \; \forall i \in M_1 \\
& a_i^T x \leq b_i \; \forall i \in M_2 \\
& a_i^T x = b_i \; \forall i \in M_3 \\
& x_j \geq 0 \;\; \forall j \in N_1 \\
& x_j \leq 0 \;\; \forall j \in N_2
\end{aligned}
$$

- This in turn can be written equivalently as

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{s.t.} \quad & Ax \geq b
\end{aligned}
$$

# Standard Form

- To solve a linear program, we must put it in the following *standard form*:

$$
\begin{aligned}
& min\, c^T x \\
& s.t. \quad Ax = b \\
& \qquad\quad x \geq 0
\end{aligned}
$$

- Eliminate free variables: Replace $x_j$ with $x_j^+ - x_j^-$, where $x_j^+, x_j^- \geq 0$.
- Eliminate inequality constraints: Given an inequality $a_i^T x \geq b_i$, we can rewrite it as:

$$
\begin{aligned}
a_i^T x - s_i &= b_i \\
s_i &\geq 0
\end{aligned}
$$

# A Little Linear Algebra Review

**Definition 1.** *A finite collection of vectors* $x_1, \ldots, x_k \in \mathbf{R}^n$ *is linearly dependent if* $\exists \lambda \in \mathbf{R}^k$ *such that* $\sum_{i=1}^k \lambda_i x^i = 0$. *Otherwise, the vectors are linearly independent.*

Let $A$ be a square matrix. Then, the following statements are equivalent:

- The matrix $A$ is invertible.
- The matrix $A^T$ is invertible.
- The determinant of $A$ is nonzero.
- The rows of $A$ are linearly independent.
- The columns of $A$ are linearly independent.
- For every vector $b$, the system $Ax = b$ has a unique solution.
- There exists some vector $b$ for which the system $Ax = b$ has a unique solution.

# A Little More Linear Algebra Review

**Definition 2.** *A nonempty subset* $S \subseteq \mathbf{R}^n$ *is called a* subspace *if* $\alpha x + \gamma y \in S \ \forall x, y \in S$ *and* $\forall \alpha, \gamma \in \mathbf{R}$.

**Definition 3.** *A* linear combination *of a collection of vectors* $x^1, \ldots x^k \in \mathbf{R}^n$ *is any vector* $y \in \mathbf{R}^n$ *such that* $y = \sum_{i=1}^k \lambda_i x^i$ *for some* $\lambda \in \mathbf{R}^k$.

**Definition 4.** *The* span *of a collection of vectors* $x^1, \ldots x^k \in \mathbf{R}^n$ *is the set of all linear combinations of those vectors.*

**Definition 5.** *Given a subspace* $S \subseteq \mathbf{R}^n$, *a collection of linearly independent vectors whose span is* $S$ *is called a* basis *of* $S$. *The number of vectors in the basis is the* dimension *of the subspace.*

# Subspaces and Bases

- A given subspace has an infinite number of bases.
- Each basis has the same number of vectors in it.
- If $S$ and $T$ are subspaces such that $S \subset T \subset \mathbf{R}^n$, then a basis of $S$ can be extended to a basis of $T$.
- The span of the columns of a matrix $A$ is a subspace called the *column space* or the *range*, denoted $range(A)$.
- The span of the rows of a matrix $A$ is a subspace called the *row space*.
- The dimensions of the column space and row space are always equal. We call this number $rank(A)$.
- Clearly, $rank(A) \leq \min\{m, n\}$. If $rank(A) = \min\{m, n\}$, then $A$ is said have *full rank*.
- The set $\{x \in \mathbf{R}^n | Ax = 0\}$ is called the *null space* of $A$ (denoted $null(A)$) and has dimension $n - rank(A)$.

# Polyhedra

**Definition 6.** *A* polyhedron *is a set of the form* $\{x \in \mathbf{R}^n | Ax \geq b\}$, *where* $A \in \mathbf{R}^{m \times n}$ *and* $b \in \mathbf{R}^m$.

**Definition 7.** *A set* $S \subset \mathbf{R}^n$ *is* bounded *if there exists a constant* $K$ *such that* $|x_i| < K \ \forall x \in S, \forall i \in [1, n]$.

**Definition 8.** *Let* $a \in \mathbf{R}^n$ *and* $b \in \mathbf{R}$ *be given.*

- *The set* $\{x \in \mathbf{R}^n | a^T x = b\}$ *is called a* hyperplane.
- *The set* $\{x \in \mathbf{R}^n | a^T x \geq b\}$ *is called a* half-space.

Notes:

- A given hyperplane forms the *boundary* of a corresponding half-space.
- A vector $a \in \mathbf{R}^n$ is *orthogonal* to the hyperplane defined by $a$.
- A polyhedron is the intersection of a finite number of half-spaces.

# Convex Sets

**Definition 9.** *A set* $S \subseteq \mathbf{R}^n$ *is* convex *if* $\forall x, y \in S, \lambda \in [0,1]$, *we have* $\lambda x + (1-\lambda)y \in S$.

**Definition 10.** *Let* $x^1, \ldots, x^k \in \mathbf{R}^n$ *and* $\lambda \in \mathbf{R}^k$ *be given such that* $\lambda^T \mathbf{1} = 1$.

- *The vector* $\sum_{i=1}^k \lambda_i x^i$ *is said to be a* convex combination *of* $x^1, \ldots, x^k$.
- *The* convex hull *of* $x_1, \ldots, x_k$ *is the set of all convex combinations of these vectors.*

Notes:

- The convex hull of two points is a line segment.
- A set is convex if and only if for any two points in the set, the line segment joining those two points lies entirely in the set.

# Extreme Points and Vertices

Let $\mathcal{P} \subseteq \mathbf{R}^n$ be a given polyhedron.

**Definition 11.** *A vector $x \in \mathcal{P}$ is an extreme point of $\mathcal{P}$ if $\nexists y, z \in \mathcal{P}, \lambda \in (0, 1)$ such that $x = \lambda y + (1 - \lambda)z$.*

**Definition 12.** *A vector $x \in \mathcal{P}$ is an vertex of $\mathcal{P}$ if $\exists c \in \mathbf{R}^n$ such that $c^T x < c^T y \; \forall y \in \mathcal{P}, x \neq y$.*

Notes:

- These definitions are purely geometric and easy to work with.
- However, to do anything algorithmic, we need an algebraic characterization.
- Thus, we will consider the matrix defining the polyhedron (not unique!).

# Binding Constraints

Consider a polyhedron $\mathcal{P} = \{x \in \mathbf{R}^n | Ax \geq b\}$.

**Definition 13.** *If a vector $\hat{x}$ satisfies $a_i^T \hat{x} = b_i$, then we say the corresponding constraint is binding.*

**Theorem 1.** *Let $\hat{x} \in \mathbf{R}^n$ be given and let $I = \{i \mid a_i^T \hat{x} = b_i\}$ represent the set of constraints that are binding at $\hat{x}$. Then the following are equivalent:*

- *There exist $n$ vectors in the set $\{a_i \mid i \in I\}$ that are linearly independent.*
- *The span of the vectors $\{a_i \mid i \in I\}$ is $\mathbf{R}^n$.*
- *The system of equations $a_i^T \hat{x} = b_i, i \in I$ has the unique solution $\hat{x}$.*

If the vectors $\{a_j \mid j \in J\}$ for some $J \subseteq [1, m]$ is linearly independent, we will say that the corresponding constraints are also linearly independent.

# Basic Solutions

Consider a polyhedron $\mathcal{P} = \{x \in \mathbf{R}^n | Ax \geq b\}$ and let $\hat{x} \in \mathbf{R}^n$ be given.

**Definition 14.** *The vector $\hat{x}$ is a* basic solution *if there exist $n$ linearly independent, binding constraints at $\hat{x}$.*

**Definition 15.** *If $\hat{x}$ is a basic solution and $\hat{x} \in \mathcal{P}$, then $\hat{x}$ is* basic feasible solution*.*

**Theorem 2.** *If $\mathcal{P}$ is nonempty and $\hat{x} \in \mathcal{P}$, then the following are equivalent:*

- *$\hat{x}$ is a vertex.*
- *$\hat{x}$ is an extreme point.*
- *$\hat{x}$ is a basic feasible solution.*

# Basic Feasible Solutions in Standard Form

- To obtain a basic solution, we must set $n - m$ of the variables to zero.
- We must also obtain a set of linearly independent constraints.
- Therefore, the variables we pick cannot be arbitrary.

**Theorem 3.** *A vector $\hat{x} \in \mathbf{R}^n$ is a basic solution if and only if $A\hat{x} = b$ and there exist indices $B(1), \ldots, B(m)$ such that:*

- *The columns $A_{B(1)}, \ldots, A_{B(m)}$ are linearly independent, and*
- *If $i \neq B(1), \ldots, B(m)$, then $\hat{x}_i = 0$.*

# Some Terminology

- If $\hat{x}$ is a basic solution, then $\hat{x}_{B(1)}, \ldots, \hat{x}_{B(m)}$ are the *basic variables*.
- The columns $A_{B(1)}, \ldots, A_{B(m)}$ are called the *basic columns*.
- Since they are linearly independent, these columns form a *basis* for $\mathbf{R}^m$.
- A set of basic columns form a *basis matrix*, denoted $B$. So we have,

$$B = \left[A_{B(1)}\ A_{B(2)} \cdots A_{B(m)}\right], \quad x_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}$$

# The Full Row Rank Assumption

**Theorem 4.** *Let* $\mathcal{P} = \{x \in \mathbf{R}^n | Ax \geq b, x \geq 0\}$ *for some* $A \in \mathbf{R}^{m \times n}$ *with* $rank(A) = k$. *If rows* $a_{i_1}^T, a_{i_2}^T, \ldots, a_{i_k}^T$ *are linearly independent, then*

$$\mathcal{P} = \{x \in \mathbf{R}^n \mid a_{i_1}^T x = b_{i_1}, a_{i_2}^T x = b_{i_2}, \ldots, a_{i_k}^T x = b_{i_k}, x \geq 0\}.$$

Notes:

- This tells us we can eliminate any row we like from $A$, as long as we leave $k$ *linearly independent rows*.
- This allows us to always assume $A$ has full row rank.

# Degeneracy

**Definition 16.** *A basic solution* $\hat{x}$ *is called* degenerate *if more than* $n$ *of the constraints are binding at* $\hat{x}$*.*

Notes:

- In a standard form polyhedron, the vector $\hat{x}$ is degenerate if more than $n - m$ of its components are zero.
- Note that degeneracy is *not* independent of representation.
- Degeneracy is important because it can cause problems in some algorithms for linear programming.
- This also gives us a glimpse of the importance of formulation in linear programming.

# Optimality in Linear Programming

- For linear optimization, a finite optimal cost is equivalent to the existence of an optimal solution.
- Since any linear programming problem can be written in standard form, we can derive the following result:

**Theorem 5.** *Consider the linear programming problem of minimizing* $c^T x$ *over a nonempty polyhedron. Then, either the optimal cost is* $-\infty$ *or there exists an optimal solution which is an extreme point.*

# Representation of Polyhedra

**Theorem 6.** *A nonempty, bounded polyhedron is the convex hull of its extreme points.*

**Theorem 7.** *The convex hull of a finite set of vectors is a polyhedron.*

Notes:

- Graphically, these results may seem obvious, but it takes some work to prove them.
- These results are fundamental to the theory of linear optimization.

# Feasible and Improving Directions

**Definition 17.** *Let $\hat{x}$ be an element of a polyhedron $\mathcal{P}$. A vector $d \in \mathbf{R}^n$ is said to be a* feasible direction *if there exists $\theta \in \mathbf{R}_+$ such that $\hat{x} + \theta d \in \mathcal{P}$.*

**Definition 18.** *Consider a polyhedron $\mathcal{P}$ and the associated linear program $\min_{x \in \mathcal{P}} c^T x$ for $c \in \mathbf{R}^n$. A vector $d \in \mathbf{R}^n$ is said to be an* improving direction *if $c^T d < 0$.*

Notes:

- Once we find a feasible, improving direction, we want to move along that direction as far as possible.
- Recall that we are interested in extreme points.
- The simplex method moves between adjacent extreme points using improving directions.

# Constructing Feasible Search Directions

- Consider a BFS $\hat{x}$, so that $\hat{x}_N = 0$.
- Any feasible direction must increase the value of at least one of the nonbasic variables.
- We will consider the *basic directions* that increase the value of exactly one of the nonbasic variables, say variable $j$. This means

$$\begin{aligned} d_j &= 1 \\ d_i &= 0 \text{ for every nonbasic index } i \neq j \end{aligned}$$

- In order to remain feasible, we must also have $Ad = 0$, which means

$$0 = Ad = \sum_{i=1}^{n} A_i d_i = \sum_{i=1}^{m} A_{B(i)} d_{B(i)} + A_j = B d_B + A_j \Rightarrow d_B = -B^{-1} A_j$$

# Constructing Improving Search Directions

- Now we know how to construct feasible search directions.
- In order to ensure they are improving, we must have $c^T d < 0$.

**Definition 19.** *Let $\hat{x}$ be a basic solution, let $B$ be an associated basis matrix, and let $c_B$ be the vector of costs of the basic variables. For each $j$, we define the* reduced cost *$\bar{c}_j$ of variable $j$ by*

$$\bar{c}_j = c_j - c_B^T B^{-1} A_j.$$

- The reduced cost is the amount of change in the objective function per unit increase in the corresponding variable.
- The basic direction associated with variable $j$ is improving if and only if $\bar{c}_j < 0$.

# Optimality Conditions

**Theorem 8.** *Consider a basic feasible solution $\hat{x}$ associated with a basis matrix $B$ and let $\bar{c}$ be the corresponding vector of reduced costs.*

- *If $\bar{c} \geq 0$, then $\hat{x}$ is optimal.*
- *If $\hat{x}$ is optimal and nondegenerate, then $\bar{c} \geq 0$.*

Notes:

- The condition $\bar{c} \geq 0$ implies there are no feasible improving directions.
- However, $\bar{c}_j < 0$ does not ensure the existence of an improving, feasible direction unless the current BFS is nondegenerate

.

# The Step Length

- The distance we can move without becoming infeasible is the *step length*.
- We can move until one of the nonnegativity constraints is violated.
- If $d \geq 0$, then the step length is $\infty$ and the linear program is unbounded.
- if $d_i < 0$, then $\hat{x}_i + \theta d_i \geq 0 \Rightarrow \theta \geq -\frac{\hat{x}_i}{d_i}$.
- Therefore, we can compute the step length explicitly as

$$\theta^* = \min_{\{i | d_i < 0\}} -\frac{\hat{x}_i}{d_i}$$

- Note that we need only consider the basic variables in this computation.

# The Tableau Method

- This is the standard method for solving LPs by hand.
- We update the matrix $B^{-1}[b|A]$ as we go.
- To do this, we use elementary row operations.
- In addition, we also keep track of the reduced costs in row “zero”.
- This gives us all the data we need at each iteration.

# What the Tableau Looks Like

- The tableau looks like this

| | |
|---|---|
| $-c_B^T B^{-1} b$ | $c^T - c_B^T B^{-1} A$ |
| $B^{-1} b$ | $B^{-1} A$ |

- In more detail, this is

| | | | |
|---|---|---|---|
| $-c_B^T x_B$ | $\bar{c}_1$ | $\cdots$ | $\bar{c}_n$ |
| $x_{B(1)}$<br>$\vdots$<br>$x_{B(m)}$ | $B^{-1} A_1$ | $\cdots$ | $B^{-1} A_n$ |

# Implementing the Tableau Method

1. Start with the tableau associated with a specified BFS and associated basis $B$.
2. Examine the reduced costs in row zero and select a *pivot column* with $\bar{c}_j < 0$ if there is one. Otherwise, the current BFS is optimal.
3. Consider $u = b^{-1}A_j$, the $j\text{th}$ column of the tableau. If no component of $u$ is positive, then the LP is unbounded.
4. Otherwise, compute the step size using the minimum ratio rule and determine the *pivot row*.
5. Scale the pivot row so that the *pivot element* becomes zero.
6. Add a constant multiple of the pivot row to each other row of the tableau so that all other elements of the pivot column become zero.
7. Go to Step 2.

# The Revised Simplex Method

The revised simplex method is used in practice and is almost identical to the tableau method except that we only update $B^{-1}$.

1. Start with a specified BFS $\hat{x}$ and the associated basis inverse $B^{-1}$.
2. Compute $p^T = c_B^T B^{-1}$ and the reduced costs $\bar{c}_j = c_j - p^T A_j$.
3. If $\bar{c} \geq 0$, then the current solution is optimal.
4. Select the entering variable $j$ and compute $u = B^{-1} A_j$.
5. If $u \leq 0$, then the LP is unbounded.
6. Determine the step size $\theta^* = \min_{\{i | u_i > 0\}} \frac{\hat{x}_i}{u_i}$.
7. Determine the new solution and the leaving variable $i$.
8. Update $B^{-1}$.
9. Go to Step 1.

# Pivot Selection

- The process of removing one variable and replacing from the basis and replacing it with another is called *pivoting*.
- We have the freedom to choose the leaving variable from among a list of candidates.
- How do we make this choice?
- The reduced cost tells us the cost in the objective function for each unit of change in the given variable.
- Intuitively, $c_j$ is the cost for the change in the variable itself and $-c_B^T B^{-1} A_j$ is the cost of the compensating change in the other variables.
- This leads to the following possible rules:
  - Choose the column with the most negative reduced cost.
  - Choose the column for which $\theta^* |\bar{c}_j|$ is largest.

# Obtaining an Initial Basic Feasible Solution

- If the origin is feasible, then finding an initial BFS is easy.
- Suppose we are given an LP $\min\{c^T X | Ax = b, x \geq 0\}$ already in standard form where the origin is not feasible.
- To obtain an initial BFS, solve the following auxiliary LP.

$$
\begin{aligned}
min \quad & \sum_{i=1}^{m} y_i \\
s.t. \quad & Ax + y = b \\
& x \geq 0 \\
& y \geq 0
\end{aligned}
$$

- If the optimal value for this problem is zero, then we obtain a feasible solution.
- Otherwise, the original problem was infeasible.
- This is usually called the *Phase I LP*.

# Two-phase Simplex and the Big M method

- In the above method, we first solve the Phase I LP to obtain a BFS.
- Using that BFS, we start Phase II, which is solving the original problem.
- Another approach is to combine Phase I and Phase II.
- In this approach, we add the artificial variables and then change the objective function to

$$\sum_{j=1}^{n} c_j x_j + M \sum_{i=1}^{m} y_i$$

- $M$ has to be large enough to force the artificial variables to zero.
- If the optimal solution has any artificial variables at nonzero level, then the original problem was infeasible.
- Otherwise, we obtain an optimal solution.
- In practice, two-phase simplex is usually used.

# Computational Efficiency of the Simplex Method

- The efficiency of the method depends on the number of iterations.
- The number of iterations depends on how many extreme points are visited.
- It is easy to construct an example where there are $2^n$ extreme points and all of them are visited.
- This means in the worst case, the simplex method requires an exponential number of iterations.
- No matter how well the method performs in practice, we are limited by the diameter of the polyhedron.
- Not much is known about the diameter of complex polyhedra.
- All things considered, the simplex method performs very well in practice.