Introduction to Mathematical Programming
IE496

Quiz 2 Review

Dr. Ted Ralphs

Reading for The Quiz

- Material covered in detail in lecture
  - Bertsimas 4.1-4.5, 4.8, 5.1-5.5, 6.1-6.3
- Material covered briefly in lecture
  - Bertsimas 4.6, 4.9

Deriving the Dual Problem

- Consider a standard form LP $\min\{c^T x : Ax = b, x \geq 0\}$.
- To derive the *dual problem*, we use Lagrangian relaxation and consider the function

$$g(p) = \min_{x \geq 0} \left[c^T x + p^T (b - Ax) \right]$$

  in which infeasibility is penalized by a vector of *dual prices* p.
- For every vector p, $g(p)$ is a lower bound on the optimal value of the original LP.
- To achieve the best bound, we considered maximizing $g(p)$, which is equivalent to

$$\begin{array}{ll} maximize & p^T b \\ s.t. & p^T A \leq c \end{array}$$

- This LP is the dual to the original one.

From the Primal to the Dual

We can dualize general LPs as follows

PRIMAL	minimize	maximize	DUAL
constraints	$\geq b_i$ $\leq b_i$ $= b_i$	≥ 0 ≤ 0 free	variables
variables	≥ 0 ≤ 0 free	$\leq c_j$ $\geq c_j$ $= c_j$	constraints

Relationship of the Primal and the Dual

The following are the possible relationships between the primal and the dual:

	Finite Optimum	**Unbounded**	**Infeasible**
Finite Optimum	Possible	Impossible	Impossible
Unbounded	Impossible	Impossible	Possible
Infeasible	Impossible	Possible	Possible

Strong Duality and Complementary Slackness

Theorem 1. *(Strong Duality) If a linear programming problem has an optimal solution, so does its dual, and the respective optimal costs are equal.*

Theorem 2. *If x and p are feasible primal and dual solutions, then x and p are optimal if and only if*

$$\begin{aligned} p^T(Ax - b) &= 0, \\ (c^T - p^T A)x &= 0. \end{aligned}$$

- From complementary slackness, we can derive a number of alternative *optimality conditions*.
- The simplex algorithm always maintains complementary slackness

LPs with General Upper and Lower Bounds

- In many problems, the variables have explicit nonzero upper or lower bounds.
- These upper and lower bounds can be dealt with *implicitly* instead of being included as constraints.
- In this more general framework, all nonbasic variables are fixed at either their upper or lower bounds.
- For minimization, variables eligible to enter the basis are either
  - Variables at their lower bounds with negative reduced costs, or
  - Variables at their upper bounds with positive reduced cost.
- When no such variables exist, we are at optimality.
- For maximization, we can just reverse the signs.

Economic Interpretation

- The dual variables tell us the marginal change in the objective function per unit change in the right-hand side of a constraint.
- Hence, we can interpret the dual variables as prices associated with the resources defined by each constraint.
- The dual constraints state that the total price of the constituent resources in each column should be no more than the given objective function value.
- This is the same as having the reduced costs nonnegative.
- Complementary slackness says that, in fact, for all variables with a positive value (basic variables), these two costs must be equal.
- The dual program tries to maximize the price of all given resources subject to the dual constraints.

The Dual Simplex Method

- This leads to a dual version of the simplex method in tableau form.
- Recall the simplex tableau

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

- In the dual simplex method, the basic variables are allowed to take on negative values, but we keep the reduced costs nonnegative.
- The pivot row is any row with a negative right-hand side.
- We choose the pivot column using a ratio test that ensures the reduced costs remain nonnegative (dual feasibility).

Optimal Bases

- Note that a given basis determines both a unique solution to the primal and a unique solution to the dual.

$$
\begin{aligned}
x_B &= B^{-1}b \\
p^T &= c_B^T B^{-1}
\end{aligned}
$$

- Both the primal and dual solutions are basic and either one, or both, may be feasible.
- If they are both feasible, then they are both optimal and the basis is also optimal.
- Both versions of the simplex method go from one adjacent basic solution to another until reaching optimality.
- Both versions either terminate in a finite number of steps or cycle.

Geometric Interpretation of Optimality

- Suppose we have a problem in inequality form, so that the dual is in standard form, and a basis B.
- If I is the index set of binding constraints at the corresponding (nondegenerate) BFS, and we enforce complementary slackness, then dual feasibility is equivalent to

$$\sum_{i \in I} p_i a_i = c.$$

- In other words, the objective function must be a nonnegative combination of the binding constraints.
- We can easily picture this graphically.

Farkas' Lemma

Proposition 1. *Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ be given. Then exactly one of the following holds:*

1. $\exists x \geq 0$ *such that* $Ax = b$.
2. $\exists p$ *such that* $p^T A \geq 0^T$ *and* $p^T b < 0$.

- This is closely related to the geometric interpretation of optimality.
- In words, this says that either
  - The vector b is in the cone C formed by all nonnegative linear combination of the columns of A, or
  - There is a hyperplane separating b from C.
  - From this lemma, we can derive much of optimization theory.

Recession Cones and Extreme Rays

- The *recession cone* associated with a polyhedron in inequality form is $\{d \in \mathbb{R}^n | Ad \geq 0\}$.
- The *extreme rays* of this cone are also called the extreme rays of the polyhedron.
- Note that a polyhedron has a finite number of "non-equivalent" extreme rays.

Theorem 3. *Consider the LP $\min\{c^T x | Ax \geq b\}$ and assume the feasible region has at least one extreme point. The optimal cost is equal to $-\infty$ if and only if some extreme ray d satisfies $c^T d < 0$.*

- When simplex terminates due to unboundedness, there is a column j with negative reduced cost for which basic direction j belongs to the recession cone.
- It is easy to show that this basic direction is an extreme ray of the recession cone.

Representation of Polyhedra

Theorem 4. *Let* $\mathcal{P} = \{x \in \mathbb{R}^n\}$ *be a nonempty polyhedron with at least one extreme point. Let* $x^1, \ldots, x^k$ *be the extreme points and* $w^1, \ldots, w^r$ *be the extreme rays. Then*

$$P = \left\{ \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j \mid \lambda_i \geq 0, \theta_j \geq 0, \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

Corollary 1. *A nonempty polyhedron is bounded if and only if it has no extreme rays.*

Corollary 2. *A nonempty bounded polyhedron, is the convex hull of its extreme points.*

Corollary 3. *Every element of a polyhedral cone can be expressed as a nonnegative linear combination of extreme rays.*

The Converse of the Representation Theorem

Definition 1. *A set Q is finitely generated if it is of the form*

$$P = \left\{ \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j \mid \lambda_i \geq 0, \theta_j \geq 0, \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

for given vectors $x^1, \ldots, x^k$ and $w^1, \ldots, w^r$ in $\mathbb{R}^n$.

Theorem 5. *Every finitely generated set is a polyhedron. The convex hull of finitely many vectors is a bounded polyhedron, also called a polytope.*

The Fundamental Idea of Sensitivity Analysis

- Using the simplex algorithm to solve a standard form problem, we know that if B is an optimal basis, then two conditions are satisfied:
  - $B^{-1}b \geq 0$
  - $c^T - c_B^T B^{-1} A \geq 0$
- When the problem is changed, we can check to see how these conditions are affected.
- When using the simplex method, we always have B^{-1} available, so we can easily recompute the appropriate quantities.
- If the change causes the optimality conditions to be violated, we can usually re-solve from the current basis using either primal or dual simplex.

Adding New Variables and Constraints

- To add a new column A_{n+1},
  - Simply compute the reduced cost, $c_{n+1} - C_B B^{-1} A_{n+1}$, and the new column in the tableau, $B^{-1} A_{n+1}$.
  - Re-optimize using primal simplex.
- To add an inequality constraint $a_{m+1}^T x \geq b_{m+1}$,
  - Add a slack variable into the basis and compute the new tableau row, $[a^T B^{-1} A - a_{m+1}^T \;\; 1]$.
  - Re-optimize using dual simplex.
- To add an equality constraint $a_{m+1}^T x = b_{m+1}$,
  - Follow the same steps as above, except introduce an artificial variables instead of a slack variable.
  - The artificial variable should be given a large cost in the objective function to drive it out of the basis.

Local Sensitivity Analysis

- For changes in the right-hand side,
  - Recompute the values of the basic variables, $B^{-1}b$.
  - Re-solve using dual simplex if necessary.
- For a changes in the cost vector,
  - Recompute the reduced costs.
  - Re-solve using primal simplex.
- For changes in a nonbasic column A_j
  - Recompute the reduced cost, $c_j - c_B B^{-1} A_j$.
  - Recompute the column in the tableau, $B^{-1} A_j$.
- For all of these changes, we can compute ranges within which the current basis remains optimal.

Global Dependence on the Cost and Right-hand Side Vectors

- Consider a family of polyhedra parameterized by the vector b

$$\mathcal{P}(b) = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$$

  and the function $F(b) = \min_{x \in \mathcal{P}(b)} c^T x$.

- Consider the extreme points $p^1, \ldots, p^N$ of the dual polyhedron.
- There must be an extremal optimum to the dual and so

$$F(b) = \max_{i=1,\ldots,N} (p^i)^T b$$

- Hence, $F(b)$ is a piecewise linear convex function.

  Theorem 6. *Suppose that the linear program* $\min\{c^T x | Ax = \hat{b}, x \geq 0\}$ *is feasible and that the optimal cost is finite. Then* p *is an optimal solution to the dual if and only if it is a subgradient of* F *at* $\hat{b}$*.*

- A similar analysis applies to the cost vector.

Parametric Programming

- For a fixed matrix A, vectors b and c, and direction d and consider the problem

$$\begin{aligned} min \ (c+\theta d)^T x & \\ s.t. \qquad\qquad Ax &= b \\ x &\geq 0 \end{aligned}$$

- If $g(\theta)$ is the optimal cost for a given θ, then as before

$$g(\theta) = \min_{i=1,\dots,N} (c+\theta d)^T x^i$$

where $x^1, \dots, x^N$ are the extreme points of the feasible set.

- Hence, $g(\theta)$ is piecewise linear and concave.

The Parametric Simplex Method

- Determine an initial feasible basis.
- Determine the interval $[\theta_1, \theta_2]$ for which this basis is optimal.
- Determine a variable j whose reduced cost is nonpositive for $\theta \geq \theta_2$.
- If the corresponding column has no positive entries, then the problem is unbounded for $\theta > \theta_2$.
- Otherwise, rotate column j into the basis.
- Determine a new interval $[\theta_2, \theta_3]$ in which the current basis is optimal.
- Iterate to find all breakpoint $\geq \theta_1$.
- Repeat the process to find breakpoints $\leq \theta_1$.

Large-scale Linear Programming

- For large LPs, we really need only consider
  - Constraints that are binding at optimality.
  - Variables that are basic at optimality.
- We derived methods for dynamically generating variables and constraints during the solution process.
- For variables, the methods are based on calculating the reduced costs for columns that do not currently appear in the tableau.
- Finding the column with the most negative reduced cost is an optimization problem that can sometimes be solved efficiently.
- For constraints, the methods are similar but are based on finding the constraint with the most negative slack value.
- These methods are needed for efficiency and in cases where we don't want to generate the constraint matrix *a priori*.