DS 320 Algorithms for Data Science	**Linear Programming**
Spring 2024	Prof. Kira Goldner

Formulating the Primal

There are 3 main steps to formulating a linear program.

Step 1: Identify your decision variables. What are you trying to *decide* in this problem? Whether or not an item is taken into a set? How many of a good you sell? Often we want some x_i or x_{ij} to indicate these decisions, where ideally these variables are 0/1 or some integer of how many we would have.

Step 2: Formulating linear constraints. What are the constraints of the problem you're trying to solve, in English? Identify them, even write them out, and then try to write them in terms of notation and the decision variables. Do not multiply the variables together or the program is not linear. In "normal form," in a maximization problem, your variations times coefficients should sum to be less than or equal to something.

Step 3: Formulate the objective function. What is the goal of the problem? What are you trying to maximize? It should be a function of your decision variables, and it should be linear. All in all, your program should be of the form:

$$\begin{aligned} \max \quad & \mathbf{c}^T \mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \end{aligned}$$

Taking the Dual

There are 4 main steps to taking the dual of a primal linear program.

$$\begin{aligned} \min \quad & \mathbf{b}^T \mathbf{y} \\ \text{subject to} \quad & \mathbf{A}^T\mathbf{y} \geq \mathbf{c} \end{aligned}$$

Step 1: Introduce a variable for each primal constraint. Let the first constraint be y_1, the second be y_2, and so forth.

Step 2: Formulate the objective function. Assuming the primal is in normal form, the objective function of the dual is the sum of each new variable y_i multiplied with the right-hand side of its constraint in the primal (b_i); you are now minimizing.

Step 3: Formulate the constraints. Each constraint in the dual corresponds back to a variable in the primal. Suppose x_1 is the first primal variable—for its constraint in the dual, find x_1's coefficients in each constraint in the primal, and add them to x_1's dual constraint multiplied by the variable representing that constraint. That is, if $4\,x_1$ appears in the primal constraint for which we introduced y_2, then add $4\,y_2$ to the left-hand side of x_1's constraint in the dual. The right-hand side is the coefficient of x_1 in the primal's objective. The sign should be $\geq$.

Step 4: Sanity check. The dual problem should be the primal problem transposed, with the objective coefficients now forming the right-hand side of the constraints, the previous RHS of the constraints now forming the objective coefficients, and the LHS of the constraints coefficients transposed. Your min/max should have switched in the objective, and your $\leq$ and $\geq$ should have switched in the constraints.

Highlights

Some key points about linear programming:

- Linear programs are solvable in polynomial time (elipsoid method, simplex in average case).
- The objective of the dual gives an upper bound on the (maximization) objective of the primal for *any* feasible solutions. (Weak Duality)
- The objectives of the primal and dual are equal if and only if the solutions to the primal and dual are optimal. (Strong Duality)
- The complementary slackness conditions hold if and only if the solutions to the primal and dual are optimal.

An Example: Quality Lumbering

A lumber company can produce either pallets or high quality lumber. It cannot produce more than 200 units (thousand board feet) of lumber per day, which maxes out usage of their kiln, and it cannot produce more than 600 pallets per day. Its main saw can process at most 400 logs per day. 1 unit of lumber requires 1.4 high quality logs, and one pallet requires 0.25 low quality logs. High quality logs used for lumber cost \$200 per log, and low quality logs used for pallets cost \$4 per log. Processing lumber costs \$200 per unit, and processing pallets costs \$5. A unit of lumber sells for \$490 per unit, and a pallet sells for \$9. To maximize profits, how many pallets and units of lumber should the lumber company produce?

The Primal

Step 1: Decision Variables. We need to determine the number of pallets x_P, units of lumber x_L, and low and high quality logs q_L and q_H.

Step 2: Constraints. We know that lumber capacity is at most 200 units per day: $x_L \leq 200$ and pallets are at most 600 per day: $x_P \leq 600$. The main saw process 400 logs of any time, so $q_L + q_H \leq 400$. 1 unit of lumber requires 1.4 high quality logs: $1.4x_L \leq q_H$. Notice that if $x_L = 1$ it forces $q_H = 1.4$. Similarly, one pallet requires 0.25 low quality logs: $0.25x_P \leq q_L$.

Step 3: Objective. High quality logs used for lumber cost \$200 per log, and low quality logs used for pallets cost \$4 per log. Processing lumber costs \$200 per unit, and processing pallets costs \$5. A unit of lumber sells for \$490 per unit, and a pallet sells for \$9. That is, per unit, lumber earns \$290 and a pallet earns \$4. To maximize profit, our objective is then

$$290\, x_L + 4\, x_P - 200\, q_H - 4\, q_L.$$

All together, that gives for our primal:

$$
\begin{array}{rrr}
\max & 290\,x_L + 4\,x_P - 200\,q_H - 4\,q_L & & \\
\text{subject to} & x_L \le 200 & \text{(Lumber capacity)} & y_1 \\
 & x_P \le 600 & \text{(Pallet capacity)} & y_2 \\
 & 1.4\,x_L \le q_H & \text{(Lumber requirement)} & y_3 \\
 & 0.25\,x_P \le q_L & \text{(Pallet requirement)} & y_4 \\
 & q_L + q_H \le 400 & \text{(Main saw)} & y_5 \\
 & x_L, x_P, q_L, q_H \ge 0 & \text{(non-negativity)} &
\end{array}
$$

The Dual

Step 1: Variable per primal constraint. The first step is simply what's written above to the right—labeling each primal constraint with the y_i's.

Step 2: Objective function. The objective function for the dual will be min instead of max, and will pair the RHS of each primal constraint with the corresponding dual variable and sum them:

$$\min \quad 200\,y_1 + 600\,y_2 + 400\,y_5.$$

Step 3: Constraints. For an example, let us create the dual constraint corresponding to primal variable x_L. The variable x_L appears in y_1's constraint with a coefficient of 1, so we add $1\,y_1$. It appears in y_3's constraint with a coefficient of 1.4, so we add $1.4\,y_3$. It does not appear in any other constraints (except for non-negativity). It appears in the primal objective with a coefficient of 290, so our right-hand side is 290. This gives the dual constraint

$$y_1 + 1.4\,y_3 \ge 290. \qquad (x_L)$$

Repeating this process for each of the primal variables will give four dual constraints:

$$
\begin{array}{rr}
y_1 + 1.4\,y_3 \ge 290 & x_L \\
y_2 + 0.25\,y_4 \ge 4 & x_P \\
y_5 - y_3 \ge -200 & q_H \\
y_5 - y_4 \ge -4 & q_L
\end{array}
$$

All together, adding in non-negativity, we get our dual:

$$
\begin{array}{rrr}
\min & 200\,y_1 + 600\,y_2 + 400\,y_5 & \\
\text{subject to} & y_1 + 1.4\,y_3 \ge 290 & x_L \\
 & y_2 + 0.25\,y_4 \ge 4 & x_P \\
 & y_5 - y_3 \ge -200 & q_H \\
 & y_5 - y_4 \ge -4 & q_L \\
 & y_i \ge 0 & \forall i \quad \text{(non-negativity)}
\end{array}
$$

Step 4: Sanity check. The primal was not originally written in normal form—we need to rewrite y_3 and y_4's constraints as $1.4\,x_L - q_h \le 0$ and $0.25\,x_P - q_L \le 0$ respectively, but by doing so, we'll see that our transpose sanity check does indeed check out.