Duality in Linear Programming

Gary D. Knott
Civilized Software Inc.
12109 Heritage Park Circle
Silver Spring MD 20906
phone:301-962-3711
email:knott@civilized.com
URL:www.civilized.com

May 10, 2013

0.1 Duality in Linear Programming

Consider the following standard primal linear programming problem. (Here we impose $n \geq 1$ and $m \geq 1$. Also note that in this section the symbol c is not reserved to be an $(n+m)$-covector; it may have differing numbers of components in differing contexts.)

[Determine $x \in (\mathcal{R}^n)^\top$ to maximize $c^{\mathrm{T}}x$ subject to $A^o x \leq b$ and $x \geq 0$ where A^o is an $m \times n$ matrix, and $c \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$].

Now note that if $y \geq 0$ with $y \in (\mathcal{R}^m)^\top$ then $A^o x \leq b$ implies $y^{\mathrm{T}}A^o x \leq y^{\mathrm{T}}b$. This is because $y^{\mathrm{T}}A^o x \leq y^{\mathrm{T}}b$ is just the sum of the non-negatively-scaled inequalities $y_i[(A^o \text{ row } i, x) \leq b_i]$, *i.e.*, $\sum_{1\leq i\leq m} y_i(A^o \text{ row } i, x) \leq \sum_{1\leq i\leq m} y_i b_i$.

Now let $P = \{x \in (\mathcal{R}^n)^\top \mid A^o x \leq b \text{ and } x \geq 0\}$. Suppose x is a feasible point for our primal linear programming problem, so $x \in P$, *i.e.*, $A^o x \leq b$ and $x \geq 0$. Also suppose we can choose $y_1, \ldots, y_m$ defining an m-covector y such that $(y^{\mathrm{T}}A^o)_i \geq c_i$ and $y_i \geq 0$ for $i = 1, \ldots, n$. Then $c^{\mathrm{T}}x \leq y^{\mathrm{T}}A^o x$, and we have $y^{\mathrm{T}}A^o x \leq y^{\mathrm{T}}b$, so altogether we have $c^{\mathrm{T}}x \leq y^{\mathrm{T}}A^o x \leq y^{\mathrm{T}}b$. We see that $y^{\mathrm{T}}A^o x$ and $y^{\mathrm{T}}b$ are both upper-bounds of our primal objective function $\zeta(x) = c^{\mathrm{T}}x$ for admissible choices of x and y, and $y^{\mathrm{T}}b$ is independent of x, so for any admissible choice of y, $y^{\mathrm{T}}b$ is an upper-bound of $c^{\mathrm{T}}x$ for all feasible points $x \in P$.

> **Exercise 0.1:** Why must we assume both $A^o x \leq b$ and $x \geq 0$ to conclude that $c^{\mathrm{T}}x \leq y^{\mathrm{T}}A^o x \leq y^{\mathrm{T}}b$?

If we want to make $y^{\mathrm{T}}b = b^{\mathrm{T}}y$ a close upper-bound for $c^{\mathrm{T}}x$, we will want to choose y, subject to $y^{\mathrm{T}}A^o \geq c^{\mathrm{T}}$ and $y \geq 0$, to make $b^{\mathrm{T}}y$ as small as possible. This is just a linear programming problem:

[determine $y \in (\mathcal{R}^m)^\top$ to minimize $b^{\mathrm{T}}y$ subject to $(A^o)^{\mathrm{T}}y \geq c$ and $y \geq 0$ where A^o is an $m \times n$ matrix, and $c \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$].

This linear programming problem is called the standard *dual problem* associated with our standard primal linear programming problem. If this dual problem has a feasible point then it has an optimal point. Moreover, for every feasible point, y, of our dual problem, we have $c^{\mathrm{T}}x \le b^{\mathrm{T}}y$ for all feasible points x of the corresponding primal problem. (Note if this dual problem is non-empty, *i.e.*, has a feasible point, and the only optimal points for this dual problem are points at infinity, then the value of the objective function $b^{\mathrm{T}}y$ will be $-\infty$ at such an optimal point and the primal problem can have no solution. This is because $b^{\mathrm{T}}y$ is *minimized* when y is an optimal point.)

Exercise 0.2: Suppose our dual problem has a feasible point. Why must $b^{\mathrm{T}}\hat{y} = -\infty$ when all the optimal points of our dual problem are points at infinity? Hint: consider how we might go from a finite feasible point to a point-at-infinity optimal point.

Exercise 0.3: If $b = 0$, is the standard dual problem guaranteed to have a solution?

Exercise 0.4: Show that the dual problem of the linear programming problem: [determine $y \in (\mathcal{R}^m)^\top$ to minimize $b^{\mathrm{T}}y$ subject to $(A^o)^{\mathrm{T}}y \ge c$ and $y \ge 0$ where A^o is an $m \times n$ matrix, and $c \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$] is just our original primal programming problem: [determine $x \in (\mathcal{R}^n)^\top$ to maximize $c^{\mathrm{T}}x$ subject to $A^o x \le b$ and $x \ge 0$].

Solution 0.4: Constructing the dual problem of a given primal linear programming problem is a *syntactic* process. When we state a given linear programming problem in the standard form: [determine $x \in (\mathcal{R}^n)^\top$ to maximize $c^{\mathrm{T}}x$ subject to $A^o x \le b$ and $x \ge 0$ where A^o is an $m \times n$ matrix, and $c \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$]. Then the dual problem is: [determine $y \in (\mathcal{R}^m)^\top$ to minimize $b^{\mathrm{T}}y$ subject to $(A^o)^{\mathrm{T}}y \ge c$ and $y \ge 0$].

This dual problem is equivalent to the standard-form problem: [determine $y \in (\mathcal{R}^m)^\top$ to maximize $-b^{\mathrm{T}}y$ subject to $(-A^o)^{\mathrm{T}}y \le -c$ and $y \ge 0$].

Syntactically, we replace c^{T} with $-b^{\mathrm{T}}$, b with $-c$, and A^o with $(-A^o)^{\mathrm{T}}$ in our standard primal linear programming problem in order to construct its dual problem in standard form.

In array form, the array $\begin{bmatrix} A^o & b \\ c^{\mathrm{T}} & 0 \end{bmatrix}$ describing our primal problem becomes the array $\begin{bmatrix} (-A^o)^{\mathrm{T}} & -c \\ -b^{\mathrm{T}} & 0 \end{bmatrix}$ describing our dual problem. This is sometimes summarized by saying the dual problem is the *negative transpose* of the primal problem.

Now note that the negative transpose of $\begin{bmatrix} (-A^o)^{\mathrm{T}} & -c \\ -b^{\mathrm{T}} & 0 \end{bmatrix}$ is just $\begin{bmatrix} A^o & b \\ c^{\mathrm{T}} & 0 \end{bmatrix}$ again. Thus the dual of the dual of our primal problem is our primal problem itself.

Exercise 0.5: Show that the dual problem of the linear programming problem: [determine $x \in (\mathcal{R}^k)^\top$ to maximize $c^{\mathrm{T}}x$ subject to $Ax = b$ and $x \ge 0$ where A is an $m \times k$ matrix, and $c \in (\mathcal{R}^k)^\top$, and $b \in (\mathcal{R}^m)^\top$] is the problem: [determine $w \in (\mathcal{R}^m)^\top$ to minimize $b^{\mathrm{T}}w$ subject to $A^{\mathrm{T}}w \ge c$]. Note c and x are here k-covectors. This fluidity in the meanings of the symbols c and x is a potential source of confusion herein; beware. (We shall sometimes write c^o to indicate $c^o \in (\mathcal{R}^n)^\top$, but this is not uniformly done because it is often counterproductive.)

Solution 0.5: Put this problem is standard form by writing $Ax = b$ as $\begin{bmatrix} A \\ -A \end{bmatrix} x \le \begin{bmatrix} b \\ -b \end{bmatrix}$.

Then the primal problem array is $\begin{bmatrix} A & b \\ -A & -b \\ c^{\mathrm{T}} & 0 \end{bmatrix}$ and the corresponding negative transpose array for the dual problem is $\begin{bmatrix} -A^{\mathrm{T}} & A^{\mathrm{T}} & -c \\ -b^{\mathrm{T}} & b^{\mathrm{T}} & 0 \end{bmatrix}$, and this specifies the dual problem as: [determine $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ to maximize $\begin{bmatrix} -b^{\mathrm{T}} & b^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ subject to $\begin{bmatrix} -A^{\mathrm{T}} & A^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \le -c$ and $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \ge 0$ where y_1 and y_2 are m-covectors].

And this problem can be restated as: [determine w to minimize $b^{\mathrm{T}} w$ subject to $A^{\mathrm{T}} w \ge c$]. This is because

$$\begin{aligned} \max_{y_1,y_2} \begin{bmatrix} -b^{\mathrm{T}} & b^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \max_{y_1,y_2} -b^{\mathrm{T}} y_1 + b^{\mathrm{T}} y_2 \\ &= \max_{y_1,y_2} b^{\mathrm{T}} [-y_1 + y_2] \\ &= \max_{w} b^{\mathrm{T}} (-w) \\ &= \min_{w} b^{\mathrm{T}} w \end{aligned}$$

where $w = y_1 - y_2$. And $\begin{bmatrix} -A^{\mathrm{T}} & A^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \le -c$ and $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \ge 0$ is equivalent to $A^{\mathrm{T}}(-y_1 + y_2) \le -c$ and $y_1 \ge 0$ and $y_2 \ge 0$, which is equivalent to $A^{\mathrm{T}} w \ge c$ with $w \in (\mathcal{R}^m)^{\top}$ otherwise unrestricted. (Given w, we may determine $y_1 \ge 0$ and $y_2 \ge 0$ as follows. Take $(y_1)_i = w_i$ when $w_i \ge 0$ and take $(y_1)_i = 0$ otherwise. And take $(y_2)_i = -w_i$ when $w_i < 0$ and take $(y_2)_i = 0$ otherwise.)

Note if $k = n+m$ and $A = [A^o \quad I]$ where A^o is an $m \times n$ matrix and $c = \begin{bmatrix} c^o \\ 0 \end{bmatrix}$ with $c^o \in (\mathcal{R}^n)^{\top}$, then this dual problem becomes [determine $w \in (\mathcal{R}^m)^{\top}$ to minimize $b^{\mathrm{T}} w$ subject to $(A^o)^{\mathrm{T}} w \ge c^o$ and $w \ge 0$].

Exercise 0.6: Show that the linear programming problem: [determine w to minimize $b^{\mathrm{T}} w$ subject to $A^{\mathrm{T}} w \ge c$] is equivalent to the linear programming problem: [determine z to maximize $b^{\mathrm{T}} z$ subject to $A^{\mathrm{T}} z \le -c$].

Exercise 0.7: Show that the dual problem of the slack-variable form of our primal linear programming problem is equivalent to the standard problem dual to our standard primal problem.

Solution 0.7: Put the slack-variable form of our primal problem in standard form as: [determine x to maximize $[(c^o)^{\mathrm{T}} \quad 0] \begin{bmatrix} x^o \\ t \end{bmatrix}$ subject to $\begin{bmatrix} A & I \\ -A & -I \end{bmatrix} \begin{bmatrix} x^o \\ t \end{bmatrix} \le \begin{bmatrix} b \\ -b \end{bmatrix}$]. Here t represents the slack variables.

Then the corresponding array representation is $\begin{bmatrix} A^o & I & b \\ -A^o & -I & -b \\ (c^o)^{\mathrm{T}} & 0 & 0 \end{bmatrix}$ and the corresponding negative transpose array is $\begin{bmatrix} (-A^o)^{\mathrm{T}} & (A^o)^{\mathrm{T}} & -c^o \\ (-I)^{\mathrm{T}} & I^{\mathrm{T}} & 0 \\ -b^{\mathrm{T}} & b^{\mathrm{T}} & 0 \end{bmatrix}$, and this specifies the dual problem: [determine $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ to maximize $\begin{bmatrix} -b^{\mathrm{T}} & b^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ subject to $\begin{bmatrix} (-A^o)^{\mathrm{T}} & (A^o)^{\mathrm{T}} \\ (-I)^{\mathrm{T}} & I^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq \begin{bmatrix} -c^o \\ 0 \end{bmatrix}$]. And this problem is equivalent to [determine $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ to maximize $-b^{\mathrm{T}}(y_1 - y_2)$ subject to $(-A^o)^{\mathrm{T}}y_1 + (A^o)^{\mathrm{T}}y_2 \leq -c^o$ and $-y_1 + y_2 \leq 0$] which is in turn equivalent to [determine w to maximize $-b^{\mathrm{T}}w$ subject to $(-A^o)^{\mathrm{T}}w \leq -c^o$ and $-w \leq 0$] where $w = y_1 - y_2$, and finally, this is equivalent to [determine w to minimize $b^{\mathrm{T}}w$ subject to $(A^o)^{\mathrm{T}}w \geq c^o$ and $w \geq 0$].

Note the array form we used to introduce the negative transpose dual is not the form we use in the tableau representation of the simplex algorithm. This is because we require $\begin{bmatrix} A^o \\ c^{\mathrm{T}} \end{bmatrix}$ to transform to $\begin{bmatrix} (-A^o)^{\mathrm{T}} & -c \end{bmatrix}$ corresponding to $(A^o)^{\mathrm{T}} \geq c$, while we need $\begin{bmatrix} b \\ 0 \end{bmatrix}$ to transform to $\begin{bmatrix} -b^{\mathrm{T}} & 0 \end{bmatrix}$; if we were to change the sign of c, we would need to change the sign of b also.

0.1.1 The Strong Duality Theorem

Let P be the polyhedron of our primal linear programming problem: [determine $x \in (\mathcal{R}^n)^\top$ to maximize $c^{\mathrm{T}}x$ subject to $A^o x \leq b$ and $x \geq 0$ where A^o is an $m \times n$ matrix, and $c \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$] and let Q be the polyhedron of the associated dual linear programming problem: [determine $y \in (\mathcal{R}^m)^\top$ to minimize $b^{\mathrm{T}}y$ subject to $(A^o)^{\mathrm{T}}y \geq c$ and $y \geq 0$]. Thus $P = \{x \in (\mathcal{R}^n)^\top \mid A^o x \leq b \text{ and } x \geq 0\}$ and $Q = \{y \in (\mathcal{R}^m)^\top \mid (A^o)^{\mathrm{T}}y \geq c \text{ and } y \geq 0\}$.

We saw above that if $x \in P$ and $y \in Q$, then $c^{\mathrm{T}}x \leq b^{\mathrm{T}}y$. This is called the *weak-duality theorem*. Thus if $c^{\mathrm{T}}x = b^{\mathrm{T}}y$ with $x \in P$ and $y \in Q$, then x must be an optimal point of P that maximizes $c^{\mathrm{T}}x$ and y must be an optimal point of Q that minimizes $b^{\mathrm{T}}y$. This is because $b^{\mathrm{T}}y \geq \max\{c^{\mathrm{T}}v \mid v \in P\}$. And when $b^{\mathrm{T}}y = c^{\mathrm{T}}x$, $b^{\mathrm{T}}y$ is the least possible upper-bound of $\{c^{\mathrm{T}}v \mid v \in P\}$, and thus $b^{\mathrm{T}}y = c^{\mathrm{T}}x = \max\{c^{\mathrm{T}}v \mid v \in P\}$, so x must be a solution of our primal linear programming problem. Similarly, $c^{\mathrm{T}}x$ is the greatest possible lower-bound of $\{b^{\mathrm{T}}w \mid w \in Q\}$, and thus y must be a solution of our dual linear programming problem when $c^{\mathrm{T}}x = b^{\mathrm{T}}y$.

Also we have the following corollary of the weak duality theorem. If the polyhedron P of our primal linear programming problem contains a point at infinity x for which the objective function value $c^{\mathrm{T}}\hat{x} = \infty$, then, since $b^{\mathrm{T}}y \geq c^{\mathrm{T}}\hat{x}$ whenever $y \in Q$, $b^{\mathrm{T}}y$ must also be ∞ for all $y \in Q$. But this is impossible if Q is non-empty, since Q then contains a finite feasible point y', *i.e.*, Q cannot be composed entirely of points at infinity, and then the dual objective function value $b^{\mathrm{T}}y'$ is finite and the minimal value of $b^{\mathrm{T}}y$ must be no greater than $b^{\mathrm{T}}y'$. Hence Q must be empty. Thus if our primal linear programming problem is unbounded with an unbounded objective function value, then the dual linear programming problem must be empty *i.e.*, $Q = \emptyset$. Conversely, if Q is non-empty, our primal problem must have a finite optimal point.

Similarly, if the polyhedron Q of our dual linear programming problem contains a point at infinity $\hat{y}$ as an optimal point for which the objective function vaue $b^{\mathrm{T}}\hat{y} = -\infty$, then the primal linear programming problem must be empty, *i.e.*, $P = \emptyset$. And conversely, if P is non-empty, our dual problem must have a finite optimal point.

Remember a linear programming problem may have an associated unbounded polyhedron without its objective function being $\pm\infty$, but if the objective function is $\pm\infty$ then the associated polyhedron is necessarily unbounded.

If either our primal or dual linear programming problem has a finite optimal point, then so does the other problem. And in fact, the values of their respective objective functions must be identical! This fact is called the *strong-duality theorem*.

Exercise 0.8: Give an example where P is unbounded but the dual problem is non-empty and has a finite optimal point.

Exercise 0.9: Give an example where $P = \emptyset$ and $Q = \emptyset$.

Solution 0.9:
primal: [determine x to maximize x_1 subject to $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} x \le \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ and $x \ge 0$].
dual: [determine y to minimize $-y_1$ subject to $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} y \ge \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $y \ge 0$].

Altogether we have the following possiblities for our standard-form primal and dual problems.

1. Our primal problem has a finite optimal point $\hat{x}$ and our dual problem has a finite optimal point $\hat{y}$ and $c^{\mathrm{T}}\hat{x} = b^{\mathrm{T}}\hat{y}$. Also, if one problem has a face of finite solutions of dimension greater than 0, then the finite solution of the other problem is an overdetermined vertex.

2. Our primal problem has a point-at-infinity optimal point $\hat{x}$ and $c^{\mathrm{T}}\hat{x} = \infty$, and our dual problem has no solution, (Q is empty.)

3. Our dual problem has a point-at-infinity optimal point $\hat{y}$ and $b^{\mathrm{T}}\hat{y} = -\infty$, and our primal problem has no solution, (P is empty.)

4. Neither our primal problem or our dual problem have solutions, (P is empty and Q is empty.)

Recall we have a criterion that specifies exactly when $P = \{x \in (\mathcal{R}^n)^\top \mid A^o x \le b \text{ and } x \ge 0\}$ is empty, namely: P is empty exactly when $b \notin K_0 = \{p \in (\mathcal{R}^m)^\top \mid p = [A^o \quad I]x \text{ and } x \ge 0\}$ which is equivalent to the criterion: there exists $y \in (\mathcal{R}^m)^\top$ such that $\begin{bmatrix} (A^o)^{\mathrm{T}} \\ I \end{bmatrix} y \ge 0$ and $b^{\mathrm{T}} y < 0$.

The possibilities above are somewhat subtle when P or Q is unbounded. For example, consider the primal problem: [find $x \in (\mathcal{R}^2)^\top$ to maximize $c^{\mathrm{T}}x$ subject to $x \in P$ where $P = \{x \in (\mathcal{R}^2)^\top \mid A^o x \le b \text{ and } x \ge 0\}$ with $A^o = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$ and $b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, *i.e.*, $P = \left\{x \in (\mathcal{R}^2)^\top \mid \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and } x \ge 0\right\}$]. Note $b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ means P is a cone with apex 0. The dual problem is: [find $y \in (\mathcal{R}^2)^\top$ to minimize 0 subject to $y \in Q$ where $Q = \left\{y \mid \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \ge \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \text{ and } y \ge 0\right\}$].

In the case where $c = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, all the points in P, including points at infinity, are optimal points, but $c^{\mathrm{T}}x = 0$ for $x \in P$; and the polyhedron Q is the singleton set $\{(0,0)^{\mathrm{T}}\}$ so the solution of the dual problem is the finite point $\hat{y} = 0^{\mathrm{T}}0 = 0$. Also note $(0,0)^{\mathrm{T}}$ is an overdetermined vertex of Q. (And $(0,0)^{\mathrm{T}}$ is an overdetermined vertex of P as well.)

In the case where $c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, we have the set of points $\{(\infty, x_2)^{\mathrm{T}} \mid x_2 \geq 0\} \subset P$ as optimal points of our primal problem with $c^{\mathrm{T}}x = \infty$ for $x \in \{(\infty, x_2)^{\mathrm{T}} \mid x_2 \geq 0\}$; all these optimal points are points at infinity. And the dual problem has $Q = \emptyset$.

In the case where $c = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, we have the set of points $\{(0, x_2)^{\mathrm{T}} \mid x_2 \geq 0\} \subset P$ as optimal points of our primal problem with $c^{\mathrm{T}}x = 0$ for $x \in \{(0, x_2)^{\mathrm{T}} \mid x_2 \geq 0\}$; all these optimal points are finite $(\mathcal{R}^2)^{\top}$-points. And the dual problem has every point of Q as an optimal point, where $Q = \{y \in (\mathcal{R}^2)^{\top} \mid 0 \leq y_1 \leq 1 \text{ and } y_1 \geq y_2 \geq 0\}$ and $b^{\mathrm{T}}y = 0^{\mathrm{T}}y = 0$ for $y \in Q$. Note Q is bounded.

> **Exercise 0.10:** With respect to the example discussed above, what is the set Q, and the set of solutions of our primal problem, and the set of solutions of our dual problem, when $c^{\mathrm{T}} = (-1, -1)$? What are the solution sets when $c^{\mathrm{T}} = (-1, 1)$?

The fact that, if $\hat{x} \in P$ is a finite optimal point of our standard-form primal problem then there exists $\hat{y} \in Q$ such that $\hat{y}$ is a finite optimal point of our standard-form dual problem, and conversely, and moreover $c^{\mathrm{T}}\hat{x} = b^{\mathrm{T}}\hat{y}$, can be proven as follows.

Suppose $\hat{x}$ is a finite optimal point of our primal linear programming problem with $\hat{x}$ a vertex of P. Then the gradient covector c belongs to the normal cone of $\hat{x}$ with respect to P.

Recall that $P = H_1 \cap \cdots \cap H_n \cap H_{n+1} \cap \cdots \cap H_{n+m}$ where $H_i = \{x \in (\mathcal{R}^n)^{\top} \mid (-e_i^{\mathrm{T}}, x) \leq 0\}$ for $i = 1, \ldots, n$, and $H_{n+j} = \{x \in (\mathcal{R}^n)^{\top} \mid (a_j^{\mathrm{T}}, x) \leq b_j\}$ for $j = 1, \ldots, m$ where $a_j = A^o$ row j. Here $-e_1^{\mathrm{T}}, \ldots, -e_n^{\mathrm{T}}, a_1^{\mathrm{T}}, \ldots, a_m^{\mathrm{T}}$ are all outwardly-directed normal vectors of the half-spaces whose intersection defines the polyhedron P. Thus $P = \left\{ x \in (\mathcal{R}^n)^{\top} \mid \begin{bmatrix} A^o \\ -I \end{bmatrix} x \leq \begin{bmatrix} b \\ 0 \end{bmatrix} \right\}$.

Let $H_{i-1}, \ldots, H_{i_q}$ be the half-spaces whose boundaries contain the optimal vertex $\hat{x}$; this means the bounding hyperplanes $\partial H_{i_1}, \ldots, \partial H_{i_q}$ intersect at the vertex $\hat{x}$.

Let $v_1, \ldots, v_q$ denote the outwardly-directed normal vectors of the half-spaces $H_{i_1}, \ldots, H_{i_q}$. Because $\partial H_{i_1} \cap \cdots \cap \partial H_{i_q} = \{\hat{x}\}$, we have $q \geq n$ and $dim(\{v_1, \ldots, v_q\}) = n$.

The normal cone at $\hat{x}$ with respect to P is the polyhedral cone $\mathcal{N}_{\hat{x}} = cone_0(v_1, \ldots, v_q)$, and we have seen that $c \in \mathcal{N}_{\hat{x}}$ because $\hat{x}$ is an optimal vertex of P that maximizes (c, x). Thus c is a conic-combination of the covectors $v_1, \ldots, v_q$. (Remember a conic-combination is a linear combination of vectors (or covectors) with non-negative scalar coefficients.) The covectors $v_1, \ldots, v_q$ all appear as *columns* of the $m \times (n+m)$ matrix $\begin{bmatrix} (A^o)^{\mathrm{T}} & -I \end{bmatrix}$. (Recall $\mathcal{N}_{\hat{x}}$ is just the polar dual cone of the 0-apex cone $H_{i_1} \cap H_{i_2} \cap \cdots \cap H_{i_q} - \hat{x}$.)

Carathéodory's theorem for polyhedral cones asserts that we can write c as a conic-combination of a subset of n of the covectors $v_1, \ldots, v_q$. Thus we can write

$$c = \begin{bmatrix} (A^o)^{\mathrm{T}} & -I \end{bmatrix} \begin{bmatrix} y' \\ z \end{bmatrix}$$

where $y' \in (\mathcal{R}^m)^\top$ and $z \in (\mathcal{R}^n)^\top$ with $y' \geq 0$ and $z \geq 0$, and at most n of the components of the covector $\begin{bmatrix} y' \\ z \end{bmatrix}$ are positive and the rest are zero.

Exercise 0.11: Show that the columns of the matrix $\begin{bmatrix} (A^o)^{\mathrm{T}} & -I \end{bmatrix}$ are the outwardly-directed normal covectors $a_1^{\mathrm{T}}, \ldots, a_m^{\mathrm{T}}, -e_1^{\mathrm{T}}, \ldots, -e_n^{\mathrm{T}}$.

Thus $c = (A^o)^{\mathrm{T}} y' - z$, and since y' is a non-negative m-covector and z is a non-negative $(n-m)$-covector, we have $c \leq (A^o)^{\mathrm{T}} y'$ and $y' \geq 0$. Note this is valid in the special case where $c = 0$.

Exercise 0.12: Explain why the relations $c \leq (A^o)^{\mathrm{T}} y'$ and $y' \geq 0$ are valid when $P = \{0\}$. What if an entire facet of P is comprised of optimal points of P?

Exercise 0.13: Show that if $\hat{x} \neq 0$, then at least one facet of P containing $\hat{x}$ lies in a non-orthant-boundary hyperplane among $\partial H_{n+1}, \ldots, \partial H_{n+m}$, *i.e.*, at least one of the inequalities $A\hat{x} \leq b$ is tight and at least one component of y' is positive.

The relations $c \leq (A^o)^{\mathrm{T}} y'$ and $y' \geq 0$ mean $y' \in Q$ (!) So our dual problem has y' as a feasible point and therefore our dual problem is not empty and hence has an optimal point $\hat{y}$ with $b^{\mathrm{T}}\hat{y} \leq b^{\mathrm{T}} y'$. And $c^{\mathrm{T}}\hat{x} \leq b^{\mathrm{T}}\hat{y}$ by the weak duality theorem, so $b^{\mathrm{T}}\hat{y} > -\infty$ and hence $\hat{y}$ is finite when $\hat{x}$ is finite.

Suppose $\hat{y}$ is a finite optimal point of our dual linear programming problem with $\hat{y}$ a vertex of Q. Then the dual problem gradient covector $-b$ belongs to the normal cone of $\hat{y}$ with respect to Q.

Since $Q = \left\{ y \in (\mathcal{R}^m)^\top \mid \begin{bmatrix} -(A^o)^{\mathrm{T}} \\ -I \end{bmatrix} y \leq \begin{bmatrix} -c \\ 0 \end{bmatrix} \right\}$ and the gradient vector of our dual problem is $-b$, we may follow the same line of argument as used above to characterize $-b \in \mathcal{N}_{\hat{y}}$ by writing

$$-b = \begin{bmatrix} -A^o & -I \end{bmatrix} \begin{bmatrix} x' \\ v \end{bmatrix}$$

where $x' \in (\mathcal{R}^n)^\top$ and $v \in (\mathcal{R}^m)^\top$ with $x' \geq 0$ and $v \geq 0$ and at most m of the components of the covector $\begin{bmatrix} x' \\ v \end{bmatrix}$ are positive and the rest are zero. Here the columns of $\begin{bmatrix} -A^o & -I \end{bmatrix}$ are outward-directed normals of the half-space boundaries whose intersection defines $\hat{y}$. Thus $-b = -A^o x' - v$ where $x' \geq 0$ and $v \geq 0$, so $b = A^o x' + v$ and $x' \geq 0$ and $v \geq 0$, and thus $A^o x' \leq b$ with $x' \geq 0$.

The relations $A^o x' \leq b$ and $x' \geq 0$ mean $x' \in P$ so our primal problem has x' as a feasible point and therefore our primal problem is not empty and hence also has an optimal point $\hat{x}$. And we have $c^{\mathrm{T}}\hat{x} \leq b^{\mathrm{T}}\hat{y}$ by the weak duality theorem, so $c^{\mathrm{T}}\hat{x} < \infty$ and hence $\hat{x}$ is finite when $\hat{y}$ is finite.

Now let us return to the representation of the gradient covector c as a conic combination of n of the outwardly-directed covectors $-e_1^{\mathrm{T}}, \ldots, -e_n^{\mathrm{T}}, a_1^{\mathrm{T}}, \ldots, a_m^{\mathrm{T}}$ where $a_j = A^o$ row j. Let $j_1, \ldots, j_n$ be the indices in $\{1, \ldots, n+m\}$ of the hyperplanes $\partial H_{j_1}, \ldots, \partial H_{j_n}$ from among the hyperplanes

$\partial H_1, \ldots, \partial H_{n+m}$ such that $\partial H_{j_1}, \ldots, \partial H_{j_n}$ are linearly-independent hyperplanes that intersect at $\{\hat{x}\}$. The corresponding constraints are *tight* at $\hat{x}$ and hence we may choose the covector $\begin{bmatrix} y' \\ z \end{bmatrix}$ so that zero or more of the non-negative coefficients in the covector $\begin{bmatrix} y' \\ z \end{bmatrix}$ that multiply the outwardly-directed normals of the hyperplanes $\partial H_{j_1}, \ldots, \partial H_{j_n}$ in the columns of $[(A^o)^{\mathrm{T}} \quad -I]$ will be positive. and *all the other components of* $\begin{bmatrix} y' \\ z \end{bmatrix}$ *will be zero.*

Exercise 0.14: Show that $\{j_1, \ldots, j_n\} \subseteq \{i_1, \ldots, i_q\}$.

Exercise 0.15: Explain why we said 'zero or more' above.

Let $p = \begin{bmatrix} y' \\ z \end{bmatrix}$, so that p row $1 : m = y'$ and p row $(m+1) : (m+n) = z$. Let $t = b - A^o\hat{x}$. Define the sequence $J = \langle j_1, \ldots, j_n \rangle$ and define the sequence $K = \langle 1, \ldots, n+m \rangle - J$. (The ordering of the indices in J and K are not important.) We have p row $J \geq 0$ and p row $K = 0$. The hyperplane normals indexed by K include all the slack constraints. Thus $[(A^o)^{\mathrm{T}} \quad -I]$ col K contains all the columns of $[(A^o)^{\mathrm{T}} \quad -I]$ where $(A^o \text{ row } i)^{\mathrm{T}}\hat{x} < b_i$ for $1 \leq i \leq m$ or $-e_i^{\mathrm{T}}\hat{x} < 0$ for $1 \leq i \leq n$.

The components of p corresponding to these slack columns lie in p row K and are thus all zero. The non-zero components of p all lie in p row J and correspond to tight columns of $[(A^o)^{\mathrm{T}} \quad -I]$ where $(A^o \text{ row } i, \hat{x}) = 0$ or $-\hat{x}_i = 0$. This means $[(b - A^o\hat{x})^{\mathrm{T}} \quad \hat{x}^{\mathrm{T}}]p = 0$! That is

$$[(b - A^o\hat{x})^{\mathrm{T}} \quad \hat{x}^{\mathrm{T}}] \begin{bmatrix} y' \\ z \end{bmatrix} = 0.$$

Thus $(b - A^o\hat{x})^{\mathrm{T}}y' + \hat{x}^{\mathrm{T}}z = 0$.

Exercise 0.16: Show that $t^{\mathrm{T}}y' + \hat{x}^{\mathrm{T}}z = 0$ where $t = b - A^o\hat{x}$, and hence $t^{\mathrm{T}}y' = 0$ and $\hat{x}^{\mathrm{T}}z = 0$.

Recall we have $c = (A^o)^{\mathrm{T}}y' - z$. Thus

$$\begin{aligned} b^{\mathrm{T}}y' &= \hat{x}^{\mathrm{T}}(A^o)^{\mathrm{T}}y' - \hat{x}^{\mathrm{T}}z \\ &= \hat{x}^{\mathrm{T}}((A^o)^{\mathrm{T}}y' - z) \\ &= \hat{x}^{\mathrm{T}}c \\ &= c^{\mathrm{T}}\hat{x}. \end{aligned}$$

Therefore, as we have seen above, y' must be an optimal point of our dual problem and we have $b^{\mathrm{T}}\hat{y} = c^{\mathrm{T}}\hat{x}$ where $\hat{y} = y'$. The conclusion that y' is an optimal point of Q is a consequence of the corollary of the weak duality theorem discussed above.

Similarly, we may follow the line of reasoning parallel to that used above to establish the identity

$$[((-A^o)^{\mathrm{T}}\hat{y} + c)^{\mathrm{T}} \quad -\hat{y}^{\mathrm{T}}] \begin{bmatrix} x' \\ v \end{bmatrix} = 0.$$

Thus $-\hat{y}^{\mathrm{T}}A^o x' + c^{\mathrm{T}}x' - \hat{y}^{\mathrm{T}}v = 0$. But then $c^{\mathrm{T}}x' = \hat{y}^{\mathrm{T}}(A^o x' + v) = \hat{y}^{\mathrm{T}}b = b^{\mathrm{T}}\hat{y}$ since we showed that $v + A^o x' = b$ above. Therefore x' must be an optimal point of our primal problem and again we have $c^{\mathrm{T}}\hat{x} = b^{\mathrm{T}}\hat{y}$ where $\hat{x} = x'$. (We also have $t = v$ because $x' = \hat{x}$.) [QED]

We have $c^{\mathrm{T}}x \leq y^{\mathrm{T}}A^o x \leq b^{\mathrm{T}}y$ when $x \in P$ and $y \in Q$; and when $\hat{x} \in P$ maximizes $c^{\mathrm{T}}x$ subject to $x \in P$ and $\hat{y} \in Q$ minimizes $b^{\mathrm{T}}y$ subject to $y \in Q$, we have $c^{\mathrm{T}}\hat{x} = \hat{y}^{\mathrm{T}}A^o\hat{x} = b^{\mathrm{T}}\hat{y}$.

Now we can deduce $\hat{y}^{\mathrm{T}}(b - A^o\hat{x}) = 0$ from $\hat{y}^{\mathrm{T}}A^o\hat{x} = b^{\mathrm{T}}\hat{y}$. And we can deduce $\hat{x}^{\mathrm{T}}((A^o)^{\mathrm{T}}\hat{y} - c) = 0$ from $c^{\mathrm{T}}\hat{x} = \hat{y}^{\mathrm{T}}A^o\hat{x}$. The identities $\hat{y}^{\mathrm{T}}(b - A^o\hat{x}) = 0$ and $\hat{x}^{\mathrm{T}}((A^o)^{\mathrm{T}}\hat{y} - c) = 0$ are called the *complementary slackness conditions*.

Exercise 0.17: We have already obtained the complementary slackness conditions above. Identify where they are derived in the proof of the strong duality theorem.

Let $t = b - A^o\hat{x}$; we have $t \geq 0$ because $A^o\hat{x} \leq b$. The elements of the m-covector t are the slack variables of our primal linear programming problem. Similarly, Let $z = (A^o)^{\mathrm{T}}\hat{y} - c$; we have $z \geq 0$ because $(A^o)^{\mathrm{T}}\hat{y} \geq c$. The elements of the n-covector z are the slack variables of our dual linear programming problem.

The complementary slackness conditions are then $\hat{y}t = 0$ and $\hat{x}z = 0$. Since $\hat{y} \geq 0$, $t \geq 0$, $\hat{x} \geq 0$, and $z \geq 0$, these conditions imply that if $\hat{y}_i \neq 0$, $t_i = 0$, and if $t_i \neq 0$, $\hat{y}_i = 0$. And similarly, if $\hat{x}_i \neq 0$, $z_i = 0$, and if $z_i \neq 0$, $\hat{x}_i = 0$. In fact, unless $\hat{x}$ is an overdetermined vertex of P, exactly one of $\hat{x}_i$ and z_i are zero, and unless $\hat{y}$ is an overdetermined vertex of Q, exactly one of $\hat{y}_i$ and t_i are zero

0.1.2 Computing a Solution of the Dual Problem

The simplex algorithm can be applied to our dual problem recast as: [determine $y \in (\mathcal{R}^m)^\top$ to maximize $-b^{\mathrm{T}}y$ subject to $(-A^o)^{\mathrm{T}}y \leq -c^o$ and $y \geq 0$ where A^o is an $m \times n$ matrix, and $c^o \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$], but in fact the simplex algorithm applied to the corresponding primal problem, (*i.e.*, to the dual of this problem,) also produces an optimal point of this dual problem when our primal problem has a finite optimal point (!) We can see this as follows.

Recall our primal problem is: [determine $x^o \in (\mathcal{R}^n)^\top$ to maximize $(c^o)^{\mathrm{T}}x^o$ subject to $A^o x^o \leq b$ and $x^o \geq 0$ where A^o is an $m \times n$ matrix, and $c^o \in (\mathcal{R}^n)^\top$, $x^o \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$]. This corresponds to the slack-variable form problem: [determine $x \in (\mathcal{R}^{n+m})^\top$ to maximize $c^{\mathrm{T}}x$ subject to $Ax = b$ and $x \geq 0$ where $A = [A^o \quad I]$ is an $m \times (n+m)$ matrix, and $c^{\mathrm{T}} = [(c^o)^{\mathrm{T}} \quad 0] \in \mathcal{R}^{n+m}$, $x^{\mathrm{T}} = [(x^o)^{\mathrm{T}} \quad t^{\mathrm{T}}] \in \mathcal{R}^{n+m}$, (the elements of the covector x row $(n+1) : m = t$ constitute our slack variables,) and $b \in (\mathcal{R}^m)^\top$].

When the simplex algorithm applied to our slack-variable primal problem halts with a finite optimal point $\hat{x} \in (\mathcal{R}^{n+m})^\top$, we have a length-$n$ basis sequence N and a corresponding orthant-boundary constraint sequence $Z = \langle 1, \ldots, n+m \rangle - N$, and $\hat{x}$ row $N = w$ and and $\hat{x}$ row $Z = 0$ and $A\hat{x} = b$. We also have the reduced gradient covector d where $d^{\mathrm{T}} = (c^{\mathrm{T}} \text{ col } N)B^{-1}A - c^{\mathrm{T}}$ and where B is the non-singular $m \times m$ matrix A col N. And $z = c^{\mathrm{T}}\hat{x}$. Recall $w = B^{-1}b$.

Now $d^{\mathrm{T}}\hat{x} = (c^{\mathrm{T}} \text{ col } N)B^{-1}A\hat{x} - c^{\mathrm{T}}\hat{x}$. We know $d \geq 0$ since this is the termination condition in the simplex algorithm for the discovery of a finite optimal point. Also d row $N = 0$. Moreover, $c^{\mathrm{T}}\hat{x} = (c^{\mathrm{T}} \text{ col } N)B^{-1}b$ since $-d^{\mathrm{T}}\hat{x} = 0$ and $A\hat{x} = b$. Alternatively,

$$(c^{\mathrm{T}} \text{ col } N)B^{-1}b \quad = \quad (c^{\mathrm{T}} \text{ col } N)(\hat{x} \text{ row } N)$$

$$\begin{aligned} &= (c^{\mathrm{T}} \text{ col } N)(\hat{x} \text{ row } N) + (c^{\mathrm{T}} \text{ col } Z)(\hat{x} \text{ row } Z) \\ &= c^{\mathrm{T}}\hat{x}. \end{aligned}$$

Exercise 0.18: Show that $d^{\mathrm{T}}\hat{x} = 0$. Hint: look at d row N, d row Z, $\hat{x}$ row N, and $\hat{x}$ row Z.

Thus we have $c^{\mathrm{T}}\hat{x} = (c^{\mathrm{T}} \text{ col } N)B^{-1}b$. Let $\hat{y}^{\mathrm{T}} = (c^{\mathrm{T}} \text{ col } N)B^{-1}$. Then $c^{\mathrm{T}}\hat{x} = \hat{y}^{\mathrm{T}}b = b^{\mathrm{T}}\hat{y}$. So $\hat{y}$ might potentially be an optimal point of our dual problem: [determine $y \in (\mathcal{R}^m)^\top$ to minimize $b^{\mathrm{T}}y$ subject to $(A^o)^{\mathrm{T}}y \geq c$ and $y \geq 0$ where A^o is an $m \times n$ matrix, and $c \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$]. Indeed $\hat{y}$ will be an optimal point if $\hat{y}$ is a feasible point of our dual problem; this is a consequence of the strong duality theorem.

But, we have the termination condition: $0 \leq d^{\mathrm{T}} = \hat{y}^{\mathrm{T}}A - c^{\mathrm{T}}$, and $A = [A^o \quad I]$ and $c = \begin{bmatrix} c^o \\ 0 \end{bmatrix}$, so $0 \leq d^{\mathrm{T}} = [\hat{y}^{\mathrm{T}}A^o - c^o \quad \hat{y}^{\mathrm{T}}]$. Thus $0 \leq \hat{y}^{\mathrm{T}}A^o - (c^o)^{\mathrm{T}}$ and $0 \leq \hat{y}^{\mathrm{T}}$, *i.e.*, $(A^o)^{\mathrm{T}}\hat{y} \geq c^o$ and $\hat{y} \geq 0$. Thus $\hat{y}$ is a feasible point of our dual problem, so $\hat{y}$ is in fact an optimal point of our dual problem.

Note $\hat{y}^{\mathrm{T}} = d^{\mathrm{T}}$ col $(n+1) : (n+m)$. Thus our terminal tableau exhibits not only $\hat{x}$ (as $\hat{x}$ row $N = w$ and $\hat{x}$ row $Z = 0$,) but also $\hat{y}$ (as $\hat{y}^{\mathrm{T}} = d^{\mathrm{T}}$ col $(n+1) : (n+m)$,) and the common value of the primal and dual objective functions is given by $z = c^{\mathrm{T}}\hat{x} = b^{\mathrm{T}}\hat{y}$.

Exercise 0.19: What are the values of the n slack variables associated with the dual problem?

Also, we can obtain a constructive proof of the strong duality theorem from the above specification of $\hat{y}$. Given that $\hat{x}$ is a finite optimal point of our primal problem and $\hat{y}$ is a feasible point of our dual problem and $b^{\mathrm{T}}\hat{y} = c^{\mathrm{T}}\hat{x}$, we may conclude that $\hat{y}$ is an optimal point of our dual problem. Since the weak duality theorem requires $\min_{y \in Q} b^{\mathrm{T}}y \geq \max_{x \in P} c^{\mathrm{T}}x = c^{\mathrm{T}}\hat{x}$, and $\min_{y \in Q} b^{\mathrm{T}}y \leq b^{\mathrm{T}}\hat{y} = c^{\mathrm{T}}\hat{x}$, we see that $\min_{y \in Q} b^{\mathrm{T}}y \geq c^{\mathrm{T}}\hat{x} = b^{\mathrm{T}}\hat{y} \geq \min_{y \in Q} b^{\mathrm{T}}y$ so $\min_{y \in Q} b^{\mathrm{T}}y = b^{\mathrm{T}}\hat{y}$. And when $\hat{x}$ is a finite optimal point of our primal problem, we see that $\hat{y} = B^{-1\mathrm{T}}(c \text{ row } N)$ is a feasible point of our dual problem with the associated objective function value $b^{\mathrm{T}}\hat{y} = c^{\mathrm{T}}\hat{x}$, so $\hat{y}$ is an optimal point of our dual problem.

0.1.3 The Geometry of Duality

Recall our standard primal linear programming problem:

[determine $x \in (\mathcal{R}^n)^\top$ to maximize $c^{\mathrm{T}}x$ subject to $A^o x \leq b$ and $x \geq 0$ where A^o is an $m \times n$ matrix, and $c \in (\mathcal{R}^n)^\top$, and $b \in (\mathcal{R}^m)^\top$].

We may introduce m slack variables $t_1, \ldots, t_m$ defined in terms of x by $t = b - A^o x$, necessarily with $t \geq 0$.

Now consider the following slack-variable formulation of our primal linear programming problem. (Recall we assume $n \geq 1$ and $m \geq 1$.)

[Determine $x \in (\mathcal{R}^{n+m})^\top$ to maximize $c^{\mathrm{T}}x$ subject to $Ax = b$ and $x \geq 0$ where A is an $m \times (n+m)$ rank m matrix, and $c \in (\mathcal{R}^{n+m})^\top$, and $b \in (\mathcal{R}^m)^\top$].

Here we have *redefined* x to be an $(n+m)$-covector with the m added slack-variable components $x_{n+1} = t_1$, …, $x_{n+m} = t_m$, and we have *redefined* c to be an $(n+m)$-covector by extending it with c row $(n+1) : (n+m) = 0$, and we define A to be the $m \times (n+m)$ matrix $[A^o \quad I]$. The corresponding polyhedron is $P^* = \{x \in (\mathcal{R}^{n+m})^\top \mid Ax \leq b \text{ and } x \geq 0\}$. (In fact the following results do not depend on the matrix A being of the form $[A^o \quad I]$; the following is valid when A is an arbitrary $m \times (n+m)$ rank m matrix.)

The corresponding dual linear programming problem is:

[determine $y \in (\mathcal{R}^m)^\top$ to minimize $b^\mathrm{T} y$ subject to $A^\mathrm{T} y \geq c$ where A is an $m \times (n+m)$ rank m matrix, and $c \in (\mathcal{R}^{n+m})^\top$, and $b \in (\mathcal{R}^m)^\top$].

Let $k = n+m$ so that A is an $m \times k$ matrix. The slack-variable primal problem above can be written as:

[determine $x \in (\mathcal{R}^k)^\top$ to maximize $c^\mathrm{T} x$ subject to $A(x-d)$ and $x \geq 0$ where $d \in (\mathcal{R}^k)^\top$ and satisfies $Ad = b$].

Note when $A = [A^o \quad I]$, we can choose $d = \begin{bmatrix} 0 \\ b \end{bmatrix}$. When there is no covector d such that $Ad = b$, our slack-variable primal problem has no solution; otherwise the m equations $Ax = Ad$ specify a $(k-m)$-dimensional flat F in $(\mathcal{R}^k)^\top$. Even if d exists such that $Ad = b$, our problem will only have a solution if d can be chosen such that $d \geq 0$. Let us assume our slack-variable primal problem has a feasible point, *i.e.*, the corresponding polyhedron $P^* = F \cap O_k^+$ is non-empty.

Prabhakaran [Pra02] gives a way to "lift" the dual problem from $(\mathcal{R}^m)^\top$ to $(\mathcal{R}^k)^\top$. The basic idea is that the m-dimensional polyhedron $Q^* = \{y \in (\mathcal{R}^m)^\top \mid A^\mathrm{T} y \geq c\}$ of the dual problem can be mapped one-to-one *into* an m-dimensional flat G in $(\mathcal{R}^k)^\top$ orthogonal to the $(k-m)$-dimensional flat $F \subseteq (\mathcal{R}^k)^\top$ where $F = \{x \in (\mathcal{R}^k)^\top \mid Ax = b\} = \{x \in (\mathcal{R}^k)^\top \mid Ax = Ad\}$ is the flat associated with the slack-variable formulation of our primal linear programming problem.

First introduce k slack variables $s_1, \ldots, s_k$ to write the dual problem as:

[determine $s \in (\mathcal{R}^k)^\top$ and $y \in (\mathcal{R}^m)^\top$ to minimize $b^\mathrm{T} y$ subject to $s = A^\mathrm{T} y - c$ and $s \geq 0$].

Note that the constraints $s = A^\mathrm{T} y - c$ define s in terms of $y \in (\mathcal{R}^m)^\top$ and thus confine s to an m-dimensional flat in $(\mathcal{R}^k)^\top$. The map $y \to A^\mathrm{T} y - c$ is the map that carries Q^* into the flat G orthogonal to F. When there are no covectors s and y such that $A^\mathrm{T} y = c + s$ with $s \geq 0$, our dual problem has no solution. Let us assume our dual problem has a feasible point, *i.e.*, the corresponding polyhedron Q^* is non-empty.

> **Exercise 0.20:** Show that if our primal slack-variable form problem has a feasible point and our dual slack-variable form problem also has a feasible point, then both our primal and our dual problem have finite optimal points.

Now recall $k - m = n$ and define $A^\perp$ as a $n \times k$ rank n matrix whose rows form a basis of $rowspace(A)^\perp$. Note $A^\perp A^\mathrm{T} = O_{n \times m}$. (When $A = [A^o \quad I_{m\times m}]$, $A^\perp$ can be taken as $[I_{n\times n} \quad -(A^o)^\mathrm{T}]$.) The dual problem constraint $s + c = A^\mathrm{T} y$ can now be transformed to yield $A^\perp(s+c) = 0$ because $A^\perp(s+c) = A^\perp A^\mathrm{T} y$ and $A^\perp A^\mathrm{T} = O_{n\times m}$ and $A^\perp$ has rank n.

Exercise 0.21: Show that when $Q^* \neq \emptyset$, $s+c \in rowspace(A)^\top$. And then, since $rank(A) = m$, there exists a unique m-covector y such that $A^\mathrm{T} y = s + c$.

Exercise 0.22: Show that $A^\mathrm{T} Q^* = \{A^\mathrm{T} y \in (\mathcal{R}^k)^\top \mid A^\mathrm{T} y \geq c\} = \{s + c \in (\mathcal{R}^k)^\top \mid s = A^\mathrm{T} y - c \text{ and } s \geq 0\} = \{s + c \in (\mathcal{R}^k)^\top \mid A^\perp(s + c) = 0 \text{ and } s \geq 0\}$.

Also the dual problem objective function $b^\mathrm{T} y$ can be written as $d^\mathrm{T}(c + s)$ since we have assumed there is a covector $d \in (\mathcal{R}^k)^\top$ such that $b^\mathrm{T} = d^\mathrm{T} A^\mathrm{T}$, so $b^\mathrm{T} y = d^\mathrm{T} A^\mathrm{T} y$, and we have $c + s = A^\mathrm{T} y$. And thus $b^\mathrm{T} y = d^\mathrm{T}(c + s)$ and we see that choosing y to minimize $b^\mathrm{T} y$ is equivalent to choosing s to minimize $d^\mathrm{T} s$.

Thus our dual problem is transformable to:

[determine $s \in (\mathcal{R}^k)^\top$ to minimize $d^\mathrm{T} s$ subject to $A^\perp(s + c) = 0$ and $s \geq 0$].

The n equations $A^\perp s = -A^\perp c$ specify an m-dimensional flat G in $(\mathcal{R}^k)^\top$ where $G = \{s \in (\mathcal{R}^k)^\top \mid A^\perp s = -A^\perp c\}$. Recall $F = \{x \in (\mathcal{R}^k)^\top \mid Ax = Ad\}$. Note $-c \in G$ and $d \in F$. When our slack-variable primal and dual problems both have feasible points, we have $ldim(F) = k - m = n$ and $ldim(G) = m$, and in fact $F \perp G$.

Exercise 0.23: Show that $G = flat(A^\mathrm{T} Q^*) - c$.

Since the rows of $A^\perp$ are normal to the rows of A, the $(\mathcal{R}^k)^\top$-flats F and G are orthogonal, and since the subspaces $F - d$ and $G + c$ are likewise orthogonal and their dimensions sum to k, they are complementary to one-another in $(\mathcal{R}^k)^\top$, and $(F - d) \cap (G + c) = \{0\}$. Hence there is a unique $(\mathcal{R}^k)^\top$-point a such that $F \cap G = \{a\}$. The point a is computable as the unique solution of the k equations $\begin{bmatrix} A \\ A^\perp \end{bmatrix} a = \begin{bmatrix} Ad \\ -A^\perp c \end{bmatrix}$. (Note $\begin{bmatrix} A \\ A^\perp \end{bmatrix}$ is a $k \times k$ non-singular matrix.)

Exercise 0.24: Is the point a necessarily a feasible point of either the slack-varible form of our primal linear programming problem or of the slack-varible form of our dual linear programming problem? That is, does $a \geq 0$ necessarily hold?

Now, $A(x - d) = A((x - a) - (d - a)) = A(x - a)$ since $A(d - a) = 0$, (because $a \in F$.) And $A^\perp(s+c) = A^\perp((s-a)-(-c-a)) = A^\perp(s-a)$ since $A^\perp(-c-a) = 0$, (because $a \in G$.) Moreover, the k-covector $\hat{x}$ that maximizes $c^\mathrm{T} x$ subject to $A(x - a) = 0$ and $x \geq 0$ also minimizes $a^\mathrm{T} x$ subject to the same constraints. This is because $c^\mathrm{T} x = (c-a+a)^\mathrm{T}(x-d+d) = -a^\mathrm{T} x+(c+a)^\mathrm{T}(x-d)+(c+a)^\mathrm{T} d$, and $A(x-d) = 0$ and $A^\perp(c+a) = 0$ implies $x-d$ and $c+a$ are perpendicular, so $(c-a)^\mathrm{T}(x-d) = 0$. Thus $c^\mathrm{T} x = -a^\mathrm{T} x + (c + a)^\mathrm{T} d$, so the covector $\hat{x}$ that yields a constrained minimum of $a^\mathrm{T} x$ is the same as the covector that yields a constrained maximum of $c^\mathrm{T} x$.

In the same way, the k-covector $\hat{s}$ that minimizes $d^\mathrm{T} s$ subject to the constraints $A^\perp(s + c) = A^\perp(s - a) = 0$ and $s \geq 0$ also minimizes $a^\mathrm{T} s$ subject to the same constraints.

Exercise 0.25: Show that $d^\mathrm{T} s = a^\mathrm{T} s + (d - a)^\mathrm{T} a$.

Thus the above transformations produce the modified primal problem:

[determine $x \in (\mathcal{R}^k)^\top$ to minimize $a^\mathrm{T} x$ subject to $A(x - a) = 0$ and $x \geq 0$].

And the modified dual problem:

[determine $s \in (\mathcal{R}^k)^\top$ to minimize $a^{\mathrm{T}}s$ subject to $A^\perp(s-a) = 0$ and $s \geq 0$].

Suppose that one of our modified primal and modified dual problems has a finite optimal point. Then the strong-duality theorem implies that both have finite solutions. Note if $a \geq 0$, both our modified primal and modified dual problems have a as a feasible point. Let x and s denote feasible points of our modified primal and modified dual problems respectively. We have $x \in F \cap O_k^+$ and $s \in G \cap O_k^+$. This means $x - d \in F - d$ and $s + c \in G + c$, so $(x-d, s+c) = 0 = x^{\mathrm{T}}(s+c) - d^{\mathrm{T}}(s+c)$. And then $b^{\mathrm{T}}y = d^{\mathrm{T}}(c+s) = x^{\mathrm{T}}(c+s) \geq c^{\mathrm{T}}x$ since $x \geq 0$ and $s \geq 0$. This is the weak-duality theorem.

Exercise 0.26: Assume there is a covector $y \in (\mathcal{R}^m)^\top$ such that $A^{\mathrm{T}}y \geq c$. Given $s \in (\mathcal{R}^k)^\top$ such that $A^\perp(s-a) = 0$ and $s \geq 0$, explain how to compute the corresponding feasible point y such that $A^{\mathrm{T}}y \geq c$. Hint: the affine transformation that maps Q^* into G is given by $v \to A^{\mathrm{T}}v - c$

Solution 0.26: Recall $A^\perp a = -A^\perp c$. Thus we have $A^\perp(s+c) = 0$, so $s + c \in \mathit{rowspace}(A)^\top$, *i.e.*, there exists a unique m-covector y such that $A^{\mathrm{T}}y = s + c$. (Note $\mathit{rtnullspace}(A^\perp) = \mathit{rowspace}(A)$.) Thus $y = (A^{\mathrm{T}})^+(s+c)$. The k equations $A^{\mathrm{T}}y = s + c$ are "overdetermined", but s is defined exactly so that these equations have a (unique) solution. And since $s \geq 0$, $A^{\mathrm{T}}y = s + c$ ensures that $A^{\mathrm{T}}y \geq c$.

The modified primal and modified dual problems obtained above have a geometrical interpretation where the solution to our modified primal problem, if it exists, is an $(\mathcal{R}^k)^\top$-covector $\hat{x}$ such that $\hat{x}$ is an extreme point of the $(k-m)$-dimensional degenerate polyhedron $F \cap O_k^+$ lying in the intersection of some distinguished collection of m (and possibly more) of the hyperplanes defining the boundary of O_k^+, while the solution to our modified dual problem, if it exists, is an $(\mathcal{R}^k)^\top$-covector $\hat{s}$ such that $\hat{s}$ is an extreme point of the m-dimensional degenerate polyhedron $G \cap O_k^+$ lying in the intersection of the other $k-m$ hyperplanes defining the boundary of O_k^+, and possibly more such ∂O_k^+-hyperplanes as well. And the connection between duality and orthogonality is clearly seen since the flats F and G are orthogonal.

Exercise 0.27: Show that $(\hat{x} - a, \hat{s} - a) = 0$.

If we "start" at the point a common to F and G, and move in F to the extreme point $\hat{x}$ and also move in G to the extreme point $\hat{s}$, then as we move, the distance between the two moving points gets steadily larger until we reach the "semi-corners" of the boundary of O_k^+ where the optimal points $\hat{x}$ and $\hat{s}$ are found. Note that the motion of these two points need not be consistent with the two gradient directions obtained from the common gradient direction a projected in F and in G; this is because a need not lie in O_k^+, and if so, both points will move from outside O_k^+ to the boundary of O_k^+.

Exercise 0.28: Under what conditions does the angle between these two moving points also grow steadily larger, or at least not diminish, as we move until the maximum possible separating distance is attained at the two optimal points?

Exercise 0.29: Let $\alpha, \beta \in \mathcal{R}^+$ and let $p, q, a \in \mathcal{R}^2$ with $p \neq 0$ and $q \neq 0$ and $p \neq q$. Let C denote the cone $\mathit{cone}_0(p, q)$. Show that when $a \in C \cup (-C)$, the "full"-angle in $[0, 2\pi)$ between the two vectors $\alpha p + a$ and $\beta q + a$ increases as α and β increase, unless $a = 0$. (We need to admit full-angles greater than π for this to be generally true for $a \in C$.)

References

[Pra02] Manoj Prabhakaran. Visualizing LP duality. *Princeton CS Dept.*, 2002.