Chapter 2

Zachary Campbell

March 20, 2018

In this chapter we will take a step back and look at primal-dual algorithms in more generality. The goal will be to describe a method of solving a set of primal-dual algorithms for *network design problems*. In network design problems we are given a graph $G = (V, E)$ and a cost c_{ij} for each edge $(i, j) \in E$, and the goal is to find a minimum/maximum-cost subset $E^{'} \subset E$ that satisfies some criteria. Our maximum-cost matching is an example. There are other common examples that we will explore later on, but for now it suffices to just think of these problems as choosing subsets of our graph according to some stipulations. Throughout, we will be looking at undirected graphs.

1 The Classical Primal-Dual Method

We begin by looking at what's known as the "classical" primal-dual method, which is concerned with linear programs for polynomial-time solvable optimization problems. This will allow us to build up a framework for a more general primal-dual method that we can use for approximation algorithms - i.e. for problems that are known to be NP-hard. **(Not sure if this is within the scope of the thesis – discuss with Jim)

Let's consider the linear program

$$\text{minimize } \mathbf{c}^T\mathbf{x} \tag{1}$$

$$\text{subject to } \; A\mathbf{x} \geq \mathbf{b} \tag{2}$$

$$\mathbf{x} \geq 0 \tag{3}$$

and its dual

$$\text{maximize } \mathbf{b}^T\mathbf{y} \tag{4}$$

$$\text{subject to } A^T\mathbf{y} \leq \mathbf{c} \tag{5}$$

$$\mathbf{y} \geq 0. \tag{6}$$

We first define a concept that we will use throughout the rest of this thesis.

Definition (Complementary slackness). Given two linear programs in the form above, the *primal complementary slackness conditions* are the conditions which, given primal solution x, are necessary for a dual solution y:

$$x_j > 0 \implies A^j\mathbf{y} = c_j,$$

where A^j is the jth column of A. Similarly, the *dual complementary slackness conditions* are the conditions which, given dual solution y, are necessary for a primal solution x:

$$y_i > 0 \implies A_i\mathbf{x} = b_i,$$

where A_i is the ith row of A. Together, these conditions give us necessary and sufficient conditions for solving the primal-dual system, which we prove. The (maximization) primal slackness variables are given by $\mathbf{s} = \mathbf{b} - A\mathbf{x}$. The dual slackness variables are given by $\mathbf{t} = A^T\mathbf{y} - \mathbf{c}$.

Theorem . Let $\mathbf{x}$ be a primal feasible solution, and $\mathbf{y}$ a dual feasible solution. Let $\mathbf{s}$ and $\mathbf{t}$ be the corresponding slackness variables. Then $\mathbf{x}$ and $\mathbf{y}$ are optimal solutions if and only if the following two conditions hold:

$$x_j t_j = 0 \quad \forall j \tag{7}$$

$$y_i s_i = 0 \quad \forall i. \tag{8}$$

Proof. Let $u_i = y_i s_i$ and $v_j = x_j t_j$, and $\mathbf{u} = \sum_i u_i$, $\mathbf{v} = \sum_j v_j$. Then $\mathbf{u} = 0$ and $\mathbf{v} = 0$ if and only if (7) and (8) hold. Also,

$$\begin{aligned}\mathbf{u}+\mathbf{v} &= \sum y_i s_i + \sum x_j t_j \\ &= \sum y_i(b_i - A_i x_i) + \sum x_j(A_j^T y_j - c_j) \\ &= \sum b_i y_i - \sum c_j x_j,\end{aligned}$$

so we get that $c^T\mathbf{x} = b^T\mathbf{y}$ if and only if $u + v = 0$, which proves the statement. □

The general "tug-of-war" between the primal and dual suggests an economic interpretation of slackness conditions. We can think of our primal (maximization) problem as concerned with profit given some constraints on resources, i.e. a resource allocation problem. The dual can be interpreted as a valuation of the resources – it tells us the availability of a resource, and its price. So if we have optimal $\mathbf{x}$ and $\mathbf{y}$, we can interpret slackness as follows: if there is slack in a constrained primal resource i ($s_i > 0$), then additional units of that resource must have no value ($y_i = 0$); if there is slack in the dual price constraint ($t_j > 0$) there must be a shortage of that resource ($x_j = 0$).

We now give an example of complementary slackness in action. Let's look back to our maximum weight matching problem. Recall the primal linear program for maximum-weight matching:

$$\text{maximize} \quad \sum_{i,j} c_{ij} x_{ij} \tag{9}$$

$$\text{subject to} \quad \sum_j x_{ij} \leq 1, \qquad \forall i \in L, \tag{10}$$

$$\sum_i x_{ij} \leq 1, \qquad \forall j \in R, \tag{11}$$

$$x_{ij} \in \{0, 1\}. \tag{12}$$

and its dual

$$\text{minimize} \quad \sum_i u_i + \sum_j v_j \tag{13}$$

$$\text{subject to} \quad u_i + v_j \geq c_{ij} \qquad \forall i \in L,\ j \in R, \tag{14}$$

$$u_i, v_j \in \{0, 1\}. \tag{15}$$

The format here is a little different, since our primal is a maximization problem and the dual is a minimization, but it's easy enough to reverse the roles. It's easy to see our corresponding primal complementary slackness conditions are

$$x_{ij} > 0 \implies u_i + v_j = c_{ij}. \tag{16}$$

The dual complementary slackness conditions are

$$u_i > 0 \implies \sum_j x_{ij} = 1, \tag{17}$$

$$v_j > 0 \implies \sum_i x_{ij} = 1. \tag{18}$$

In general, the slackness conditions guide us in our algorithm – they tells us how, given a solution to one of the problems, we should augment the solution to the other. For example, the algorithm we presented for maximum-weight matching/minimum vertex cover intializes with a solution to both the primal and dual that satisfies conditions (8) and (10); the algorithm then at each step works to decrease the number of conditions in (9) that are unsatisfied, while maintaining satisfiability of (8) and (10). This method is not unique to the Hungarian algorithm. In fact, the Hungarian algorithm paved the way for this general method, which we describe presently.

Looking back at the original linear programs at the beginning of this chapter, suppose we have a dual feasible solution $\mathbf{y}$. We can then state the problem of finding a feasible primal solution $\mathbf{x}$ that obeys our complementary slackness conditions as another *restricted* linear program. Define the sets $J = \{j \mid A^j\mathbf{y} = c_j\}$ and $I = \{i \mid y_i = 0\}$. So J tells us which dual constraints (5) are tight, given the solution $\mathbf{y}$, and I tells us which y_i are 0. What we want to do is give a linear program to find a solution $\mathbf{x}$ that minimizes the "violation" of the complementary slackness conditions and the primal constraints. We will have slack variables s_i which will describe the difference between $A_i\mathbf{x}$ and b_i for $i \notin I$. We do this because we want to look at all $y_i > 0$ where the we do not have that $A_i\mathbf{x} = b_i$. So part of our objective function will be to minimize the sum of these s_i. We also want to minimize the sum over variables x_j where $j \notin J$. This is because we want to see if there are any x_j such that $A^j\mathbf{y} \neq c_j$. So we give the following restricted primal linear program:

$$\text{minimize} \sum_{i \notin I} s_i + \sum_{j \notin J} x_j \tag{19}$$

$$\text{subject to} \quad A_i\mathbf{x} \geq b_i \quad i \in I, \tag{20}$$

$$A_i\mathbf{x} - b_i = s_i \quad i \notin I, \tag{21}$$

$$\mathbf{x} \geq 0, \tag{22}$$

$$\mathbf{s} \geq 0. \tag{23}$$

Observe that if this restricted primal has a feasible solution $(\mathbf{x}, \mathbf{s})$ such that the objective function is 0, then $\mathbf{x}$ is a feasible primal solution that satisfies the complementary slackness conditions for the dual solution $\mathbf{y}$. This means that $\mathbf{x}$ and $\mathbf{y}$ are optimal primal and dual solutions. If, however, the optimal solution to this restricted primal has value greater than 0, more work is required. We can consider the dual of the restricted primal:

$$\text{maximize } \mathbf{b}^T\mathbf{y}' \tag{24}$$

$$\text{subject to} \quad A^j\mathbf{y}' \leq 0 \qquad j \in J, \tag{25}$$

$$A^j\mathbf{y}' \leq 1 \qquad j \notin J, \tag{26}$$

$$y_i' \geq -1 \quad i \notin I, \tag{27}$$

$$y_i' \geq 0 \qquad i \in I. \tag{28}$$

What we want here is to improve our dual solution. By assumption, the optimal solution to this linear program's primal is greater than 0, so we know that this dual has a solution $\mathbf{y}'$ such that $\mathbf{b}^T\mathbf{y}' > 0$. What we want is the existence of some $\epsilon > 0$ such that $\mathbf{y}'' = \mathbf{y} + \epsilon\mathbf{y}'$ is a feasible dual solution. In particular, a solution of this form will be an improvement on our original soluation $\mathbf{y}$.

[SEE P.148 Goemans, Williamson for ϵ calculation].

It's not immediately clear why reducing our original linear programs to a series of linear programs is heplful. However, note that the vector **c** has totally disappeared in the restricted primal and its dual. Recall that in the original linear program, **c** gave us the edge-costs on our graph. So this method reworks our original weighted problem into unweighted parts, which are often easier to solve. Oftentimes, it is the case that we can interpret these unweighted problems as purely combinatorial, which means that instead of actually solving the problem with linear programming, we can solve it by combinatorial methods. Using a combinatorial algorithm to find a solution **x** that obeys the complementary slackness conditions, or to find an improved dual solution **y**, is oftentimes more efficient.

Let's revisit our maximum matching linear program, and its dual. We will begin with the dual feasible solution $u_i = \max_j c_{ij}$ for all i, and $v_j = 0$ for all j. We will then try to find a corresponding primal solution that minimizes the violation of the primal constraints, and complementary slackness conditions. We will have slack variables s_i, s_j corresponding to dual slackness; for u_i, v_j that are not equal to zero, we want to minimize the distance, or *slackness* between $\sum x_{ij}$ and 1. Additionally, we want to minimize the values on x_{ij} where $u_i + v_i \neq c_{ij}$. In fact, for any edge $(i, j) \notin J$, we want the corresponding x_{ij} to be zero, since if this weren't the case we would be certainly violating our complementary slackness conditions. This corresponds to the following restricted primal

$$
\begin{aligned}
&\text{minimize} \sum_{i \in L} s_i + \sum_{j \in R} s_j && && (29)\\
&\text{subject to} \quad \sum_j x_{ij} - s_i = 1 && \forall i, && (30)\\
&\qquad\qquad\quad \sum_i x_{ij} - s_j = 1 && \forall j, && (31)\\
&\qquad\qquad\qquad\qquad x_{ij} = 0 && (i, j) \notin J, && (32)\\
&\qquad\qquad\qquad\qquad x_{ij} \geq 0 && (i, j) \in J, && (33)\\
&\qquad\qquad\qquad\qquad \mathbf{s} \geq 0. && && (34)
\end{aligned}
$$

The dual of this is given by the following linear program:

$$\text{maximize} \sum_{i \in L} u_{i'} + \sum_{j \in R} v_{j'} \tag{35}$$

$$\text{subject to} \qquad u_{i'} + v_{j'} \leq 0 \qquad (i, j) \in J, \tag{36}$$

$$u_{i'} \leq 1, \tag{37}$$

$$v_{j'} \leq \tag{38}$$