Min Weight Perfect Matching

Jamie Fuente

August 2020

1 Introduction

For our final report, we investigated the weighted matching problem on graphs. Given a weighted graph G = (V,E,w) with w $\geq$ 0, a maximum weight matching produces a matching M such that w(M) is as large as possible. Note that for an arbitrary graph G, it is always possible to reduce the matching problem to the maximum weight perfect matching problem by adding a dummy node if $|V|$ is odd and adopting the convention w((u,v)) = 0 if $(u,v) \notin E$, making G complete. First we will investigate the classical approach which solves the perfect matching problem (and thus the matching problem). Later we examine an approximation algorithm that solves the matching problem in linear time, but because it is an approximation, it cannot guarantee the resulting matching is perfect.

This problem has a wide number of applications. For example, when G is a bipartite graph, this is often referred to as the "assignment problem" as it is used to assign workers to tasks based on their qualifications. A non-bipartite example is Christofides algorithm, which uses a minimum weight matching to find a $\frac{3}{2}$ approximation for the metric traveling salesman problem in polynomial time.

2 The MWPM LP and its dual

We can formulate the minimum weight perfect matching problem as an ILP as follows (ILP 1):

$$min \sum_{e \in E} w_e x_e$$

$$s.t. \sum_{e \in \partial(v)} x_e = 1 \qquad \forall v \in V$$

$$x_e \geq 0 \quad x_e \in \{0,1\} \qquad \forall e \in E$$

In the case that G is a bipartite graph, it is well known that the binary constraints can be dropped (see section 8.6 of Bondy and Murty). Unfortunately when G is not bipartite, the existence of blossoms can cause problems. We can augment the above LP by adding constraints that eliminate the problems caused by blossoms. If M is some matching, recall that a blossom occurs when a cycle of 2k +1 nodes (k > 0) in G contains k edges of M. In the blossom algorithm, if we find an augmenting path that contains a blossom and attempt to take its symmetric difference with M, the root of the blossom will be doubly matched. We'll address this problem as follows: Let B be the set of odd subsets of V with more than 1 node, so if $S \in B$, then $S \subseteq V, |S| = 2k+1$, and $k > 0$. Now, create an LP1 by augmenting the LP relaxation of ILP 1 with constraints $\sum_{e \in E(S)} x_e \leq \frac{|S|-1}{2}, \quad \forall S \in B$. These constraints eliminate the possibility of a doubly matched blossom root in the augmented matching. Bellow is the full LP with blossom constraints (LP1) and its dual (DLP 1)

$$
\begin{aligned}
& min \sum_{e \in E} w_e x_e \\
s.t. & \sum_{e \in \partial(v)} x_e = 1 \qquad \forall v \in V \\
& \sum_{e \in E(S)} x_e \leq \frac{|S|-1}{2}, \quad \forall S \in B \\
& x_e \geq 0 \qquad \forall e \in E
\end{aligned}
$$

$$
\begin{aligned}
& max \sum_{v \in V} 1 * y_v + \sum_{S \in B} \frac{|S|-1}{2} * y_s \\
s.t. \quad & y_u + y_v + \sum_{u,v \in S} y_s \leq w_{(u,v)} \quad \forall (u,v) \in E \\
& y_s \leq 0 \qquad \forall S \in B
\end{aligned}
$$

In 1964, Edmonds showed that adding these constraints produces an LP which is guaranteed to have an integral solution. Unfortunately there are $2^{n-1} - n$ such constraints (dual variables), so naively implementing this LP would be intractable.

Luckily, Edmonds also showed that there is an algorithm based on the primal-dual simplex method that can solve the problem in $O(n^4)$ operations (and can be optimized to run in $O(n^3)$).

3 Complementary Slackness

Recall that the complementary slackness and tightness conditions of linear programs require that x is an optimal solution to LP1 and y is an optimal solution to DLP1 if and only if

$$\left(\frac{|S|-1}{2} - \sum_{e \in E(S)} x_e\right) y_s = 0 \tag{1}$$

$$\left(1 - \sum_{e \in \partial(v)} x_e\right) y_v = 0 \tag{2}$$

$$\left(y_u + y_v + \sum_{u,v \in S} y_{(u,v)} - w_{(u,v)}\right) x_{(u,v)} = 0 \tag{3}$$

Note that (2) is required simply for primal feasibility. Let y be some basic feasible solution to the DLP1. We will attempt to check whether y is an optimal solution to DLP1 by doing the following: Let J be the set of edges $(u, v) \in E$ s.t. $y_u + y_v + \sum_{u,v \in S} y_s < w_{(u,v)}$ and let K be the set of $S \in B$ s.t. $y_s < 0$. If complementary slackness is to be achieved, it must be the case that $x_e = 0 \quad \forall e \in J$, and $\sum_{e \in E(S)} x_e = \frac{|S|-1}{2} \quad \forall S \in K$. Finally, because $\sum_{e \in \partial(v)} x_e = 1 \quad \forall v \in V$, every vertex must be matched. We construct the the residual primal (LP2) and residual dual (DLP2) as follows:

$$\begin{gathered}
min \sum_{v \in V} \bar{x_v} + 2 \sum_{S \in K} \bar{x_s} \\
s.t. \sum_{e \in \partial(v)} \bar{x_e} + \bar{x_v} = 1 \qquad \forall v \in V \\
\sum_{e \in E(S)} \bar{x_e} + \bar{x_s} = \frac{|S|-1}{2}, \quad \forall S \in K \\
\bar{x_v}, \bar{x_s}, \bar{x_e} \geq 0 \quad \forall v \in V, s \in B, e \in E \\
\bar{x_e} = 0 \quad \forall e \in J
\end{gathered}$$

$$\begin{gathered}
max \sum_{v \in V} \bar{y_v} + \sum_{S \in B} \frac{|S|-1}{2} * \bar{y_s} \\
s.t. \quad \bar{y_u} + \bar{y_v} + \sum_{u,v \in S} \bar{y_s} \leq 0 \quad \forall (u, v) \in E \setminus J \\
\bar{y_v} \leq 1, \quad \forall v \in V \\
\bar{y_s} \leq 2 \quad \forall S \in K \\
\bar{y_s} \leq 0 \quad \forall S \in B \setminus K
\end{gathered}$$

Note that weights are no longer considered in the residual problem. To check that y is optimal, we require for the optimal objective value of LP2 (and hence of DLP2 by strong duality) to be 0. This problem looks strikingly familiar to our original LP1. In fact, it is merely a cardinally matching problem on a subgraph of G! For the objective of LP2 to be zero and y to be optimal, we need only find a perfect matching in the graph $(V, E \setminus J)$. However there is one more requirement; All $S \in K$ must be a blossom in $(V, E \setminus J)$. We therefore reduce the problem to finding a maximum cardinally matching in G^*, which is $(V, E \setminus J)$ with each blossom $S \in K$ contracted to a single node.

4 Proof of correctness

We proceed to prove correctness, as well as the following claim for the start of each iteration:
1) $x_e \in \{0, 1\} \quad \forall e \in E$ and these give a valid matching on $(V, E \setminus J)$
2) If $S \in K, then \quad \sum_{e \in E(S)} x_e = \frac{|S|-1}{2}$
3) For any $S_1, S_2 \in K$ with $S_1 \cap S_2 \neq \emptyset$, we have $S_1 \subseteq S_2$ or $S_2 \subseteq S_1$

We proceed by induction on the number of iterations.
Base case: iteration 1
At initialization we have $x_e = 0 \quad \forall e \in E$, which is the trivial matching and satisfies 2) and 3) vacuously. We also initialize $y_v = \frac{1}{2} min\{w(e) | e \in E\} \quad \forall v \in V$ and $y_s = 0 \quad \forall S \in B$.
Inductive step: Assume the claim holds for n = 0, 1, ... , k:
Suppose we start iteration k. By IH, the current values of x_e are binary and give a valid matching on $(V, E \setminus J)$. By b) and c), the maximal elements of K are valid blossoms of $(V, E \setminus J)$ and shirking them to produce G^* is an appropriate operation by the correctness of Edmond's blossom algorithm. At this stage, we can continue to run Edmond's blossom algorithm to find a maximum cardinality matching on G^*. If the blossom algorithm ever finds a blossom β of G^* that contains a blossom node v_s for some $S \in K$ and shrinks β to create v_β, simply expand v_s in v_β to maintain the property that all blossoms are maximal subsets of V in B.

Now, it is easy to see that the number of unmatched nodes in a matching on G^* is equal to the number of unmatched nodes in a matching on $(V, E \setminus J)$ under the restriction that if $S \in K$, then $\quad \sum_{e \in E(S)} x_e = \frac{|S|-1}{2}$. This is because if S is a blossom of on $(V, E \setminus J)$, there is exactly one node v of S that is not matched to another node of S. If v is matched, then v_s after shrinking will be matched to v's partner. If v is unmatched, v_s will be unmatched, so the number of unmatched nodes is the same. Going in the opposite direction, if v_s is unmatched, after expanding it will only have k matches and 2k + 1 nodes, so there is still one unmatched node. If v_s is matched to u in G^*, it can be expanded and then the edges of S can be arranged such that u's neighbor in S is the root of S. In both cases the number of unmatched nodes remains the same. Now let M be the maximum cardinality matching on G^*.

Let p be the objective value of LP2 at $\bar{x}^*$ which corresponds to the matching M. The objective value will be the number of nodes unmatched in $(V, E \setminus J)$. If we can show that there is a feasible solution to DLP2 with the same objective value, we will be done. Let G^{**} be the graph return after Edmond's blossom algorithm finds a maximum cardinality matching (It may contract blossoms). Now let O be the set of nodes of G^{**} (which may or may not be true nodes of V) that are reachable from an unmatched node on an even length alternating path. Now let I be the nodes of G^{**} that are matched to nodes in O. Since the only nodes in O that do not have a corresponding node in I are the unmatched nodes themselves (reachable by a path of length 0), $|O| - |I| = p$. First we will let B_O be the set of all $S \in B$ s.t. S is a blossom of G^* and therefore shrunk to become a node of O. We let B_I be the set of all blossoms of G^* that shrunk to become nodes of I. Finally, starting with O we will create O^* by repeatedly expanding all blossom in O^* while keeping the vertices of the blossom in O^* until $O^* \subseteq V$. We will do the same to create I^*. Then for all $v \in O^*$ set $\bar{y_v} = 1$, for all $v \in I^*$ set $\bar{y_v} = -1$, for all $s \in B_o$ set $\bar{y_s} = -2$, and for all $s \in B_I$ set $\bar{y_s} = 2$. Set all other DLP2 variables to 0.

Now the objective value of DLP2 is $|O^*| - |I^*| + \sum_{S \in B_I}(|S|-1) - \sum_{S \in B_o}(|S| - 1) = |O| - |I| = p$.

The only thing left to check is whether this solution is feasible. Since we selected our values to be feasible variable values, we need only check that the constraint corresponding to edges is satisfied. Let (u,v) $\in E \setminus J$. If (u,v) is an edge of some blossom in B_I, then both u and v are in I^*, so the 2 value of y_s will be offset and the constraint is satisfied. Now suppose $v \in O^*$. If $e \in M$ then $u \in I^*$, so the value of y_v is offset. If $u \in O^*$ then (u,v) must be in a blossom with root O^*, so there exists some $S \in B$ with $(u, v) \in S$ and $\bar{y_s} = -2$, hence both values are offset. If $u \notin O^*, v \notin I^*$ then we have found an augmenting path, contradicting that M is optimal. It follows that we have solved the LP2 optimally by finding a maximum cardinality matching. All that is left to do is set $x_e = \bar{x_e} \quad \forall e \in E$ and augment the variables of DLP1 as specified by the primal-dual simplex algorithm. One notes that the only y_s variables that may move from 0 are those that correspond to blossoms of G^* with respect to M. Since M is optimal, the only sets S added to K satisfy the blossom constraint with equality. By construction, $\bar{x_e}$ gives a valid matching on G with $\sum_{e \in E(S)} \bar{x_e} + \bar{x_s} = \frac{|S|-1}{2}, \quad \forall S \in K$.
So both 1) and 2) are satisfied. Lastly, if S is a blossom of G^* with respect to M and $S \cap T \neq \emptyset$, since no two blossoms of G^* can intersect each other, the blossom node v_T must have been absorbed by S. So T was already shrunk and $T \subseteq S$, giving 3).

5 References

Edmonds, J. 1965a. Maximum matching and a polyhedron with 0, 1-vertices. J. Res. Nat. Bur. Stand. Sect. B 69B, 125–130. J. Edmonds, "Paths, trees and flowers", Canadian J. Math. 17, 449-467 (1965).

Mucha, M.; Sankowski, P. (2004), "Maximum Matchings via Gaussian Elimination", Proc. 45th IEEE Symp. Foundations of Computer Science, pp. 248–255

Papadimitriou, C. H., and Steiglitz, K. (1982),Combinatorial Optimization: Algorithms and Complexity, Prentice-Hall, Englewood Cliffs, NJ.

Vijay V. Vazirani 2013, A Simplification of the MV Matching Algorithm and its Proof. arXiv:1210.4594

C. Witzgall and C. T. Zahn, Jr, , Modification of Edmonds' Maximum Matching Algorithm. JOURNAL OF RESEARCH of the National Bureau of Standards-B. Mathematics and Mathematical Physics Vol. 69B, Nos. 1 and 2, January-June 1965 Modification of Edmonds' Maximum Matching Algorithm