Homework 2 Solutions

Exercise 1. Maximum Bipartite Subgraphs (*30 pts*)

The *maximum bipartite subgraph* problem is defined as follows. Given an undirected graph $G = (V, E)$, the goal is to find a subset of the edges $S \subseteq E$ such that (V, S) is bipartite and $|S|$ is maximized. Consider the following linear programming relaxation:

$$\max \sum_{ij \in E} x_{ij}$$
$$\sum_{ij \in C} x_{ij} \leq |C| - 1, \quad \text{for all odd cycles } C \in E,$$
$$0 \leq x_{ij} \leq 1, \quad \text{for all edges } ij \text{ in } E. \qquad (P_{bipartite})$$

a) (*15 pts*) Give a polynomial-time separation oracle for $(P_{bipartite})$.

Solution: Let $\mathbf{x}^*$ be an optimal solution for $(P_{bipartite})$, and let $y_{ij} = 1 - x^*_{ij}$. Then for a cycle C, we have:

$$\sum_{ij \in C} x_{ij} \leq |C| - 1 \iff \sum_{ij \in C} y_{ij} \geq 1.$$

Given a graph with nonnegative edge weights $\{y_{ij}\}$, we want to determine if every odd cycle has weight at least 1, i.e. if the above constraint is satisfied. To do this, we create a new bipartite graph G' with vertex set $V_L = V_R = V$. For each edge $ij \in E$, we add the edge from i_L to j_R and from j_L to i_R. There is an odd cycle in G containing vertex i iff there is a path from i_L to i_R in G'. So for each $v \in V$, we can find the shortest path from v_L to v_R in G' and if these paths all have value at least 1, then $\mathbf{x}^*$ is feasible. Otherwise, we have found an odd cycle corresponding to a violated constraint.

b) (*5 pts*) Find a graph that exhibits an integrality of at least $\frac{10}{9}$. (The graph should contain at most ten vertices.)

Solution: Consider K_5. By case analysis, we can see that the maximum cut has six edges. Setting $y_{ij} = \frac{2}{3}$ for each edge $ij \in E$ yields a feasible solution with objective value $\frac{20}{3}$ (since there are ten edges). So the integrality gap is at least:

$$\frac{\frac{20}{3}}{6} = \frac{10}{9}.$$

c) (*10 pts*) There are graphs for which the polytope $(P_{bipartite})$ is integer. Such graphs are called *weakly bipartite*. Suppose G is a weakly bipartite graph. Prove that there is a partition of the vertices of G such that every triangle in G is partitioned by the cut. In other words, there exists a subset $U \subset V$ such that for every triangle t in G, $E(U, V \setminus U) \cap t \neq \emptyset$.

Solution: Set $y_{ij} = \frac{2}{3}$ for each edge $ij \in E$. Since the smallest possible odd cycle in G is a triangle (i.e. 3-cycle), this solution satisfies all odd cycle constraints. In other words, for every odd cycle C, we have:

$$\sum_{ij \in C} y_{ij} \; = \; \frac{2}{3}|C| \; \leq \; |C| - 1.$$

So $\mathbf{y}$ is a feasible solution for $(P_{bipartite})$.

We can decompose $\mathbf{y}$ into a convex combination of extreme points of $(P_{bipartite})$ (e.g. via Carathéodory's Theorem; Exercise 2, TD 6). Since, by assumption, $(P_{bipartite})$ is an integer polytope for G, each of these extreme points corresponds to a bipartite subgraph. We refer to this set of bipartite subgraphs as $\mathcal{S} = \{S_1, S_2, \ldots, S_\ell\}$, where $S_i \subseteq E$. For each triangle, a bipartite subgraph S_i contains either zero or two edges from the triangle. Our goal is to show that each bipartite subgraph in the decomposition contains exactly two edges from each triangle. But maybe a bipartite subgraph contains zero edges from a triangle. However, recall that the solution $\mathbf{y}$ has value 2 on each triangle and $\mathbf{y}$ can be expressed as a convex combination of the subgraphs in $\mathcal{S}$:

$$\mathbf{y} = \sum_{i=1}^{\ell} \lambda_i S_i, \quad \text{where } \lambda_i > 0.$$

Thus, if any S_i contains zero edges from some triangle, another S_j must contain *more* than two edges from some triangle, which is impossible. So we conclude that every bipartite subgraph $S_i \in \mathcal{S}$ contains exactly two edges from each triangle in G.

Exercise 2. Randomized Rounding (Problem 5.8 from Williamson-Shmoys) (*20 pts*)

Consider a variation of the maximum satisfiability problem in which all variables occur positively in each clause. Each clause has a nonnegative weight w_j and there is an additional nonnegative weight $v_i \geq 0$ for each boolean variable x_i. The goal is now to set the boolean variables to maximize the total weight of the satisfied clauses plus the total weight of variables set to be false.

a) (*5 pts*) Give an integer programming formulation for this problem, with 0-1 variables y_i to indicate whether or not x_i is set to true.

Solution: Let z_j be a binary variable denoting whether or not the clause C_j is satisfied, and let y_j be a binary variable denoting whether or not the variable x_i is true. Let $\mathcal{C}$ and $\mathcal{X}$ denote the sets of clauses and variables, respectively, in the input formula.

$$\begin{aligned} \max \sum_{C_j \in \mathcal{C}} w_j z_j + \sum_{x_i \in \mathcal{X}} v_i(1 - y_i) & \\ \sum_{x_i \in C_j} y_i \geq z_j, \quad & \text{for all clauses } C_j \in \mathcal{C}, \\ y_i, z_j \in \{0, 1\}, \quad & \text{for all } x_i \in \mathcal{X},\ C_j \in \mathcal{C}. \end{aligned} \tag{P_{SAT+}}$$

b) (*15 pts*) Consider the linear programming relaxation of the integer program in part a). Show that a randomized rounding of this linear program in which variable x_i is set to true with probability $1 - \lambda + \lambda y_i^*$ gives a $2(\sqrt{2}-1)$-approximation algorithm for some appropriate setting of λ; note that $(2(\sqrt{2}-1) \approx .828$. In other words, show that this randomized rounding results in a solution whose *expected* value is at least $2(\sqrt{2}-1)$ times the value of an optimal solution.

Solution: Let $\{y_i^*, z_j^*\}$ denote an optimal solution for (P_{SAT+}). We set each variable $x_i \to 1$ with probability $1 - \lambda + \lambda y_i^*$

$$\begin{aligned} \Pr[\text{Clause } C_j \text{ is not satisfied}] &= \prod_{x_i \in C_j} (1 - (1 - \lambda + \lambda y_i^*)) \\ &= \prod_{x_i \in C_j} (\lambda - \lambda y_i^*)) \,. \end{aligned}$$

Since $\lambda - \lambda y_i^* \geq 0$, when $\lambda > 0$, we can apply the arithmetic-geometric mean inequality.

$$\begin{aligned} \prod_{x_i \in C_j} (\lambda - \lambda y_i^*) &\leq \left(\frac{1}{\ell_j} \sum_{x_i \in C_j} (\lambda - \lambda y_i^*) \right)^{\ell_j} \\ &= \left(\lambda - \frac{\lambda}{\ell_j} \sum_{x_i \in C_j} y_i^* \right)^{\ell_j} \\ &\leq \left(\lambda - \frac{\lambda}{\ell_j} z_j^* \right)^{\ell_j} . \end{aligned}$$

So, we have:

$$\begin{aligned} \Pr[\text{Clause } C_j \text{ is satisfied}] &= 1 - \Pr[\text{Clause } C_j \text{ is not satisfied}] \\ &\geq 1 - \left(\lambda - \frac{\lambda}{\ell_j} z_j^* \right)^{\ell_j} \\ &\geq \left[1 - \left(\lambda - \frac{\lambda}{\ell_j} \right)^{\ell_j} \right] z_j^* . \end{aligned} \tag{1}$$

We can apply Fact 4 from Lecture 8 to obtain the inequality on Line (1).

Consider the following inequality, where c is a positive integer.

$$1 - \left(\lambda - \frac{\lambda}{c}\right)^c \quad \geq \quad \lambda$$

When $c = 1$, then this inequality is true when $\lambda \leq 1$. When $c > 2$ and $\lambda \leq 1$ (and $\lambda > 0$), we have:

$$1 - \left(\lambda - \frac{\lambda}{c}\right)^c \quad \geq \quad 1 - \left(\lambda - \frac{\lambda}{2}\right)^2.$$

When $c = 2$, we can verify that the inequality holds for $\lambda = 2(\sqrt{2} - 1)$.

Now we can determine the expected value of the solution W.

$$\begin{aligned}
\mathbb{E}[W] &= \sum_{C_j \in \mathcal{C}} w_j z_j^* + \sum_{x_i \in X} v_i y_i^* \\
&= \sum_{C_j \in \mathcal{C}} w_j \cdot \Pr[\text{Clause } C_j \text{ is satisfied}] + \sum_{x_i \in \mathcal{X}} v_i \cdot \Pr[x_i \text{ is false}] \\
&\geq \lambda(\sum w_j z_j^* + \sum v_i y_i^*) \\
&= 2(\sqrt{2} - 1) \cdot OPT.
\end{aligned}$$

Exercise 3. Traveling Salesman Problem (*30 pts*)

Let $G = (V, E)$ be a complete, weighted graph. Each edge $ij \in E$ has weight $w_{ij} \geq 0$ and these weights obey the triangle inequality: $w_{ij} + w_{jk} \geq w_{ik}$ for all $i, j, k \in V$. The *traveling salesman problem* is to find a minimum weight Hamiltonian tour in G. Consider the following integer program:

$$\begin{aligned}
&\min \sum_{ij \in E} w_{ij} x_{ij} \\
&\sum_{j \in \delta(i)} x_{ij} = 2, \ \text{for all } i \in V, \\
&\sum_{\substack{ij \in E: \\ i \in S, j \notin S}} x_{ij} \geq 2, \ \text{for all cuts } S \subset V, \\
&x_{ij} \in \{0, 1\}.
\end{aligned} \tag{P_{tsp}}$$

1. (*10 pts*) Prove that the problem of finding an optimal solution for (P_{tsp}) is equivalent to that of finding a minimum weight Hamiltonian tour in G.

   Solution: Any Hamiltonian tour in G will obey all of the constraints in (P_{tsp}). Conversely, suppose we are given an optimal solution to the integer program (P_{tsp}). It must be a set of cycles, because each vertex has degree 2. To show it is a single cycle, we observe that if it is not a single cycle, then one of the cut constraints is violated.

2. (~~*10 pts*~~, *0 pts*) Relax the integrality requirement in (P_{tsp}) to obtain the constraint: $0 \leq x_{ij} \leq 1$. Let $OPT_f(G)$ denote the optimal value for the resulting linear programming relaxation.

   Let T denote a minimum weight spanning tree in G and let $w(MST(G))$ denote its weight. Let $J \subset V$ denote the vertices in V that have an odd-degree in T. Show that there is a minimum weight perfect matching on J with value at most $\frac{OPT_f(G)}{2}$.

   Solution: The solution for this question is based on the following claim[1].

[1] A proof of Claim 1 can be found in [SW90]. Since its proof is non-trivial, the problem statement should have included the claim and stated that it can be assumed to be true. Due to this oversight, the question was too hard and will not be counted.

Claim 1 *Let $G = (V, E)$ be a complete graph with nonnegative edge weights obeying triangle inequality. Define $G = (V', E(V'))$ for $V' \subset V$ to be the subgraph of G induced on the vertices of V'. Then $OPT_f(G') \leq OPT_f(G)$.*

Define $G' = (J, E(J))$ to be the subgraph of G induced on the vertex set J, which are the vertices with odd degree in a minimum weight spanning tree T. Let $\mathbf{x}^*$ denote an optimal solution for the relaxation of (P_{tsp}) on the graph G'. Let $y_{ij} = \frac{x^*_{ij}}{2}$ and note that $\mathbf{y}$ is a feasible solution for the perfect matching polytope on G'. We showed in Lecture 6 that this polytope is integer. Thus, there exists a perfect matching on the vertices in J with value at most $OPT_f(G')/2$, which by Claim 1 is at most $OPT_f(G)/2$.

3. (*10 pts*) Conclude that the traveling salesman problem has a solution with weight at most $w(MST(G)) + \frac{OPT_f(G)}{2}$.

   Solution: Let $M \subset E$ be a perfect matching on the vertices in J. If we add M to T, we obtain a graph where every vertex has even degree (i.e. an Eulerian graph) and whose cost is $w(T) + w(M) \leq w(T) + \frac{OPT_f(G)}{2}$. Since this graph is Eulerian, there is an Eulerian tour with the same cost. Using the triangle inequality, we can shortcut this Eulerian tour to obtain a Hamiltonian tour on G of no greater cost.

Exercise 4. Triangle Packing (*20 pts + 20 pts extra credit*)
Let $G = (V, E)$ be an unweighted, undirected graph. Let T be the set of all triangles in E; a triangle is a triple of edges $\{e_1, e_2, e_3\}$ such that these three edges form a cycle with 3 edges. Our goal is to find a set of edges $S \subset E$ such that $S \cap t \neq \emptyset$ for all $t \in T$ and $|S|$ is minimized. We call this the *triangle hitting set problem.* Let $OPT(G)$ denote the size of an optimal triangle hitting set.

1. (*10 pts*) Give a 3-approximation for the triangle hitting set problem.

   Solution: We can greedily find a triangle t in G, add all edges in t to the solution set S (and remove them from E), and continue. At then end of this process, the set of triangles we have found T' is a set of edge disjoint triangles, so it is a lower bound on the optimal size of a triangle hitting set. Moreover, our solution set S has size $|S| = 3 \cdot |T'| \leq 3 \cdot OPT$, so it is a 3-approximation.

2. (*10 pts*) Consider the following linear programming relaxation for the problem:

$$\begin{aligned} \min \sum_{e \in E} & x_e \\ \sum_{e \in t} x_e &\geq 1 \text{ for all triangles } t \in T, \\ x_e &\geq 0. \end{aligned} \qquad (P_{triangle})$$

   Write the dual linear program.

   Solution:

$$\begin{aligned} \max \sum_{t \in T} & y_t \\ \sum_{t : e \in t} y_t &\leq 1 \text{ for all edges } e \in E, \\ y_t &\geq 0. \end{aligned} \qquad (D_{triangle})$$

3. * (*10 pts*) Consider an optimal solution x^* for $(P_{triangle})$ with value $OPT_f(G)$. Suppose $x^*_e > 0$ for all $e \in E$. Show that in this case, $OPT(G) \leq \frac{3}{2} OPT_f(G)$. (Hint: Use complementary slackness to bound $OPT_f(G)$ in terms of $|E|$, and then find a bound on $OPT(G)$ in terms of $|E|$.)

   Solution: First, we show that there always exists a triangle hitting set $F \subset E$ with size at most $|E|/2$. Consider a maximum bipartite subgraph of G. Call this set of edges $B \subseteq E$. Then the two following facts hold for B:

(i) $|B| \geq |E|/2$, and (ii) B contains no triangles. Therefore, the following two facts hold for the complement set $F = E \setminus B$: (i) $|F| \leq |E|/2$, and (ii) F hits every triangle.

Let $\mathbf{x}^*$ be an optimal solution for $(P_{triangle})$. When $x_e^* > 0$ for all $e \in E$, we can obtain a "large" lower bound on $OPT_f(G)$. By complementary slackness, we have that for each edge $e \in E$, $\sum_{t:e \in t} y_t = 1$. In other words,

$$|E(G)| \;=\; \sum_{e \in E} 1 \;=\; \sum_{e \in E} \sum_{t:e \in t} y_t \;=\; \sum_{t \in T} \sum_{e \in t} y_t \;=\; 3 \sum_{t \in T} y_t.$$

This implies:

$$\sum_{t \in T} y_t \;=\; \frac{|E|}{3} \;=\; OPT_f(G).$$

So, we can conclude:

$$OPT(G) \;\leq\; |F| \;\leq\; \frac{|E|}{2} \;=\; \frac{|E|}{3}\frac{3}{2} \;=\; \frac{3}{2} OPT_f(G).$$

4. ** (*10 pts*) Prove that $OPT(G) \leq 2\, OPT_f(G)$.

Solution: Let $S = \{e \mid x_e^* \geq \frac{1}{2}\}$ and let $E_0 = \{e \mid x_e^* = 0\}$. Now consider the graph $E' = E \setminus (E_0 \cup S)$.

We claim that every triangle in G containing an edge $e \in E_0$ also contains an edge in S. It remains to find a triangle hitting set for triangles with all edges e where $0 < x_e^* < \frac{1}{2}$, which are exactly the edges in E'. Note that $\mathbf{x}^*$ restricted to E' is a feasible solution for $(P_{triangle})$ on $G' = (V, E')$. Thus, the optimal solution for $(P_{triangle})$ on G' is at most that of $\mathbf{x}^*$ restricted to E'. We can resolve for the optimal solution of $(P_{triangle})$ on G' and repeat the procedure of adding edges for which $x_e^* \geq \frac{1}{2}$ to S and deleting edges with $x_e^* = 0$ from the graph.

Finally, we reach a point where $|S| \leq 2\, OPT_f(G)$ and S hits every triangle, or we have an optimal solution for some G' in which for all edges, $x_e^* > 0$. In the latter case, we apply the solution from Part 3.

For more on triangle packing and covering, see [Kri95].

References

[Kri95] Michael Krivelevich. On a conjecture of Tuza about packing and covering of triangles. *Discrete Mathematics*, 142(1-3):281–286, 1995.

[SW90] David B. Shmoys and David P. Williamson. Analyzing the Held-Karp TSP bound: A monotonicity property with application. *Information Processing Letters*, 35(6):281–285, 1990.