MTAT.03.286: Design and Analysis of Algorithms **University of Tartu**

Final exam

January 13th, 2020

Student name: ____

Student ID: ____

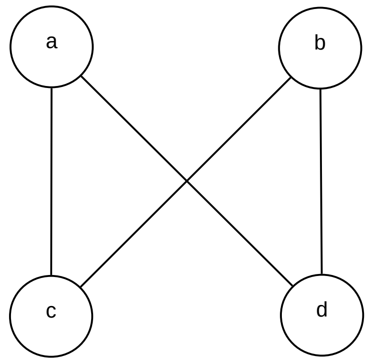
1. This exam contains 8 pages. Check that no pages are missing.
2. It is possible to collect up to 120 points. Try to collect as many points as possible.
3. Justify and prove all your answers (where applicable).
4. All facts and results that were proved or stated in the class can be used in your solution without a proof. Such results need to be rigorously formulated.
5. Any printed and written material is allowed in the class. No electronic devices are allowed.
6. Exam duration is 1 hour and 40 minutes.
7. Good luck!

Question 1	
Question 2	
Question 3	
Total	

Question 1 (35 points).

Reminder: A *vertex cover* in an undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is a subset of vertices $\mathcal{S} \subseteq \mathcal{V}$ such that every edge $e \in \mathcal{E}$ is incident with at least one vertex in $\mathcal{S}$. The *size* of $\mathcal{S}$ is a number of vertices in it.

We are interested in finding a minimum vertex cover $\mathcal{S}$ in the following undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$:



(a) For each vertex $v \in \mathcal{V}$ define an indicator variable $\mathcal{X}_v$ ($v \in \{a, b, c, d\}$):

$$\mathcal{X}_v = \begin{cases} 1 & \text{if } v \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} .$$

Formulate a problem of finding a minimum vertex cover in $\mathcal{G}$ as an *integer linear programming (ILP)* problem.

(b) Relax the ILP problem in part (a) to become a *linear programming (LP)* problem (explain the steps in your solution).

(c) What is an optimum solution to the ILP problem in part (a)?

(d) Does the choice $\mathcal{X}_a = \mathcal{X}_b = \mathcal{X}_c = \mathcal{X}_d = \frac{1}{2}$ give a feasible solution to the LP problem in (b)? Is it an optimum solution? Justify your answer.

(e) Does the choice in part (d) give an extreme point solution of the problem in part (b)?

Solution

(a)

$$\begin{aligned} \textbf{min} \quad & \mathcal{X}_a + \mathcal{X}_b + \mathcal{X}_c + \mathcal{X}_d \\ \textbf{s.t.} \quad & \mathcal{X}_a + \mathcal{X}_c \geq 1 \\ & \mathcal{X}_a + \mathcal{X}_d \geq 1 \\ & \mathcal{X}_b + \mathcal{X}_c \geq 1 \\ & \mathcal{X}_b + \mathcal{X}_d \geq 1 \\ & \mathcal{X}_a \in \{0, 1\}, \mathcal{X}_b \in \{0, 1\}, \mathcal{X}_c \in \{0, 1\}, \mathcal{X}_d \in \{0, 1\} \end{aligned}$$

(b) First, the variables $\mathcal{X}_v$ will belong to the continuous interval $[0, 1]$. Second, the restriction $\mathcal{X}_v \leq 1$ is not neccessary because the objective function is minimization, and the minimum cannot be obtained for $\mathcal{X}_v > 1$ (i.e., if we replace some $\mathcal{X}_v > 1$ in a feasible solution by $\mathcal{X}_v = 1$, we decrease the objective without violating any constraint). We obtain the following LP problem:

$$\begin{array}{rl} \textbf{min} & \mathcal{X}_a + \mathcal{X}_b + \mathcal{X}_c + \mathcal{X}_d \\ \textbf{s.t.} & \mathcal{X}_a + \mathcal{X}_c \geq 1 \\ & \mathcal{X}_a + \mathcal{X}_d \geq 1 \\ & \mathcal{X}_b + \mathcal{X}_c \geq 1 \\ & \mathcal{X}_b + \mathcal{X}_d \geq 1 \\ \multicolumn{2}{c}{\mathcal{X}_a \geq 0, \mathcal{X}_b \geq 0, \mathcal{X}_c \geq 0, \mathcal{X}_d \geq 0} \end{array}$$

(c) For example, $\mathcal{X}_a = \mathcal{X}_b = 1$, $\mathcal{X}_c = \mathcal{X}_d = 0$, is a feasible solution to the ILP problem in (a) because it satisfies all the constraints. The value of the objective function is 2. This is an optimal feasible solution to ILP in (a), because from the first and the fourth inequalities we obtain $\mathcal{X}_a + \mathcal{X}_b + \mathcal{X}_c + \mathcal{X}_d \geq 2$.

(d) The choice $\mathcal{X}_a = \mathcal{X}_b = \mathcal{X}_c = \mathcal{X}_d = \frac{1}{2}$ gives a feasible solution to the LP problem in (b) (but not to the ILP problem in (a)). The value of the objective function is 2. From the first and the fourth inequalities in the LP problem in (b) we obtain $\mathcal{X}_a + \mathcal{X}_b + \mathcal{X}_c + \mathcal{X}_d \geq 2$, and therefore this solution is optimal.

(e) $\mathcal{X}_a = \mathcal{X}_b = \mathcal{X}_c = \mathcal{X}_d = \frac{1}{2}$ is not an extreme point solution. Take, for example, feasible solutions $\mathcal{X}'_a = \mathcal{X}'_b = 1$, $\mathcal{X}'_c = \mathcal{X}'_d = 0$, and $\mathcal{X}''_a = \mathcal{X}''_b = 0$, $\mathcal{X}''_c = \mathcal{X}''_d = 1$. Then, $\frac{1}{2} = \mathcal{X}_v = \frac{1}{2}\left(\mathcal{X}'_v + \mathcal{X}''_v\right)$ for all $v \in \mathcal{V}$.

Question 2 (45 points). You are given the following linear-programming (LP) problem:

$$\begin{array}{rrcrcrcl}
\textbf{max} & x_1 & + & x_2 & - & 2x_3 & & \\
\textbf{s.t.} & x_1 & + & x_2 & & & \leq & 5 \\
 & x_1 & & & + & x_3 & \leq & 3 \\
- & x_1 & + & x_2 & + & 2x_3 & \leq & 0 \\
\end{array}$$
$$x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0$$

(a) Solve this LP problem by using the simplex method.

(b) Formulate the dual LP problem.

(c) Verify that your solution of (a) is optimal by the substitution of the point $(1, 0, 0)$ into the dual problem. Explain your answer.

Solution:

(a) Write the given problem in the tableau form:

$$\left[\begin{array}{cccccc|c}
1 & 1 & 0 & 1 & 0 & 0 & 5 \\
\textcircled{1} & 0 & 1 & 0 & 1 & 0 & 3 \\
-1 & 1 & 2 & 0 & 0 & 1 & 0 \\
\hline
1 & 1 & -2 & 0 & 0 & 0 & 0
\end{array}\right]
\begin{array}{l}(1)\\(2)\\(3)\\(4)\end{array}$$

Pick a variable with the positive multiplier in the last row, for example x_1. The smaller non-negative ratio corresponds to the choice of the encircled element as a pivot (it is also possible to chose as a pivot the element in the second column and third row instead).

$$\left[\begin{array}{ccc|ccc|c}
0 & \textcircled{1} & -1 & 1 & -1 & 0 & 2 \\
1 & 0 & 1 & 0 & 1 & 0 & 3 \\
0 & 1 & 3 & 0 & 1 & 1 & 3 \\
\hline
0 & 1 & -3 & 0 & -1 & 0 & -3
\end{array}\right]
\begin{array}{l}(1) \leftarrow (1) - (2)\\(2) \leftarrow (2)\\(3) \leftarrow (2) + (3)\\(4) \leftarrow (4) - (2)\end{array}$$

Now, we chose the encircled element as a pivot.

$$\left[\begin{array}{ccc|ccc|c}
0 & 1 & -1 & 1 & -1 & 0 & 2 \\
1 & 0 & 1 & 0 & 1 & 0 & 3 \\
0 & 0 & 4 & -1 & 2 & 1 & 1 \\
\hline
0 & 0 & -2 & -1 & 0 & 0 & -5
\end{array}\right]
\begin{array}{l}(1) \leftarrow (1)\\(2) \leftarrow (2)\\(3) \leftarrow (3) - (1)\\(4) \leftarrow (4) - (1)\end{array}$$

Now, all coefficients in the last row corresponding to x_1, x_2, x_3 are non-positive. The maximum of the objective function is 5. We have for this solution $x_1 = 3$, $x_2 = 2$ and $x_3 = 0$.

(b) By using the rules presented in the class, we write down the dual problem:

$$\begin{array}{lrcrcrcl}
\text{min} & 5y_1 & + & 3y_2 & & & & \\
\text{s.t.} & y_1 & + & y_2 & - & y_3 & \geq & 1 \\
 & y_1 & & & + & y_3 & \geq & 1 \\
 & & & y_2 & + & 2y_3 & \geq & -2 \\
\end{array}$$
$$y_1 \geq 0,\ y_2 \geq 0,\ y_3 \geq 0$$

(c) Consider a substitution $y_1 = 1, y_2 = 0, y_3 = 0$. It is straightforward to see that this substitution satisfies all the constraints of the dual problem, and therefore it is a feasible point. The value of the dual objective at this point is 5, which is equal to the value of the primal problem for $x_1 = 3, x_2 = 2, x_3 = 0$. By the strong duality theorem, indeed, 5 is the optimum value of both the primal and dual problem.

Question 3 (40 points).

Reminder:

- An *edge cover* in an undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is a subset of edges $\mathcal{C} \subseteq \mathcal{E}$ such that any vertex $v \in \mathcal{V}$ is incident with at least one edge in $\mathcal{C}$. The *size* of the edge cover is a number of edges in it.
- An *independent set* in an undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is a subset of vertices $\mathcal{I} \subseteq \mathcal{V}$ such that no two vertices in $\mathcal{I}$ are connected by an edge. The *size* of the independent set is a number of vertices in it.

Example:

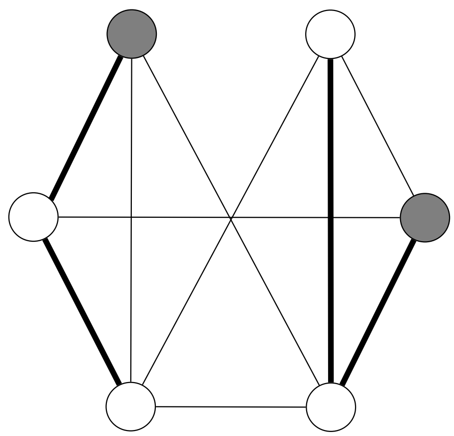


Figure 1: The bold edges form an edge cover. This edge cover has size 4, while the smallest edge cover in this graph has size 3. Moreover, the two gey vertices form an independent set.

Assume that $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is an undirected finite graph with degree of every vertex $v \in \mathcal{V}$ equal 3. Propose a simple algorithm for finding a smallest edge cover in $\mathcal{G}$ with approximation factor 3. Prove correctness of your solution.

Hint: it is possible to solve this problem by finding a maximal independent set first.

Solution 1:

Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be an undirected finite graph with degree of every vertex equal 3. Consider an algorithm, which does the following.

1. Pick any maximal independent set $\mathcal{I} \subseteq \mathcal{V}$ in $\mathcal{G}$.
2. $\mathcal{C} \leftarrow \varnothing$.
3. For every vertex $v \in \mathcal{I}$, add all three edges incident with v to the edge cover $\mathcal{C}$.

The first step can be efficiently implemented. Start with $\mathcal{I} = \varnothing$. Iteratively pick a vertex $v \in \mathcal{V}$ and check if it is adjacent to any other vertex that has already be chosen into $\mathcal{I}$. If not – add v to $\mathcal{I}$. Then, continue to the next vertex in $\mathcal{V}$.

The second step can be implemented in time linear in the size of $\mathcal{G}$.

- The constructed set $\mathcal{C}$ is an edge cover. If not, assume that there is a vertex $u \in \mathcal{V}$, which is not incident with any edge in $\mathcal{C}$. Then, this vertex could be added to $\mathcal{I}$, thus increasing its size, and preserving it being independent set. Therefore, the current $\mathcal{I}$ is not maximal – contradiction.

- Any edge cover should have size at least $|\mathcal{I}|$, because any vertex in $\mathcal{I}$ is covered by a different edge. Thus, $\mathsf{OPT} \geq |\mathcal{I}|$. The solution produced by the algorithm has size $3 \cdot |\mathcal{I}| \leq 3 \cdot \mathsf{OPT}$. Therefore, the approximation factor is 3.

□

Solution 2:

Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be an undirected finite graph with degree of every vertex equal 3. Consider an algorithm, which does the following.

1. For every vertex $v \in \mathcal{V}$, pick one of the edges incident with v, and add it to the vertex cover $\mathcal{C}$.

The proposed algorithm can be implemented in time linear in the size of the gaph: we need to iterate over all vertices and to do a fixed number of operations for each of them.

- The constructed set $\mathcal{C}$ is an edge cover. Indeed, every vertex is covered (due to the way how the edges were selected).

- Every edge can cover up to 2 vertices. Therefore, any edge cover in the graph should have size at least $|\mathcal{V}|/2$. Thus, $\mathsf{OPT} \geq |\mathcal{V}|/2$. The solution produced by the algorithm has size $\leq |\mathcal{V}| \leq 2 \cdot \mathsf{OPT}$. Therefore, the approximation factor is 2, and $2 \leq 3$.

□