ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Operations Research
ETH Zurich HG G21-22
Prof. Dr. Rico Zenklusen and Assistants
Contact: `math.opt@ifor.math.ethz.ch`

IFOR
Institute for Operations Research
Department of Mathematics

Fall 2019

Mathematical Optimization – Solutions to problem set 9

`https://moodle-app2.let.ethz.ch/course/view.php?id=4844`

Problem 1: Relaxation of the vertex cover polytope

(a) Let $S \subseteq V$. The following equivalences are true.

$$\begin{aligned} S \text{ is a vertex cover} \quad &\iff \quad \{u,v\} \cap S \neq \emptyset \quad \forall \{u,v\} \in E \\ &\iff \quad \chi^S(u) + \chi^S(v) \geq 1 \quad \forall \{u,v\} \in E \\ &\iff \quad \chi^S \in P \ . \end{aligned}$$

As incidence vectors are binary, this shows that for each vertex cover S, χ^S is in P. On the other hand, any binary vector in P can be written as χ^S for some $S \subseteq V$, and the above shows that in that case, S is a vertex cover. Thus, P is indeed a relaxation of the vertex cover polytope.

(b) Consider a point $x \in P$ that is not half-integral, i.e., x contains components not in $\{0, \frac{1}{2}, 1\}$. We show that x is not an extreme point of the polyhedron P by constructing two distinct points $x^1, x^2 \in P$ with the property that $x = \frac{1}{2}(x^1 + x^2)$. This implies that every extreme point of P is half-integral.

We call an edge $\{u,v\}$ *tight* if $x_u + x_v = 1$, otherwise we call it *loose*. Notice that there exists a constant $\varepsilon > 0$ with the property that $x - (\varepsilon, \varepsilon, \ldots, \varepsilon)$ does not violate any constraint corresponding to loose edges (ε can be chosen to be the minimum of $\frac{1}{2}(x_u + x_v - 1)$ over all loose edges $\{u,v\}$). We are going to construct points x^1 and x^2 by slightly modifying x (by increasing/decreasing some of its coordinates by $\delta \in [0, \varepsilon]$). By this, we are guaranteed that loose constraints remain valid for x^1 and x^2, and thus we only have to worry about constraints corresponding to tight edges.

Next, define the set of vertices corresponding to entries in x that are not half-integral, namely

$$U \coloneqq \left\{u \in V \colon x_u \notin \left\{0, \tfrac{1}{2}, 1\right\}\right\} \ .$$

Let $H \coloneqq G[U]$ be the induced subgraph of G on the vertex set H, i.e., $H = (U, F)$ with $F = \{e \in E \colon e \subseteq U\}$, and let H' be the subgraph of H containing only tight edges.

First, we note that every tight edge is either contained in H', or it is not incident to a vertex in U. In other words, there is no tight edge with one endpoint in U and one endpoint in $V \setminus U$. This holds since $x_u + x_v = 1$ for every tight edge $\{u,v\}$, and $u \in U$ implies $v \in U$.

Next, we notice that H' is bipartite. A partition $U = U^- \cup U^+$ certifying this is defined by

$$U^- \coloneqq \left\{u \in V \colon x_u \in \left(0, \tfrac{1}{2}\right)\right\} \quad \text{and} \quad U^+ \coloneqq \left\{u \in V \colon x_u \in \left(\tfrac{1}{2}, 1\right)\right\} \ .$$

Indeed, an edge $\{u,v\}$ connecting two vertices in U^- would imply $x_u + x_v < 1$, contradicting the feasibility of x, while an edge $\{u,v\}$ connecting two vertices in U^- would imply $x_u + x_v > 1$, implying that it is loose, contradicting the choice of H'.

We can now define x^1 and x^2 by

$$x^1_u = \begin{cases} x_u & \text{if } u \in V \setminus U \\ x_u - \delta & \text{if } u \in U^- \\ x_u + \delta & \text{if } u \in U^+ \end{cases} \quad \text{and} \quad x^2_u = \begin{cases} x_u & \text{if } u \in V \setminus U \\ x_u + \delta & \text{if } u \in U^- \\ x_u - \delta & \text{if } u \in U^+ \end{cases} \ .$$

We emphasize that for an appropriately chosen small $\delta \leq \varepsilon$, both x^1 and x^2 remain in the set $[0,1]^V$. As we mentioned earlier, to verify feasibility of x^1 and x^2 we only need to check the constraints for tight edges in H'. For these edges, however, the constraints obviously remain satisfied, as for every such edge $\{u, v\}$ and every $i \in \{1, 2\}$, we have $x_u^i + x_v^i = x_u + x_v + \delta - \delta = 1$, since u and v belong to different parts in the partition $U^- \cup U^+$ of the vertex set of H'. Finally, we notice that $x = \frac{1}{2}(x^1 + x^2)$ indeed holds true. This finishes the proof.

(c) Let x^* be an optimal solution of the linear program $\min\{w(x) \colon x \in P\}$.

We claim that the algorithm returning the set

$$U \coloneqq \left\{u \in V \colon x_u^* \geq \tfrac{1}{2}\right\}$$

is a 2-approximation for the minimum weight vertex cover problem. To see this, we first show that U is a vertex cover. Indeed, observe that for every edge $\{u, v\} \in E$, the constraint $x_u^* + x_v^* \geq 1$ implies that either at least one of the inequalities $x_u^* \geq \frac{1}{2}$ and $x_v^* \geq \frac{1}{2}$ holds true, hence at least one of the endpoints is contained in U. Additionally, we have

$$w(U) \leq 2 \cdot w(x^*) \leq 2 \cdot \mathrm{OPT} \ ,$$

where OPT denotes the value of a minimum vertex cover. Here, the first inequality follows from the fact that U contains only vertices $u \in V$ that had a value of at least $\frac{1}{2}$ in x^*, hence contributed at least $\frac{1}{2} w_v$ to the value of x^*, while now they contribute w_v to the value of U. The second inequality is true because P is a relaxation of the vertex cover polytope, and hence we have $w(x^*) \leq \mathrm{OPT}$.

This proves that indeed, U is a 2-approximate solution of the minimum vertex cover problem. Given that the linear program $\min\{w(x) \colon x \in P\}$ can be solved efficiently, we can find U efficiently.

Problem 2: Incidence Matrices and Total Unimodularity

Assume that $A \in \{0, 1\}^{V \times E}$ is totally unimodular. By the Characterization of Ghouila-Houri (Theorem 5.9 in the script) this implies existence of a partition $V = R_1 \dot{\cup} R_2$ of the vertices (which correspond to the rows of A) such that

$$\sum_{v \in R_1} A_{ve} - \sum_{v \in R_2} A_{ve} \in \{-1, 0, 1\} \quad \forall e \in E \ .$$

Note that this implies that any edge of G must have one endpoint in R_1 and one in R_2. Indeed, if $e \subseteq R_1$ for some $e \in E$, then

$$\sum_{v \in R_1} A_{ve} - \sum_{v \in R_2} A_{ve} = 2 \ ,$$

contradicting the choice of R_1 and R_2 by Ghouila-Houri. Similarly, if $e \subseteq R_2$, the above sum has value -2, which is a contradiction, as well.

Thus, $V = R_1 \dot{\cup} R_2$ is a bipartition for the graph G (all edges cross from R_1 to R_2 or vice versa), hence G is bipartite.

Problem 3: Max-flow min-cut via duality II

(a) The proof consists of three steps. In point (i) below, we prove that the constraint matrices of the primal and dual linear programs (which are transposes of each other) are totally unimodular. We do so by using the characterization of Ghouila-Houri (Theorem 5.9 in the script). In point (ii), we show that adding the constraint $y \in \mathbb{R}^V_{\geq 0}$ to the dual linear program does not change the optimal value. These first two steps together with the fact that the right hand sides of the constraints are integral imply that the feasible region of the dual linear program (intersected with $y \geq 0$) is integral. Finally, in point (iii) we show that any integral optimal solution can be transformed into an optimal $\{0, 1\}$ solution. Altogether, these steps imply that the dual linear program has an optimal $\{0, 1\}$ solution.

(i) Let us look at the constraint matrix of the primal. Observe that the coefficients of the rows coming from the constraints

$$\sum_{a\in\delta^+(v)} f_a - \sum_{a\in\delta^-(v)} f_a \quad = \quad \begin{cases} \nu & \text{if } v = s \\ -\nu & \text{if } v = t \\ 0 & \text{if } v \in V \setminus \{s,t\} \end{cases}$$

form a matrix with precisely one entry $+1$ and one entry -1 in every column (note that ν is a variable, too, and it contributes a column with the same property to the matrix). Thus, if we apply the Ghouila-Houri criterion (Theorem 5.9 in the script) to the rows of this matrix, no matter what subset of we consider, their sum will be a $\{-1,0,1\}$-vector. Consequently, this part of the constraint matrix is totally unimodular.

Also taking the constraints $u_a \geqslant f_a \geqslant 0$ and $\nu \geqslant 0$ into account corresponds to appending unit vectors to the previously discussed totally unimodular matrix, which does not destroy total unimodularity, as can be seen by applying the Ghouila-Houri criterion to the columns: The previously analyzed part of the matrix is TU, hence for every subset of the rows, Ghouila-Houri guarantees a suitable partition—but this partition can be used for the new rows, too, as every new row has a single non-zero entry only.

(ii) Let $(y,z) \in \mathbb{R}^V \times \mathbb{R}^A_{\geq 0}$ be any solution of the dual linear program. Denote by $u \in V$ the vertex with the minimal value y_r, i.e., such that $y_r \leq y_v$ for every $v \in V$. Then $(\widetilde{y}, z)$, where $\widetilde{y}_v := y_v - y_r$, is a solution of the dual linear program with the same value as (y,z), and it satisfies $\widetilde{y}_v \geq 0$. Therefore, the optimal value of the linear program remains the same if we add the non-negativity constraint $y \in \mathbb{R}^V_{\geq 0}$. Thus, without loss of generality we may consider the modified linear program (with the additional non-negativity constraint) instead of the original one.

(iii) Let $y \in \mathbb{Z}^V$ be an integral optimal solution of the modified dual linear program. Note that by integrality, $y_v \notin (y_s, y_s+1)$ for every $v \in V$, so either $y_v \leq y_s$ or $y_v \geq y_s + 1$. For every $v \in V$, define

$$\widetilde{y}_v := \begin{cases} y_s & \text{if } y_v \leq y_s \\ y_s + 1 & \text{if } y_v \geq y_s + 1 \end{cases}.$$

Let us prove that $(\widetilde{y}, z)$ is an integral optimal solution of the modified dual linear program, as well. The non-negativity constraints $y \in \mathbb{R}^V_{\geq 0}$ and $z \in \mathbb{R}^A_{\geq 0}$ are satisfied by construction. To show that for any arc $a = (v,w) \in A$, the corresponding constraint $\widetilde{y}_v - \widetilde{y}_w + z_a \geq 0$ is satisfied, we consider the four possible relationships between y_s and the variables y_v and y_w:

- If we have $y_v, y_w \leq y_s$ or $y_v, y_w \geq y_s + 1$, then $\widetilde{y}_v - \widetilde{y}_w = 0$, and since $z_a \geq 0$, the constraint corresponding to a holds.
- Similarly, if $y_v \geq y_s + 1$ and $y_w \leq y_s$, then $\widetilde{y}_v - \widetilde{y}_w = 1$, and again, as $z_a \geq 0$, the constraint is satisfied.
- Finally, if $y_v \leq y_s$ and $y_w \geq y_s + 1$, then $\widetilde{y}_v - \widetilde{y}_w = y_s - (y_s+1) \geq y_v - y_w$, so $\widetilde{y}_v - \widetilde{y}_w + z_a \geq y_v - y_w + z_a \geq 0$, as desired (the last inequality uses feasibility of (y,z)).

Thus, all constraints corresponding to arcs $a \in A$ hold. Finally, note that $\widetilde{y}_s = y_s$ and $y_t \geq y_s + 1$, so $\widetilde{y}_t = y_s + 1$, hence the last constraint $\widetilde{y}_t - \widetilde{y}_s \geq 1$ is satisfied. Therefore, $(\widetilde{y}, z)$ is an integral feasible solution of the modified dual LP. Moreover, it is an optimal solution, since the value of the solution depends only on z and (y,z) is an optimal solution.

Based on the observation in point (ii), we may further assume that $\widetilde{y}_s = 0$ (after replacing $\widetilde{y}_v$ by $\widetilde{y}_v - \widetilde{y}_s$ for all $v \in V$), which transforms the solution to one with $\widetilde{y} \in \{0,1\}^V$. This also implies that we can have $z \in \{0,1\}^A$: $z_a \geqslant 0$ is a constraint anyway, and if there exists $a \in A$ with $z_a > 1$, we can set $z_a = 1$ without affecting feasibility while not increasing the solution value. Thus, the dual linear program has an optimal $\{0,1\}$ solution.

(b) Since $(y,z) \in \{0,1\}^V \times \{0,1\}^A$, the constraint $y_t - y_s \geq 1$ implies that $y_s = 0$ and $y_t = 1$. Thus, $s \in C$ and $t \notin C$ by construction, so C is an s-t cut.

Furthermore, we show that the value $u(\delta^+(C))$ of the cut C equals the optimal value of the dual LP, which is the sum $\sum_{a\in A} u_a z_a$. To this end, we prove that

$$u_a z_a = \begin{cases} u_a & \text{if } a \in \delta^+(C) \\ 0 & \text{else} \end{cases} \quad . \tag{1}$$

By construction of C, we have $v \in C$ and $w \notin C$ for every arc $a = (v, w) \in \delta^+(C)$, which means that $y_v = 0$ and $y_w = 1$. Hence the constraint $y_v - y_w + z_a \geq 0$ together with the fact that $z_a \in \{0, 1\}$ implies that $z_a = 1$, and hence indeed $u_a z_a = u_a$. Conversely, if $a = (v, w) \notin \delta^+(C)$, then $y_v = y_w$, hence the constraint $y_v - y_w + z_a \geq 0$ is equivalent to just $z_a \geq 0$. Thus, the only constraint on z_a in the dual LP is a non-negativity constraint. If $u_a > 0$, then as the dual is a minimization problem and by optimality of (y, z), we get $z_a = 0$ and thus $u_a z_a = 0$. If $u_a = 0$, the same conclusion is immediate.

This proves (1), from which we immediately obtain $\sum_{a\in A} u_a z_a = \sum_{a\in\delta^+(C)} u_a = u(\delta^+(C))$.

(c) As remarked in the problem statement, the optimal value of the primal linear program is equal to the value of a maximum *s*-*t* flow, and the optimal value of the dual linear program is at most the value of a minimum *s*-*t* cut. By point (b) of this problem, there even exists an *s*-*t* cut whose value equals the optimal value of the dual LP, and hence the value of the dual linear program is equal to the value of a minimum *s*-*t* cut. Since both the primal and the dual LP are feasible, strong linear programming duality implies that the primal and the dual values are the same, i.e., the value of a maximum *s*-*t* flow equals the value of a minimum *s*-*t* cut.

Remark: From the above arguments (total unimodularity of the primal constraint matrix and an application of Theorem 5.8 in the script), we also directly see that if the arc capacities are integral, then there always exists an integral maximum flow.

Problem 4: Matching Polytope Relaxation
Consider the graph G, the complete graph on three vertices (i.e., a triangle), given in Figure 1.

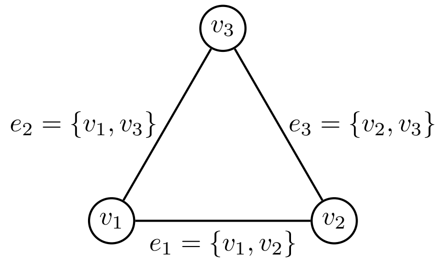

Figure 1: The triangle graph K_3.

We present two different ways to show that $P \neq P_{\mathcal{M}}$ for this graph G.

First proof of $P \neq P_{\mathcal{M}}$. Consider the system defining the polyhedron P, which is

$$\begin{array}{ccccccc} x_1 & + & x_2 & & & \leq & 1 \\ x_1 & & & + & x_3 & \leq & 1 \\ & & x_2 & + & x_3 & \leq & 1 \\ & & & & x_i & \geq & 0 \quad \forall i \in \{1, 2, 3\} \ . \end{array}$$

Since the first three constraints are linearly independent, they correspond to a basic solution of the system which, a priori, may be feasible or infeasible. This solution is obtained by replacing inequalities with equalities and computing the unique solution of the linear system. In this case we obtain the point $y = (1/2, 1/2, 1/2)$, which clearly satisfies all constraints of P and hence is feasible, i.e., $y \in P$. Thus, y is a vertex of P. However, y is not integral, and hence P is not an integral polytope, while $P_{\mathcal{M}}$ is. This implies $P \neq P_{\mathcal{M}}$.

Second proof of $P \neq P_{\mathcal{M}}$. Notice that the set set of all incidence vectors corresponding to matchings in G is given by

$$X \coloneqq \{(0,0,0),\ (1,0,0),\ (0,1,0),\ (0,0,1)\}\ ,$$

and we have $P_{\mathcal{M}} = \mathrm{conv}(X)$. Let $w = (1,1,1)$ and consider the linear optimization problem $\max\{w^\top x\colon x \in P_{\mathcal{M}}\}$, and observe that by definition of $P_{\mathcal{M}}$, we have

$$\max\{w^\top x\colon x \in P_{\mathcal{M}}\} = \max\{w^\top x\colon x \in X\} = 1\ .$$

If $P = P_{\mathcal{M}}$, we would in particular also have $\max\{w^\top x\colon x \in P\} = 1$. But $y = (1/2, 1/2, 1/2)$ is easily seen to be feasible for P, and $w^\top y = \frac{3}{2} > 1$, a contradiction. Note that this second solution did not use that y is a vertex of P.

Problem 5: Low Discrepancy Coloring

(a) Consider the polytope defined by the following natural constraints, expressing exactly the condition of good sets for integral vectors:

$$Q \coloneqq \left\{x \in \mathbb{R}^n_{\geq 0}\colon x_{\sigma_i(2t-1)} + x_{\sigma_i(2t)} = 1 \text{ for all } i \in \{1,2\} \text{ and } t \in [n/2]\right\}\ .$$

It is obvious that $Q \cap \{0,1\}^n$ corresponds exactly to the family of all good sets. We prove next that Q is integral, which implies $Q = P_{\mathcal{R}}$. We do this by showing that the constraint matrix A defining Q is totally unimodular using the Ghouila-Houri criterion (Theorem 5.9 in the script), and then apply Theorem 5.8 from the script.

To show that the conditions of the Ghouila-Houri criterion hold, let $J = \{A_{j_1}, \ldots, A_{j_t}\}$ be a subset of the rows of A. We have to find a partition $J = J_1 \,\dot\cup\, J_2$ of the rows such that

$$\sum_{j \in J_1} A_j - \sum_{j \in J_2} A_j \in \{-1, 0, 1\}^n\ . \tag{2}$$

We call a row $A_j \in J$ a *row of type* 1 if it comes from a constraint of the form $x_{\sigma_1(2t-1)} + x_{\sigma_1(2t)} = 1$, and we call it a *row of type* 2 if it comes from a constraint of the form $x_{\sigma_2(2t-1)} + x_{\sigma_2(2t)} = 1$. As every variable has non-zero coefficient in precisely one row of type 1 and precisely one row of type 2, a partition achieving (2) is given by $J_1 \coloneqq \{A_j \in J\colon A_j \text{ is of type } 1\}$ and $J_2 \coloneqq \{A_j \in J\colon A_j \text{ is of type } 2\}$.

(b) To obtain the desired set with discrepancy 2, we first notice that there always exists a good set. Indeed, since $Q = P_{\mathcal{R}}$ describes the polytope of all good sets, existence of a good set is implied from the fact that Q is non-empty: note that $\frac{1}{2} \cdot (1, 1, \ldots, 1) \in Q$, hence by integrality there must also be an integral point in $P_{\mathcal{R}}$. Such a point corresponds to a good set R.

Finally, we show that such a set R has discrepancy 2. To this end, note that for every $i \in \{1,2\}$ and every $t \in [n/2]$, the condition of good sets implies that

$$\big||\{\sigma_i(2t-1), \sigma_i(2t)\} \cap R| - |\{\sigma_i(2t-1), \sigma_i(2t)\} \setminus R|\big| = 0\ .$$

Thus, for every $i \in \{1,2\}$ and every $I = \{\sigma_i(l), \sigma_i(l+1), \ldots, \sigma_i(u)\}$, every pair $\{\sigma_i(2t-1), \sigma_i(2t)\}$, for $t \in [n/2]$, that is fully contained in I or has empty intersection with I does not contribute anything to $\big||I \cap R| - |I \setminus R|\big|$. Since at most the first and/or last element of I remain, we conclude that

$$\big||I \cap R| - |I \setminus R|\big| \leq 2\ ,$$

as desired.

Problem 6: Laminar Matroids

In order to prove that $P_{\mathcal{F}} = P$, we first observe that $P_{\mathcal{F}} \cap \{0,1\}^N = P \cap \{0,1\}^N$, which holds since for an incidence vector χ^S of some set $S \subseteq N$, we have $|S \cap L| = \chi^S(L)$ for all $L \in \mathcal{L}$, hence the conditions in $\mathcal{F}$ and those in P, restricted to integral points, are equivalent.

Thus, we prove $P_{\mathcal{F}} = P$ by verifying that P is integral. To this end, rewrite the polyhedron P in the form $P = \{x \in \mathbb{R}_{\geqslant 0} \mid Ax \leqslant b\}$, where the system $Ax \leqslant b$ of inequalities contains all constraints of the form $x(L) \leqslant b_L$ for $L \in \mathcal{L}$ and all constraints of the form $x(e) \leqslant 1$ for $e \in N$. Since the right-hand side b of the inequality constraints defining P is integral, (the entries of b are 1 or b_L for $L \in \mathcal{L}$, and we know $b_L \in \mathbb{Z}_{>0}$), it is enough to show that the constraint matrix A is totally unimodular to conclude integrality of P. Thereto, we partition the matrix into two parts A_1 and A_2, where A_1 contains all rows corresponding to constraints of the form $x(L) \leqslant b_L$ for $L \in \mathcal{L}$, and A_2 contains all rows corresponding to constraints of the form $x(e) \leqslant 1$ for $e \in N$. Note that A_2 is an identity matrix, thus the matrix $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ is totally unimodular if and only if A_1 is totally unimodular. We proceed by showing that the matrix A_1 is actually a consecutive-ones matrix, given a proper ordering of the variables. Concretely, this ordering is constructed by the following inductive rule. Given a set of variables $X \subseteq [n]$ that remain to be ordered, choose any maximal set $L \in \mathcal{L}$ that is a subset of X, namely a set $L \subseteq X$ with the property that every set $L' \in \mathcal{L}$ with $L \subsetneq L'$ is *not* a subset of X. Place all variables in L in the beginning of the ordering of X, and all variables in $X \setminus L$ after them. Now inductively apply the same procedure on the sets L and $X \setminus L$ to obtain the ordering within each set. If at some point there is no set $L \in \mathcal{L}$ with $L \subseteq X$, choose any arbitrary ordering of the elements in X and return it. The procedure starts with $X = [n]$.

The fact that $\mathcal{L}$ is a laminar family implies that the latter procedure uses *every* set $L \in \mathcal{L}$ to split some set $X \supseteq L$. It follows that the resulting ordering $i_1, \ldots, i_n$ of $[n]$ has the property that for every $L \in \mathcal{L}$ with $L \neq \emptyset$ we have $L = \{i_l, i_l + 1, \ldots, i_u\}$ for some $l \leq u$. It is now obvious that the matrix A_1 is a consecutive-ones matrix (the rows are given by χ^L for $L \in \mathcal{L}$) with respect to this variable ordering. This finishes the proof.

The figure on the problem sheet illustrates one possible ordering of the elements $e_1, \ldots, e_{12}$ such that the corresponding matrix A_1 is a consecutive-ones matrix.