Solutions

Mathematical Optimization

Exam from January 22, 2020

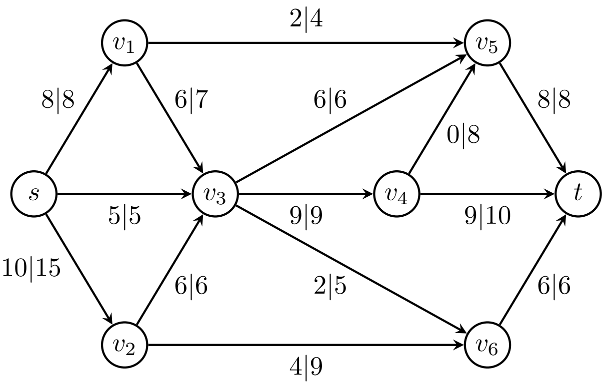
Problem 1: Short questions (4×2.5 points)

(a) Determine all legal pivot elements in the short tableau given below. For each of them, fill in a line of the given table by stating the variables leaving and entering the basis, respectively, as well as the change in the objective value if an exchange step on the corresponding pivot is performed.

	x_1	x_2	x_3	x_4	1
z	-1	0	-2	3	4
x_5	-2	3	-1	-1	-2
x_6	1	4	0	2	0
x_7	3	2	5	-1	3

Pivot element value	*Variable leaving the basis*	*Variable entering the basis*	*Change in objective value*
1	x_6	x_1	0
5	x_7	x_3	$\frac{6}{5}$

(b) In the directed graph $G = (V, A)$ below, arc capacities $u\colon A \to \mathbb{Z}_{\geq 0}$ and values of a maximal s-t flow $f\colon A \to \mathbb{R}_{\geq 0}$ are given in the form $f(a)|u(a)$ on every arc. Provide a minimum s-t cut and prove that it is indeed a minimum s-t cut.



The following vertices form a minimum s-t cut: $\boxed{s, v_1, v_2, v_3, v_5, v_6}$

Proof: The cut $S = \{s, v_1, v_2, v_3, v_5, v_6\}$ has value $u(\delta^+(S)) = 8 + 9 + 6 = 23$ which is the value of the given (maximal) flow. Thus, by the strong max-flow min-cut theorem, S is a minimum cut.

(c) Consider the integer linear program

$$\begin{array}{lccccccccccccc}
\max & & & & & - & 2x_5 & - & x_6 & - & 3x_7 & & \\
 & x_1 & & & & + & x_5 & - & \frac{5}{4}x_6 & + & \frac{1}{2}x_7 & = & \frac{19}{4} \\
 & & x_2 & & & + & x_5 & - & x_6 & + & x_7 & = & 7 \\
 & & & x_3 & & - & \frac{3}{2}x_5 & + & x_6 & - & \frac{3}{4}x_7 & = & \frac{7}{2} \\
 & & & & x_4 & + & \frac{5}{4}x_5 & - & \frac{7}{2}x_6 & + & x_7 & = & \frac{21}{4} \\
 & & & & & & & & & & x & \in & \mathbb{Z}^7_{\geq 0} \ .
\end{array}$$

The point $x^* = (\frac{19}{4}, 7, \frac{7}{2}, \frac{21}{4}, 0, 0, 0)$ is an optimal solution of the linear relaxation of the above problem, where the constraint $x \in \mathbb{Z}^7_{\geq 0}$ is replaced by $x \in \mathbb{R}^7_{\geq 0}$.

Provide a cutting plane that cuts off x^* from the feasible region of the relaxation without cutting off any solution of the integer program. You do not have to indicate how you found the cutting plane, and you do not have to prove that it has the desired properties.

Cutting plane: $\boxed{x_1 + x_5 - 2x_6 \leq 4}$

Such a cutting plane can be found by considering a fractional basic variable, e.g. x_1, and rounding down each coefficient of its corresponding constraint, in this case $x_1 + x_5 - \frac{5}{4}x_6 + \frac{1}{2}x_7 = \frac{19}{4}$.

(d) Let $f, g : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ be two functions. Does $\log f = O(\log g)$ imply $f = O(g)$? Write your answer below and prove that it is correct.

Answer: No, the implication does not hold.

Proof: Consider the functions $f(n) = n^2$ and $g(n) = n$. We have $\log n^2 = 2 \log n = O(\log n)$, i.e. $\log f = O(\log g)$, however it is not true that $f = O(g)$. Indeed, $\frac{f(n)}{g(n)} = n$ is unbounded as $n \to \infty$, thus there is no $c > 0$ s.t. $f(n) \leq c \cdot g(n)$ holds for large enough n.

Problem 2: Facets of polars and satisfying distance requirements $(4 + 6$ points)

(a) By Lemma 7.2. of the lecture notes, for every polytope Q that contains the origin in its interior, there is a one-to-one correspondence between the vertices of Q and the facets of Q°. In particular, this implies that $|\text{vertices}(Q)| = |\text{facets}(Q^\circ)|$. As both P_1 and P_2 are polytopes that contain the origin in their interior, we have

$$\begin{aligned} |\text{facets}(P_1^\circ)| &= |\text{vertices}(P_1)| \ , \\ |\text{facets}(P_2^\circ)| &= |\text{vertices}(P_2)| \ , \end{aligned}$$

and

$$\left|\text{facets}\big(\big(\operatorname{conv}(P_1 \cup P_2)\big)^\circ\big)\right| = \left|\text{vertices}\big(\operatorname{conv}(P_1 \cup P_2)\big)\right| \ ,$$

since either of P_1, P_2 containing the origin in their interior implies that $\operatorname{conv}(P_1, P_2)$ contains the origin in its interior, and since $\operatorname{conv}(P_1, P_2)$ is a polytope, again (namely the convex hull of the union of the vertices of P_1 and the vertices of P_2). Thus it remains to show that

$$\left|\text{vertices}\big(\operatorname{conv}(P_1 \cup P_2)\big)\right| \le \left|\text{vertices}(P_1)\right| + \left|\text{vertices}(P_2)\right| \ .$$

We will in fact prove a stronger property, namely

$$\text{vertices}\big(\operatorname{conv}(P_1 \cup P_2)\big) \subseteq \text{vertices}(P_1) \cup \text{vertices}(P_2) \ .$$

To see this, let $x \in \text{vertices}(\operatorname{conv}(P_1 \cup P_2))$. Then in particular, $x \in \operatorname{conv}(P_1 \cup P_2)$, and by definition of the latter, there exist $x_1 \in P_1$, $x_2 \in P_2$ and $\lambda \in [0, 1]$ such that

$$x = \lambda x_1 + (1 - \lambda) x_2 \ .$$

As we have $x_1, x_2 \in \operatorname{conv}(P_1 \cup P_2)$, x can only be a vertex of $\operatorname{conv}(P_1 \cup P_2)$ if either $\lambda \in \{0, 1\}$.

If $\lambda = 0$, then $x = x_1 \in P_1$. Assume for contradiction that $x \notin \text{vertices}(P_1)$, then $x = \mu x' + (1 - \mu) x''$ for some $x', x'' \in P_1$ and $\mu \in (0, 1)$. But as we also have $x', x'' \in \operatorname{conv}(P_1 \cup P_2)$, this shows that $x \notin \text{vertices}(\operatorname{conv}(P_1 \cup P_2))$, contradicting the assumption. Thus, we must have $x \in \text{vertices}(P_1)$ in this case.

The case where $\lambda = 1$ is analogous to the above, and allows for the conlusion that $x \in \text{vertices}(P_2)$. This completes the proof.

(b) Let $(\star)$ denote the following property of the arc lengths ℓ:

For all $v \in V \setminus \{s\}$, the shortest s-v path in G with respect to ℓ has length at least $r(v)$.

Consider the linear program

$$\begin{aligned} \min \quad & \sum_{a \in A} \ell_a \\ d_s &= 0 \\ d_v &\le d_u + \ell_a && \text{for every arc } a \text{ from } u \text{ to } v \\ d_v &\ge r(v) && \text{for every node } v \\ d &\in \mathbb{R}^V_{\ge 0} \\ \ell &\in \mathbb{R}^A_{\ge 0} \ . \end{aligned}$$

We claim that an optimal solution (d, ℓ) of this LP yields arc lengths ℓ that minimize the total arc length while satisfying $(\star)$.

- First, for a given ℓ satisfying $(\star)$, setting d_v to the actual distance from s to v for all $v \in V$ gives a feasible point of the LP. Thus, the optimal value of the LP is at most the minimum total arc length that can be achieved while satisfying $(\star)$.

- Now let (d, ℓ) be an optimal solution of the LP. We show that ℓ satisfies $(\star)$, and hence the LP value is at least the minimum total arc length that can be achieved while satisfying $(\star)$.

  To see that $(\star)$ is satisfied, it is enough to prove that for every $v \in V$, d_v is equal to the length of a shortest s-v path with respect to ℓ; then the inequality $d_v \geq r(v)$ guarantees precisely what $(\star)$ states.

  Fix a node v and let $s = v_0, v_1, \ldots, v_k = v$ be the vertices on a shortest s-v path P in G with respect to ℓ (in the order given by the path, i.e., such that $a_i \coloneqq (v_{i-1}, v_i)$ for $i \in [k]$ are the edges on the path). By the constraints in the LP, we have

  $$\sum_{i=1}^{k} d_{v_i} \leq \sum_{i=1}^{k} d_{v_{i-1}} + \ell_{a_i} \ , \quad \text{which implies} \quad d_v \leq \ell(P) \ ,$$

  hence d_v is at most the length of the shortest s-v path with respect to ℓ. If equality does not hold, then there is $i \in [k]$ such that $d_{v_i} < d_{v_{i-1}} + \ell_{a_i}$. Observe that ℓ_{a_i} appears in the LP in precisely this constraint only, hence we could decrease ℓ_{a_i} without affecting correctness of the other constraints, but while decreasing the objective value of the LP. This contradicts the assumption that (d, ℓ) is an optimal LP solution. Consequently, we must have that d_v equals the length of a shortest s-v path in G for all v, as desired.

Together, the above two points show that the LP value is equal to the shortest arc length that can be achieved while satisfying $(\star)$, and that an optimal solution gives suitable arc lengths ℓ. This proves the claim.

Note that the LP has $|V| + |A|$ many variables and $|V| + |A|$ many constraints, with $r(v)$ for $v \in V$ being the only numbers in the description. Thus, its input size is polynomial in the input size of G and r, which is the problem input, and hence we can solve the LP in time polynomial in the input size of G and r (using the given hint). Thus, we can also solve our problem in polynomial time, as desired.

Problem 3: Polyhedral description of linearly independent subsets $(2+3+5$ points$)$

(a) We first prove that $P \cap \{0,1\}^N \subseteq \{\chi^S : S \in \mathcal{I}\}$. To this end, let $x \in P \cap \{0,1\}^N$ and denote $S_1 := \{n \in N : x(n) = 1\}$. Then $x = \chi^{S_1}$ and we have to show that $S_1 \in \mathcal{I}$, i.e., that the vectors in S_1 are linearly independent. Assume for contradiction that they are not linearly independent. Then $r(S_1) = \dim(\mathrm{span}(S_1)) < |S_1| = \chi^{S_1}(S_1) = x(S_1)$, but this contradicts the assumption that $x \in P$.

Let us now prove that $P \cap \{0,1\}^N \supseteq \{\chi^S : S \in \mathcal{I}\}$. To this end, let $S \in \mathcal{I}$ be a set of linearly independent vectors. We have to show that $\chi^S \in P$, i.e., that $\chi^S(S') \leq r(S')$ for all $S' \subseteq N$. Observe that for every such S', we have $\chi^S(S') = |S \cap S'| = \dim(\mathrm{span}(S \cap S')) \leq \dim(\mathrm{span}(S')) = r(S')$, giving the desired.

(b) If $S_1, S_2 \subseteq N$ are both y-tight, we have

$$y(S_1) = r(S_1) \qquad \text{and} \qquad y(S_2) = r(S_2) \ .$$

Summing these two inequalities and using the given property of the function r, we get

$$y(S_1) + y(S_2) = r(S_1) + r(S_2) \geq r(S_1 \cap S_2) + r(S_1 \cup S_2) \ .$$

As $y \in P$, we do in particular have $y(S_1 \cap S_2) \leq r(S_1 \cap S_2)$ and $y(S_1 \cup S_2) \leq r(S_1 \cup S_2)$. Combining this with the above, we obtain

$$y(S_1) + y(S_2) \geq y(S_1 \cap S_2) + y(S_1 \cup S_2) \ .$$

But observe that this inequality is in fact an equality—hence all inequalities that we used must be equalities. In particular, we must have that $y(S_1 \cap S_2) = r(S_1 \cap S_2)$ and $y(S_1 \cup S_2) = r(S_1 \cup S_2)$, i.e., $S_1 \cap S_2$ and $S_1 \cup S_2$ are both y-tight, as desired.

(c) Let $\mathcal{F} \subseteq 2^N$ be the family of all y-tight sets. Then the equation system

$$x(F) = r(F) \quad \forall F \in \mathcal{F} \tag{1}$$

is the system of all y-tight constraints. As y is a vertex, this system (1) has y as a unique solution. In other words, the system has full rank, i.e., rank $|N|$.

Let $\mathcal{H}$ be a maximal laminar subfamily of $\mathcal{F}$. We claim that the system

$$x(H) = r(H) \quad \forall H \in \mathcal{H} \tag{2}$$

implies all constraints in system (1). This directly implies that the rank of system (2) is $|N|$, too, as desired.

To prove the claim, by the second part of the hint, it is enough to prove that

$$\chi^F \in \mathrm{span}(\{\chi^H : H \in \mathcal{H}\}) \quad \forall F \in \mathcal{F} \ , \tag{3}$$

as the two systems have the common solution y. Let us denote $Q := \mathrm{span}(\{\chi^H : H \in \mathcal{H}\})$. We prove the above by contradiction and assume that (3) does not hold, i.e., that there exist $F \in \mathcal{F}$ such that $\chi^F \notin Q$. Among all such sets, fix the one F such that

$$\mathcal{H}_F := \{H \in \mathcal{H} : H \text{ and } F \text{ are intersecting}\}$$

has smallest size. We have $\mathcal{H}_F \neq \emptyset$, for otherwise we could have added F to $\mathcal{H}$ without destroying laminarity of $\mathcal{H}$, which contradicts maximality of $\mathcal{H}$. Hence, there exists $H \in \mathcal{H}_F$. By part b and from the fact that H and F are both y-tight, we know that $H \cap F$ and $H \cup F$ are both y-tight, too. Furthermore, we have

$$\chi^F + \chi^H = \chi^{H \cup F} + \chi^{H \cap F} \ .$$

By definition, we have $\chi^F \notin Q$ and $\chi^H \in Q$. Hence, at least one of $\chi^{H \cup F}$ and $\chi^{H \cap F}$ is not in Q. However, by the first part of the hind, we have

$$|\mathcal{H}_{H \cup F}| < |\mathcal{H}_F| \qquad \text{and} \qquad |\mathcal{H}_{H \cap F}| < |\mathcal{H}_F| \ .$$

This contradicts the choice of F such that $\mathcal{H}_F$ has smallest possible size, and thus proves that (3) holds. This finishes the proof.

Problem 4: r-arborescences and separation $(2 + 8$ points)

(a) The dominant P of the r-arborescence polytope is given by

$$P = \left\{x \in \mathbb{R}^A_{\geq 0} \colon x(\delta^-(S)) \geq 1 \text{ for all } S \subseteq V \setminus \{r\},\ S \neq \emptyset\right\} \ .$$

(b) For designing a separation oracle, we stick to the standard assumption that we are given a rational representation of y, i.e., $y \in \mathbb{Q}^A$. Observe that P has two types of constraints, namely non-negativity constraints and cut constraints, which we separate over seperately:

- We first check whether non-negativity constraints hold. This is easy since there are $|A|$ many such constraints that we can check one by one explicitly. If we find a violated one, we found a separating hyperplane (namely, the violated constraint) that the oracle can return; else we can assume $y \geq 0$ from now on.
- It remains to check whether one of the constraints $x(\delta^-(S)) \geq 1$ for some $S \subseteq V \setminus \{r\}$, $S \neq \emptyset$ is violated. This can be done via minimum s-t cut computations. For each $v \in V \setminus \{r\}$, we compute a minimum r-v cut $S_v \subseteq V$ in the graph G with respect to capacities y on the arcs. If one of these cuts S_v has a value strictly less than 1, then

$$y(\delta^-(V \setminus S_v)) = y(\delta^+(S_v)) < 1 \ ,$$

hence the constraint corresponding to $V \setminus S_v$ is violated and leads to a separating hyperplane.

Otherwise, if $y(\delta^+(S_v)) \geq 1$ for all $v \in V$, we claim that y satisfies all constraints of P and therefore $y \in P$. To prove this claim, assume that there was a violated constraint, i.e., $y(\delta^-(S)) < 1$ for some $S \subseteq V \setminus \{r\}$, $S \neq \emptyset$. Then, for any $v \in S$, we have

$$y(\delta^+(S_v)) \leq y(\delta^+(V \setminus S)) = y(\delta^-(S)) < 1 \ ,$$

where the first inequality follows from the fact that $V \setminus S$ is an r-v cut, and its cut value is thus at least as large as the value of the cut S_v, which is by definition a minimum r-v cut. But the latter inequality contradicts the assumption that all cuts S_v have values at least 1, and the result follows.

Note that following the above, checking $y \in \mathbb{Q}^A_{\geq 0}$ can be done in time $O(|A|)$. Moreover, separating over the constraints $x(\delta^-(S)) \geq 1$ boils down to solving $O(|V|)$ many minimum r-v cut problems in G. Observe that in these problems, y is used as capacities, and y is rational. By scaling y appropriately, we can reduce to integral capacities without changing the solution structure, and minimum cut problems with integral arc capacities can be solved in strongly polynomial time via the Edmonds-Karp algorithm. Thus overall, the procedure runs in strongly polynomial time.

Problem 5: Two equivalent polyhedral statements (10 points)
Consider the linear program

$$\begin{array}{rrcl} \max & a^\top x & & \\ & Ax & \leq & b \\ & x & \in & \mathbb{R}^n \ . \end{array} \tag{4}$$

The feasible region of the above linear program is precisely P. As P is a non-empty polytope (in particular, this implies that P is non-empty and bounded), the linear program in (4) is feasible and bounded. Let $\nu^* \in \mathbb{R}$ denote its optimal value.

As P is the feasible region of the linear program in (4), Statement (i) is equivalent to $\nu^* \leq \beta$. Consequently, it is enough to prove that $\nu^* \leq \beta$ and statement (ii) are equivalent.

Consider the dual of the linear program in (4), which is

$$\begin{array}{rrcl} \min & b^\top y & & \\ & A^\top y & = & a \\ & y & \in & \mathbb{R}^m_{\geq 0} \ . \end{array}$$

As the primal is feasible and bounded, by LP duality, the dual is feasible and bounded, too, and it has the same objective value ν^* as the primal, hence an optimal dual solution y^* satisfies

$$a = A^\top y^* \qquad \text{and} \qquad b^\top y^* = \nu^* \ .$$

Now assume that $\nu^* \leq \beta$. Let $\lambda = y^*$ and $\Delta = \beta - \nu^*$. Then, by definition, $\lambda \in \mathbb{R}^m_{\geq 0}$ and $\Delta \in \mathbb{R}_{\geq 0}$, and we have

$$a = A^\top y^* = A^\top \lambda \qquad \text{and} \qquad \beta = \nu^* + \Delta = b^\top y^* + \Delta = b^\top \lambda + \Delta \ ,$$

i.e., statement (ii) holds.

For the other direction, assume that statement (ii) holds, i.e., there are $\lambda \in \mathbb{R}^m_{\geq 0}$ and $\Delta \in \mathbb{R}_{\geq 0}$ such that

$$a = A^\top \lambda \qquad \text{and} \qquad \beta = b^\top \lambda + \Delta \ .$$

In other words, λ is feasible for the dual linear program, hence $\nu^* \leq b^\top \lambda = \beta - \Delta \leq \beta$ follows. This finishes the proof of the desired equivalence.