Discrete Optimization
ISyE 6662 - Spring 2023
Homework 5

Instructor: Alejandro Toriello
TA: Filipe Cabral

April 10, 2023.

1. Let $N = \{1, \ldots, n\}$, and let $f : 2^N \to \mathbb{R}$ be a set function with $f(\emptyset) = 0$; define

$$P(f) = \left\{ x \in \mathbb{R}^N : \sum_{i \in S} x_i \leq f(S),\ S \subseteq N \right\}.$$

We showed in class that if f is submodular, the greedy solution is optimal for $\max_{x \in P(f)} \sum_{i \in N} w_i x_i$ with any objective vector $w \in \mathbb{R}_+$. Prove the converse: if the greedy solution is optimal for any objective vector, then f is submodular. Hint: It suffices to consider objectives of the form $w \in \{0, 1\}^N$.

Answer: It is enough to show that for every k the following inequality holds:

$$f(S_{k+1}) - f(S_k) \geq f(S_{k+2}) - f(S_k \cup \{k+2\}),$$

where $S_j := \{1, 2, \ldots, j\}$, since we can always relabel the set N and describe the submodularity property in this form.

(Primal greedy solution): Indeed, let $w_1 = \cdots = w_n = 1$. Then, the greedy optimal solution is

$$\begin{aligned}
x_1^* &= f(S_1), \\
x_2^* &= f(S_2) - x_1^* &&= f(S_2) - f(S_1), \\
&\vdots \\
x_j^* &= f(S_j) - \textstyle\sum_{k=1}^{j-1} x_k^* &&= f(S_j) - f(S_{j-1}), \\
&\vdots \\
x_n^* &= f(S_n) - \textstyle\sum_{k=1}^{n-1} x_k^* &&= f(S_n) - f(S_{n-1}).
\end{aligned}$$

It follows from the feasibility of the primal greedy solution that

$$f(S_k \cup \{k+2\}) \geq \sum_{j \in S_k \cup \{k+2\}} x_j^* = \sum_{j=1}^{k} x_j^* + x_{k+2}^* = f(S_k) + f(S_{k+2}) - f(S_{k+1}).$$

Hence, the function f is submodular.

(Dual greedy solution): Indeed, let $w_1 = \cdots = w_k = 1$, $w_{k+1} = 0$, $w_{k+2} = 1$, and $w_{k+3} = \cdots = w_n = 0$. According to the dual greedy solution, the dual optimal is

$$y_S^* = \begin{cases} w_i - w_{i+1}, & \text{if } S = S_i, \text{ and } i \in N, \\ 0, & \text{otherwise,} \end{cases}$$

for all $S \in 2^N$. This implies that the dual optimal value is $\sum_{i=1}^n f(S_i)(w_i - w_{i+1})$ for the optimization problem

$$\begin{array}{rl} \min_y & \sum_{S\subseteq N} f(S)\cdot y_S \\ \text{s.t.} & \sum_{S\subseteq N: i\in S} y_S \geq w_i, \quad \forall i \in N. \end{array}$$

Below are some remarks about the optimal solution:

- The optimal value can be described as

$$\sum_{i=1}^{n} f(S_i)(w_i - w_{i+1}) = f(S_k) - f(S_{k+1}) + f(S_{k+2}). \tag{1}$$

- The solution defined as

$$y_S = \begin{cases} 1, & \text{if } S = S_k \cup \{k+2\}, \\ 0, & \text{otherwise,} \end{cases}$$

  for all $S \in 2^N$, is feasible and has objective value $f(S_k \cup \{k+2\})$. In particular, we have that

$$f(S_k) - f(S_{k+1}) + f(S_{k+2}) \leq f(S_k \cup \{k+2\}).$$

Hence, f is submodular.

2. Recall the integral polyhedron $P = \{x \in [0,1]^n : \sum_{i=1}^n x_i \leq \mathcal{U}\}$, for some $\mathcal{U} \in \mathbb{N}$. Is P a submodular polyhedron? In other words, does there exist a submodular function f with $P = P(f)$ or $P = P_+(f)$? Justify your answer.

Answer: Consider the function $f(S) := \min(|S|, \mathcal{U})$. We show that $P = P_+(f)$.

Let $x \in P_+(f)$. Then,

- Let $S = N$ and note that $\sum_{i=1}^n x_i = \sum_{i\in N} x_i \leq f(N) \leq \mathcal{U}$.
- Let $S = \{j\}$. Then, $x_j = \sum_{i\in\{j\}} x_i \leq f(\{j\}) \leq 1$, for every $j \in N$.

This implies that $x \in P$. Conversely, let $x \in P$ and let $S \subseteq N$. There there are two options:

- If $|S| \leq \mathcal{U}$ then $\sum_{i\in S} x_i \leq \sum_{i\in S} 1 = |S| = f(S)$.
- If $|S| > \mathcal{U}$ then we have that $\sum_{i\in S} x_i \leq \sum_{i=1}^n x_i \leq |\mathcal{U}| = f(S)$.

This implies that $x \in P_+(f)$.

We just have to prove that f is a submodular function. Indeed, for any subset $S \subseteq N$ and index $i \in N\backslash S$ we have that

$$f(S \cup \{i\}) - f(S) = \begin{cases} 1, & \text{if } |S| < \mathcal{U}, \\ 0, & \text{otherwise.} \end{cases}$$

In particular, the following inequality holds

$$f(S \cup \{i\}) - f(S) \geq f(S \cup \{i,j\}) - f(S \cup \{j\}),$$

for every subset $S \subseteq N$ and indexes $i, j \in N\backslash S$. Hence, f is submodular.

3. Let (N, E) be a connected, undirected network. Consider the integer programming formulation

$$Q_I = \left\{ x \in \mathbb{Z}_+^E, y \in \mathbb{Z}_+^{2(n-2)|E|} \;\middle|\; \begin{array}{ll} x_{ij} = y_{ij}^k + y_{ji}^k, & \{i,j\} \in E,\ k \neq i,j, \\ x_{ij}\mathbb{1}_{\{i,j\}\in E} + \sum_{k\in\delta(i)\backslash j} y_{ik}^j = 1, & i \in N,\ j \in N\backslash i \end{array} \right\}.$$

Here, $\mathbb{1}$ is the indicator function, equal to one when a statement is true and zero otherwise. Note that the y variables are ordered triples and the last set of constraints ranges over ordered pairs.

a) Prove that $\text{proj}_x(Q_I)$ is the set of indicator vectors of spanning trees of (N, E). Hint: Interpret y_{ij}^k as indicating that k is on j's side of the spanning tree.

Answer: Recall that the spanning tree formulation is given by:

$$P_I = \left\{ x \in \mathbb{Z}_+^E : \sum_{e \in E[S]} x_e \leq |S| - 1, \ \emptyset \neq S \subsetneq N; \quad \sum_{e \in E} x_e = n - 1 \right\}.$$

We start by proving that $P_I \subseteq \text{proj}_x(Q_I)$. Let $x \in P_I$ and define $y \in \mathbb{Z}_+^{2(n-2)|E|}$ such that

$$y_{ij}^k = \begin{cases} 1, & \text{if } x_{ij} = 1 \text{ and } k \text{ lies in the same connected component as } j \\ & \text{when we remove the edge } \{i,j\} \text{ from the spanning tree,} \\ 0, & \text{otherwise.} \end{cases}$$

The constraint $x_{ij} = y_{ij}^k + y_{ji}^k$ is satisfied for any $\{i,j\} \in E$ and $k \neq i, j$ because if x_{ij} is equal to 1 then k lies either in the same connected component as j or i but not both at the same time, otherwise the subgraph represented by x would have a cycle. If x_{ij} is 0 then y_{ij}^k and y_{ji}^k are 0 by definition.

The second group of constraints $x_{ij} + \sum_{k \in \delta(i) \setminus j} y_{ik}^j = 1$ is also satisfied for all $i \in N$, and $j \in N \setminus i$. Indeed, if x_{ij} is equal to 1 then y_{ik}^j must equal 0 for all $k \in \delta(i) \setminus \{j\}$, otherwise if some y_{ik}^j equal 1 then x_{ik} is also 1 and j lies in the same connected component as k. This implies that there are two different paths from i to j, i.e., the subgraph represented by x has a cycle. If x_{ij} is 0 then the unique path that connects i and j have length greater than 1, which means that there exists $r \in \delta(i) \setminus \{j\}$ such that x_{ir} is 1 and there is a unique path from r to j. In particular, the variable y_{ir}^j is equal to 1 and all the variables y_{ik}^j for $k \in \delta(i) \setminus \{j, r\}$ are 0. Indeed, if y_{ik}^j is 1 for some $k \in \delta(i) \setminus \{j, r\}$ then the variable x_{ik} equals 1 and j lies in the same connected component as k, which implies that there are two different paths from i to j. Hence (x, y) belongs to Q_I.

Lets prove that $\text{proj}_x(Q_I) \subseteq P_I$. Indeed, consider $(x, y) \in Q_I$. Given $\{i,j\} \in E$, if we sum up the first group of constraints $x_{ij} = y_{ij}^k + y_{ji}^k$ over all $k \neq i, j$ we get

$$(n-2) \cdot x_{ij} = \sum_{k \in N \setminus \{i,j\}} y_{ij}^k + y_{ji}^k,$$

and if we sum it over all $\{i,j\} \in E$ we obtain

$$(n-2) \cdot \sum_{\{i,j\} \in E} x_{ij} = \sum_{\{i,j\} \in E} \sum_{k \in N \setminus \{i,j\}} (y_{ij}^k + y_{ji}^k) = \sum_{i \in N} \sum_{j \in \delta(i)} \sum_{k \in N \setminus \{i,j\}} y_{ij}^k. \tag{2}$$

Now, we sum up the second group of constraints $x_{ij} \mathbb{1}_{\{i,j\} \in E} + \sum_{k \in \delta(i) \setminus j} y_{ik}^j = 1$ over all $i \in N$, and $j \in N \setminus i$:

$$\begin{aligned} n(n-1) &= \sum_{i \in N} \sum_{j \in N \setminus \{i\}} x_{ij} \mathbb{1}_{\{i,j\} \in E} + \sum_{i \in N} \sum_{j \in N \setminus \{i\}} \sum_{k \in \delta(i) \setminus j} y_{ik}^j \\ &= 2 \cdot \sum_{\{i,j\} \in E} x_{ij} + \sum_{i \in N} \sum_{k \in N \setminus \{i\}} \sum_{j \in \delta(i) \setminus k} y_{ij}^k \\ &= 2 \cdot \sum_{\{i,j\} \in E} x_{ij} + \sum_{i \in N} \sum_{j \in \delta(i)} \sum_{k \in N \setminus \{i,j\}} y_{ij}^k, \end{aligned} \tag{3}$$

where the second equality follows from the fact that x_{ij} equals x_{ji} if $\{i,j\} \in E$, and we replace k by j in the triple sum expression. By subtracting Equation (3) from Equation (2), we conclude that $n \cdot \sum_{\{i,j\} \in E} x_{ij} = n(n-1)$, which implies the constraint $\sum_{\{i,j\} \in E} x_{ij} = n - 1$.

We now prove the constraint $\sum_{e\in E[S]} x_e \le |S|-1$ is satisfied for every nonempty set $S \subsetneq N$. The idea is the same, except that we sum over all elements of S instead of all elements of N. Given $\{i,j\} \in E[S]$, if we sum over $k \in S\backslash\{i,j\}$ in the first constraint group we get

$$(|S|-2)x_{ij} = \sum_{k\in S\backslash\{i,j\}} y_{ij}^k + y_{ji}^k,$$

and if we sum it over all $\{i,j\} \in E[S]$ we obtain

$$(|S|-2)\cdot \sum_{\{i,j\}\in E[S]} x_{ij} = \sum_{\{i,j\}\in E[S]} \sum_{k\in S\backslash\{i,j\}} (y_{ij}^k + y_{ji}^k) = \sum_{i\in S} \sum_{j\in\delta(i)\cap S} \sum_{k\in S\backslash\{i,j\}} y_{ij}^k.$$

For the second constraint group, $x_{ij}\mathbb{1}_{\{i,j\}\in E} + \sum_{k\in\delta(i)\backslash j} y_{ik}^j = 1$, we sum over all $i \in S$ and $j \in S\backslash\{i\}$:

$$\begin{aligned}
|S|(|S|-1) &= \sum_{i\in S}\sum_{j\in S\backslash\{i\}} x_{ij}\mathbb{1}_{\{i,j\}\in E} + \sum_{i\in S}\sum_{j\in S\backslash\{i\}}\sum_{k\in\delta(i)\backslash j} y_{ik}^j \\
&= 2\cdot \sum_{\{i,j\}\in E[S]} x_{ij} + \sum_{i\in S}\sum_{k\in S\backslash\{i\}}\sum_{j\in\delta(i)\backslash k} y_{ij}^k \\
&= 2\cdot \sum_{\{i,j\}\in E[S]} x_{ij} + \sum_{i\in S}\sum_{j\in\delta(i)}\sum_{k\in S\backslash\{i,j\}} y_{ij}^k \\
&\ge 2\cdot \sum_{\{i,j\}\in E[S]} x_{ij} + \sum_{i\in S}\sum_{j\in\delta(i)\cap S}\sum_{k\in S\backslash\{i,j\}} y_{ij}^k,
\end{aligned}$$

where the last inequality comes from the fact that the summation over $j \in \delta(i)$ is greater than or equal to the summation over $j \in \delta(i)\cap S$. Therefore, we have that $|S|\cdot\sum_{\{i,j\}\in E[S]} x_{ij} \le |S|(|S|-1)$, which implies the constraint $\sum_{\{i,j\}\in E[S]} x_{ij} \le |S|-1$.

b) Let Q be the linear relaxation of Q_I where we remove integrality constraints. Prove that $\text{proj}_x(Q)$ is the convex hull of indicator vectors of spanning trees.

Answer: In the proof above for the inclusion $\text{proj}_x(Q_I) \subseteq P_I$ we did not use any specific property of the integers. Indeed, the same argument proves that $\text{proj}_x(Q) \subseteq P$ if we replace the set of integers by real numbers. In summary, we have the following properties:

i. $\text{proj}_x(Q) \subseteq P$.

ii. $P = \text{conv}(P_I)$.

iii. $\text{proj}_x(Q_I) = P_I$.

Since the convex hull and the projection operators commute, we have the following identities:

$$P = \text{conv}(P_I) = \text{conv}(\text{proj}_x(Q_I)) = \text{proj}_x(\text{conv}(Q_I)) \subseteq \text{proj}_x(Q).$$

Hence, $P = \text{proj}_x(Q)$.

4. Let $N = \{1,\dots,n\}$. A collection of subsets $\mathcal{L} \subseteq 2^N$ is laminar if $S_1, S_2 \in \mathcal{L}$ implies either $S_1\cap S_2 = \emptyset$, $S_1 \subseteq S_2$ or $S_1 \supseteq S_2$. So at least one of $S_1\cap S_2$, $S_1\backslash S_2$ and $S_2\backslash S_1$ is empty. It may be helpful to think of a laminar family as a rooted tree, where N is the root, the individual elements are leaves, and other sets are intermediate nodes with adjacency determined by containment.

a) For N and laminar family $\mathcal{L}$, let $A \in \{0,1\}^{\mathcal{L}\times N}$ be the incidence matrix of $\mathcal{L}$: $a_{S,i} = 1$ when $i \in S$ and $a_{S,i} = 0$ otherwise. Prove that A is TU.

Answer: Removing a row of A is equivalent to remove a subset of the laminar family and removing a column of A is equivalent to remove an element $i \in N$ from all subsets $S \in \mathcal{L}$. Thus, any square submatrix $B \in \{0,1\}^{k\times k}$ of A is the incidence matrix of a laminar family $\mathcal{L}' \subseteq 2^{N'}$.

Since elementary row operations do not change the determinant of a matrix we can subtract the rows associated to subsets S_1 and S_2 such that $S_1 \subseteq S_2$ and redefine the subset S_2, i.e., $S_2 := S_2 \backslash S_1$. The resulting matrix $\widetilde{B}$ is the incidence matrix of a laminar family where all the subsets are disjoints. Thus, each column of $\widetilde{B}$ has at most one $+1$, which implies that

$$\det B = \det \widetilde{B} \in \{0, \pm 1\}.$$

b) Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two laminar families, and let $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$. Prove that the incidence matrix A of $\mathcal{L}$ is TU.

Answer: Let B be a square submatrix of A. Let $S_1, S_2 \in \mathcal{L}_1$ be two rows of B such that $S_1 \subseteq S_2$. Subtract S_1 from S_2 and redefine the subset S_2, i.e., $S_2 := S_2 \backslash S_1$. Analogously for the subsets of the other laminar family $\mathcal{L}_2$, that is, $S_1', S_2' \in \mathcal{L}_2$ such that $S_1' \subseteq S_2'$. Thus, each column of $\widetilde{B}$ has at most two 1's. Given a column with two 1's, the associated rows are subsets S and S' that belong to different laminar families, that is, $S \in \mathcal{L}_1$ and $S' \in \mathcal{L}_2$.

We expand the determinant for the columns with exactly one 1, which results in a submatrix $\overline{B}$. There are two possibilities for $\overline{B}$:

(i) There exists a column of 0's in $\overline{B}$. This implies that $\det \overline{B} = 0$.

(ii) All the columns of $\overline{B}$ have exactly two 1's. Thus, the matrix $\overline{B}$ is the node-edge incidence matrix of a bipartite graph, where the nodes are the rows of $\overline{B}$ and $\mathcal{L}_1 \cup \mathcal{L}_2$ is the node partition. Hence, $\overline{B}$ is TU.

Therefore,

$$\det B = \det \widetilde{B} = \det B' \in \{0, \pm 1\}.$$

Now consider a submodular function $f : 2^N \to \mathbb{R}$ with $f(\emptyset) = 0$, and recall the submodular polyhedron $P(f)$.

c) Let $A, B \subseteq N$ be two sets with $A \backslash B$, $B \backslash A$, $A \cap B \neq \emptyset$. Show that if the constraints for A and B are binding, then so are the constraints for $A \cap B$ and $A \cup B$. Show that these four binding constraints define a constraint matrix of rank three.

Answer: Indeed, we have the following relations

$$\begin{aligned} f(A) + f(B) &\geq f(A \cap B) + f(A \cup B) \\ &\geq \sum_{i \in A \cap B} x_i + \sum_{j \in A \cup B} x_j \\ &= \sum_{i \in A} x_i + \sum_{j \in B} x_j = f(A) + f(B), \end{aligned}$$

where the first inequality follows from the submodularity of f, the second inequality is a consequence of the feasibility of x, the third relation is an equality obtained by rearranging the summation, and the last equality follows from the hypothesis. Therefore, we get that $f(A) + f(B) = f(A \cap B) + f(A \cup B)$.

This implies that the constraints associated to $A \cap B$ and $A \cup B$ are also binding, otherwise if $f(A \cap B) > \sum_{i \in A \cap B} x_i$ or $f(A \cup B) > \sum_{i \in A \cup B} x_i$ we would get that $f(A \cap B) + f(A \cup B) > f(A) + f(B)$, which is a contradiction.

Consider the matrix induced by the constraints associated to A, B, $A \cap B$, and $A \cup B$:

$$M_{ij} = \begin{cases} 1, & \text{if } i = 1 \text{ and } j \in A \cap B, \\ 1, & \text{if } i = 2 \text{ and } j \in A, \\ 1, & \text{if } i = 3 \text{ and } j \in B, \\ 1, & \text{if } i = 4 \text{ and } j \in A \cup B, \\ 0, & \text{o.w..} \end{cases}$$

Since elementary row operations do not change the rank of a matrix, we can subtract the first row of M from the second and third rows. Thus, we get the matrix

$$M'_{ij} = \begin{cases} 1, & \text{if } i = 1 \text{ and } j \in A \cap B, \\ 1, & \text{if } i = 2 \text{ and } j \in A \backslash B, \\ 1, & \text{if } i = 3 \text{ and } j \in B \backslash A, \\ 1, & \text{if } i = 4 \text{ and } j \in A \cup B, \\ 0, & \text{o.w..} \end{cases}$$

If we subtract the first, second, and third rows of M' from its forth row, we get a zero row vector:

$$\widetilde{M}_{ij} = \begin{cases} 1, & \text{if } i = 1 \text{ and } j \in A \cap B, \\ 1, & \text{if } i = 2 \text{ and } j \in A \backslash B, \\ 1, & \text{if } i = 3 \text{ and } j \in B \backslash A, \\ 0, & \text{if } i = 4, \\ 0, & \text{o.w..} \end{cases}$$

Hence, $\text{rank}(M) = \text{rank}(\widetilde{M}) \leq 3$. The rank of $\widetilde{M}$ is indeed 3 because if we take one element of $A \cap B$, $A \backslash B$, and $B \backslash A$ the submatrix obtained is equivalent to the identity matrix:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

d) For any face $F \subseteq P(f)$, show that you can choose the linearly independent binding constraints defining F so they form a laminar family.

Answer: Let F be a face of $P(f)$. There exist $S_1, \ldots, S_r \subseteq N$ such that the face F is represented as

$$F = \left\{ x \in P(f) : \sum_{i \in S_k} x_i = f(S_k),\ k = 1, \ldots, r \right\}.$$

From item (c), all possible pairwise intersections and unions generated from $S_1, \ldots, S_r$ are implicit equalities for F. We can apply elementary rows operations (Gaussian elimination) to create a set partition of $S_1 \cup \cdots \cup S_k$ whose corresponding equality constraints represent F. In other words, $F = \{x \in P(f) : \sum_{i \in B} x_i = f(B),\ B \in \mathcal{L}\}$ where $\mathcal{L}$ is

$$\mathcal{L} = \left\{ B \in 2^N : B = \left(\bigcap_{k \in \mathcal{A}} S_k \right) \bigcap \left(\bigcap_{k \in \mathcal{A}^{\complement}} S_k^{\complement} \right),\ \mathcal{A} \subseteq \{1, \ldots, r\},\ |\mathcal{A}| \geq 1 \right\}.$$

Moreover, $\mathcal{L}$ is a laminar family because $\mathcal{L}$ is a set partition of $S_1 \cup \cdots \cup S_k$.

e) Conclude that for two integer-valued submodular functions $f, g : 2^N \to \mathbb{Z}$, the polyhedron $P(f) \cap P(g)$ is integral.

Answer: We prove that any face F of $P(f) \cap P(g)$ is integral. From item (d), if F is a face of $P(f) \cap P(g)$ then there are laminar families $\mathcal{L}_1, \mathcal{L}_2$ such that

$$F = \left\{ x \in P(f) \cap P(g) : \begin{array}{l} \sum_{i \in B_1} x_i = f(B_1),\ B_1 \in \mathcal{L}_1, \\ \sum_{i \in B_2} x_i = g(B_2),\ B_2 \in \mathcal{L}_2. \end{array} \right\}.$$

If $f(B)$ differs from $g(B)$ for some $B \in \mathcal{L}_1 \cap \mathcal{L}_2$ then F is empty. Otherwise, the function given by

$$h(B) = \begin{cases} f(B), & \text{if } B \in \mathcal{L}_1, \\ g(B), & \text{if } B \in \mathcal{L}_2 \end{cases}$$

is well-defined. Note that we can describe the face F as

$$F = \left\{ x \in P(f) \cap P(g) : \sum_{i \in B} x_i = h(B),\ B \in \mathcal{L}_1 \cup \mathcal{L}_2 \right\}.$$

Since the constraint matrix of $\sum_{i \in B} x_i = h(B)$ for all $B \in \mathcal{L}_1 \cup \mathcal{L}_2$ is the indicator matrix of the union of two laminar families, we have from item (b) that it is totally unimodular (TU). Hence, F is integral.