Short Course on Submodular Functions

S. Thomas McCormick Maurice Queyranne*

June 26, 2013

Abstract

Functions that assign values to subsets of a finite ground set are called *set functions*. A particular class of set functions are the *submodular* functions. Submodular functions are interesting because they have many applications in a wide range of fields, and because optimization problems involving submodular functions are typically easy to solve.

There are many important applied problems that can be formulated as minimizing or maximizing a submodular function, perhaps subject to some side constraints. In particular, Submodular Function Minimization (SFMin) asks for a subset with minimum value, and Submodular Function Maximization (SFMax) asks for a subset with maximum value. Here are some homework problems related to this material, some of which were referenced in the course lectures.

1 Problems

This is an updated and corrected version of what was handed out at the short course, with answers to all the problems. Answers are given in sans-serif font. **If you have any corrections, questions, or further comments about any of this, please email me at `Tom.McCormick@sauder.ubc.ca`.**

Question 1. In class we saw two different definitions of submodularity. First, the "factory" definition:

$$\forall S \subset T \subset T + e,\ f(T+e) - f(T) \le f(S+e) - f(S). \tag{1}$$

Second, the "classic" definition:

$$\text{for all } S, T \subseteq E,\ f(S) + f(T) \ge f(S \cup T) + f(S \cap T). \tag{2}$$

(a) Prove that definitions (1) and (2) are equivalent.

To show that (2) implies (1), apply (2) to the sets $X = S+e$ and $Y = T$ to get $f(S+e)+f(T) \ge f((S+e) \cup T) + f((S+e) \cap T) = f(T+e) + f(S)$, which is equivalent to (1).

To show that (1) implies (2), first re-write (1) as $f(S+e) - f(T+e) \ge f(S) - f(T)$ for $S \subset T \subset T+e$. Now, enumerate the elements of $Y - X$ as $e_1, e_2, \ldots, e_k$ and note that, for $i < k$,

*Sauder School of Business, University of British Columbia, Vancouver, BC V6T 1Z2 Canada. Supported by NSERC Operating Grants.

$[(X \cap Y) \cup \{e_1, e_2, \ldots, e_i\}] \subset [X \cup \{e_1, e_2, \ldots, e_i\}] \subset [X \cup \{e_1, e_2, \ldots, e_i\}] + e_{i+1}$, so the re-written (1) implies that

$$\begin{aligned} f(X \cap Y) - f(X) &\leq f((X \cap Y) + e_1) - f(X + e_1) \\ &\leq f((X \cap Y) \cup \{e_1, e_2\}) - f(X \cup \{e_1, e_2\}) \\ &\ldots \\ &\leq f((X \cap Y) \cup \{e_1, e_2, \ldots, e_k\}) - f(X \cup \{e_1, e_2, \ldots, e_k\}) \\ &= f(Y) - f(X \cup Y), \end{aligned}$$

and this is equivalent to (2).

(b) Here is an apparently very weak special case of definition (1) of submodularity. For every $S \subset E$ and every $e, g \notin S$,

$$f(S + e) - f(S) \geq f(S + e + g) - f(S + g).$$

Prove that f is submodular if and only if this weaker condition is true.

If suffices to show that for $S \subseteq T \subseteq E$ and $R \subseteq E$ such that $T \cap R = \emptyset$, that $f(T \cup R) - f(S \cup R) \leq f(T) - f(S)$. Now enumerate R as e_1, e_2, $\ldots$, e_k. Then

$$\begin{aligned} f(T \cup R) - f(S \cup R) &= f(T + e_1 + e_2 \cdots + e_k) - f(S + e_1 + e_2 \cdots + e_k) \\ &\leq f(T + e_1 + e_2 \cdots + e_{k-1}) - f(S + e_1 + e_2 \cdots + e_{k-1}) \\ &\leq \ldots \\ &\leq f(T + e_1) - f(S + e_1) \leq f(T) - f(S). \end{aligned}$$

Question 2. Suppose that f is a set function on E with $f(\emptyset) = 0$.

(a) Prove that f is modular iff there exists a vector $w \in \mathbb{R}^E$ such that $f(S) = w(S) \equiv \sum_{e \in S} w_e$ for all $S \subseteq E$. (This justifies us treating modular set functions as vectors. Modular set functions are, roughly speaking, the set function analogue of linear functions.)

($w \in \mathbb{R}^E \Rightarrow w(S)$ is modular with $w(\emptyset) = 0$): Clearly $w(\emptyset) = 0$. We want to verify that $w(S) + w(T) = w(S \cap T) + w(S \cup T)$. If $e \in S \cap T$ then it contributes $2w_e$ to both sides. If $e \in (S - T) \cup (T - S)$, then it contributes w_e to both sides. If $e \in N - (S \cup T)$ then it contributes 0 to both sides. Since we have the same contribution to both sides in all cases, the identity is verified.

(f modular with $f(\emptyset) = 0 \Rightarrow$ there is $w \in \mathbb{R}^E$ such that $f(S) = w(S)$): Set $w_e = f(\{e\})$. Modularity of f and $f(\emptyset) = 0$ imply that $f(S) + w_e = f(S) + f(e) = f(S + e) + f(\emptyset) = f(S + e)$, so induction yields that $f(S) = w(S)$ for all $S \subseteq E$.

(b) Let $G = (N, A)$ be a directed graph, and let $x \in \mathbb{R}^A$ be a vector of flows. Then for $S \subseteq N$ the set function $f(S) = x(\delta^-(S)) - x(\delta^+(S))$ is the net flow into node subset S. Prove that $f(S)$ is a modular set function with $f(\emptyset) = 0$.

Clearly $f(\emptyset) = 0$. From the proof of #3 $x(\delta^+(S)) + x(\delta^+(T)) = x(\delta^+(S \cap T)) + x(\delta^+(S \cup T)) + x(\delta^+(S - T, T - S)) + x(\delta^+(T - S, S - T))$ and $x(\delta^-(S)) + x(\delta^-(T)) = x(\delta^-(S \cap T)) + x(\delta^-(S \cup T)) + x(\delta^-(S - T, T - S)) + x(\delta^-(T - S, S - T))$. Noting that $\delta^-(T - S, S - T) = \delta^+(S - T, T - S)$, when we subtract the second from the first we get $(x(\delta^+(S)) - x(\delta^-(S))) + (x(\delta^+(T)) - x(\delta^-(T))) = (x(\delta^+(S \cap T)) - x(\delta^-(S \cap T))) + (x(\delta^+(S \cup T)) - x(\delta^-(S \cup T)))$, so $f(S)$ is modular.

Question 3. Let $G = (N, A)$ be a directed graph. For $S \subseteq N$ define $\delta(S) = \delta^+(S) \cup \delta^-(S)$. Given $l, u, w \in \mathbb{R}^A$ with $w \geq 0$ and $l \leq u$, for $S \subseteq N$ define $f_1(S) = w(\delta(S))$, $f_2(S) = w(\delta^+(S))$, and $f_3(S) = u(\delta^+(S)) - l(\delta^-(S))$.

(a) Prove that f_k is submodular on 2^N for $k = 1, 2, 3$. Give examples showing that f_k is not submodular if $w \not\geq 0$ or $u \not\geq l$, $k = 1, 2, 3$. (Note that $S \subseteq T \subseteq N \not\Rightarrow f_k(S) \leq f_k(T)$, $k = 1, 2, 3$, i.e., none of these functions is necessarily monotone, so they are not polymatroid rank functions.)

For $X, Y \subseteq N$, define $\delta^+(X, Y) = \{i \to j \in A \mid i \in X,\ j \in Y\}$ and $\delta(X, Y) = \delta^+(X, Y) \cup \delta^+(Y, X)$. I claim that $w(\delta(S)) + w(\delta(T)) = w(\delta(S \cap T)) + w(\delta(S \cup T)) + w(\delta(S - T, T - S))$. This is easy to see by considering the 16 cases of an arc starting in one of the four sets $S \cap T$, $S - T$, $T - S$, and $N - (S \cup T)$ and ending in any of the four sets, and noting that each arc a contributes exactly 0, 1, or 2 w_a to both sides of the identity. Therefore if $w \geq 0$, f_1 is submodular. Let $N = \{1, 2, 3, 4\}$, $A = \{1 \to 2, 1 \to 3, 2 \to 3, 2 \to 4, 3 \to 4\}$, $w = (0, 0, -1, 0, 0)$, $S = \{1, 2\}$, and $T = \{1, 3\}$. Then $\delta(S) = \{1 \to 3, 2 \to 3, 2 \to 4\}$, $\delta(T) = \{1 \to 2, 2 \to 3, 3 \to 4\}$, $\delta(S \cap T) = \{1 \to 2, 1 \to 3\}$, and $\delta(S \cup T) = \{2 \to 4, 3 \to 4\}$. Thus $w(\delta(S)) + w(\delta(T)) = -2 \not\geq 0 = w(\delta(S \cap T)) + w(\delta(S \cup T))$, so f_1 is not submodular when $w \not\geq 0$.

The same type of computation shows that $w(\delta^+(S)) + w(\delta^+(T)) = w(\delta^+(S \cap T)) + w(\delta^+(S \cup T)) + w(\delta^+(S - T, T - S)) + w(\delta^+(T - S, S - T))$, implying that if $w \geq 0$ that f_2 is submodular. Consider the same counterexample as above. The only change is that $\delta(T)$ is now $\{1 \to 2, 3 \to 4\}$, so that $w(\delta(S)) + w(\delta(T)) = -1 \not\geq 0 = w(\delta(S \cap T)) + w(\delta(S \cup T))$, so f_2 is not submodular when $w \not\geq 0$.

The proof of Question 1 shows that $(u(\delta^+(S)) - l(\delta^-(S))) + (u(\delta^+(T)) - l(\delta^-(T))) = (u(\delta^+(S \cap T)) - l(\delta^-(S \cap T))) + (u(\delta^+(S \cup T)) - l(\delta^-(S \cup T))) + (u - l)(\delta(S - T, T - S))$, so when $u \geq l$ $f_3(S)$ is submodular. Consider the same counterexample as above, but with $l = 0$, $u = (0, 0, -1, 0, 0)$. Now the key term $(u - l)(\delta(S - T, T - S)) = -1$ (coming from $(u - l)_{23} = -1$), so f_3 is not submodular when $u \not\geq l$.

Question 4. Suppose that f is a set function on 2^E. We say that f is a *cardinality* set function if there is some function $g : \{1, 2, \ldots, |E|\} \to \mathbb{R}$ such that $f(S) = g(|S|)$, i.e., if the value of $f(S)$ depends only on the size of S. Prove that a cardinality set function is submodular iff g is concave. [This point of view is consistent with the "decreasing returns to scale" factory definition of submodularity.]

Since we care only about the values of g at integer points, concavity is equivalent to requiring that $g(i) - g(i-1) \geq g(i+1) - g(i)$ for all i. This translates into $f(S+e) - f(S) \geq f(S+e+h) - f(S+h)$ whenever $e, h \notin S$, which is equivalent to submodularity by Question 1 (b).

Question 5. If f is a submodular function on E with $f(\emptyset) = 0$, then its associated *submodular polyhedron* is $P(f) = \{x \in \mathbb{R}^E \mid x(S) \leq f(S)\ \forall S \subseteq E\}$. If $x \in P(f)$, we call $S \subseteq E$ *x-tight* if $x(S) = f(S)$.

(a) Prove that the union and intersection of x-tight sets is x-tight.

S, T x-tight, submodularity of f, and modularity and feasibility of x imply that

$$\begin{aligned} x(S \cap T) + x(S \cup T) &= x(S) + x(T) \\ &= f(S) + f(T) \\ &\geq f(S \cap T) + f(S \cup T) \\ &\geq x(S \cap T) + x(S \cup T), \end{aligned}$$

implying that $x(S \cap T) + x(S \cup T) = f(S \cap T) + f(S \cup T)$, and then feasibility of x implies that $x(S \cap T) = f(S \cap T)$ and $x(S \cup T) = f(S \cup T)$.

(b) Suppose that $x \in P(f)$ and $S \subset E$, and that for all $i \in S$ and $j \notin S$ we know that there is an x-tight set S_{ij} containing i but not j. Prove that S is x-tight.

By (a), the set $B = \cup_{i \in S} \cap_{j \notin S} S_{ij}$ is also x-tight. Since $\cap_{j \notin S} S_{ij}$ contains i but no elements not in S, B includes all elements of S and no elements not in S, i.e., $B = S$.

Question 6. Let $G = (N, A)$ be a max flow network with return arc $t \to s$. Define $S = \{s \to i \in A\}$, i.e., the subset of arcs with tail s. For $F \subseteq S$ define $v(F)$ to be the max flow value in the network with capacities u'_a defined by $u'_a = 0$ for $a \in S - F$, and $u'_a = u_a$ otherwise, i.e., we set the capacities of arcs in $S - F$ to zero and otherwise leave the capacities alone. Thus $v(\emptyset) = 0$ and $v(S)$ is the optimal max flow value in the original G.

(a) Prove that $F_1 \subseteq F_2 \subseteq S$ implies that $v(F_1) \leq v(F_2)$, i.e., that $v(F)$ is monotone.

Let $x(F)$ be a max flow corresponding to $v(F)$, so that $\text{val}(x(F)) = v(F)$. Since $F_1 \subseteq F_2$, $x(F_1)$ is a feasible flow for the F_2 network, so $v(F_2) \geq \text{val}(x(F_1)) = v(F_2)$.

(b) Prove that $v(F)$ is submodular.

We need to show that $v(F_1) + v(F_2) \geq v(F_1 \cup F_2) + v(F_1 \cap F_2)$. Let $x(F_1 \cap F_2)$ be a max flow corresponding to $v(F_1 \cap F_2)$. Use any augmenting path algorithm starting from $x(F_1 \cap F_2)$ to extend it to a max flow $x(F_1 \cup F_2)$ corresponding to $v(F_1 \cup F_2)$. Note that for $s \to i \in F_1 \cap F_2$ we must have $x(F_1 \cup F_2)_{si} = x(F_1 \cap F_2)_{si}$, since the optimality of $x(F_1 \cap F_2)$ implies that no augmenting path in the $F_1 \cup F_2$ network can use any arc of $F_1 \cap F_2$. This implies that

$$\begin{aligned} \sum_{s \to i \in F_1} x(F_1 \cup F_2)_{si} + \sum_{s \to i \in F_2} x(F_1 \cup F_2)_{si} &= \sum_{s \to i \in F_1 \cup F_2} x(F_1 \cup F_2)_{si} + \sum_{s \to i \in F_1 \cap F_2} x(F_1 \cup F_2)_{si} \\ &= v(F_1 \cup F_2) + v(F_1 \cap F_2). \end{aligned}$$

Now conformal decomposition implies that we can transform $x(F_1 \cup F_2)$ into a flow $x'(F_1)$ feasible for the F_1 network such that $x(F_1 \cup F_2)_{si} = x'(F_1)_{si}$ for all $s \to i \in F_1$, and similarly we can transform $x(F_1 \cup F_2)$ into a flow $x'(F_2)$ feasible for the F_2 network such that $x(F_1 \cup F_2)_{si} = x'(F_2)_{si}$ for all $s \to i \in F_2$. Note that

$$\sum_{s \to i \in F_1} x'(F_1)_{si} + \sum_{s \to i \in F_2} x'(F_2)_{si} = \sum_{s \to i \in F_1 \cap F_2} x(F_1 \cap F_2)_{si} + \sum_{s \to i \in F_1 \cup F_2} x(F_1 \cup F_2)_{si}$$

since every $s \to i$ arc is counted exactly the same number of times (0, 1, or 2) on both sides.

Then $v(F_1) \geq \sum_{s\to i\in F_1} x'(F_1)_{si}$ and $v(F_2) \geq \sum_{s\to i\in F_2} x'(F_2)_{si}$, so

$$\begin{aligned} v(F_1)+v(F_2) &\geq \sum_{s\to i\in F_1} x'(F_1)_{si} + \sum_{s\to i\in F_2} x'(F_2)_{si} \\ &= \sum_{s\to i\in F_1\cup F_2} x(F_1\cap F_2)_{si} + \sum_{s\to i\in F_1\cup F_2} x(F_1\cup F_2)_{si} \\ &= v(F_1\cup F_2)+v(F_1\cap F_2). \end{aligned}$$

Note that (a) and (b) imply that $Q = \{y \in \mathbb{R}^S \mid y(F) \leq v(F)\ \forall F \subseteq S\}$ is a *polymatroid*.

The *flow polyhedron* is $P(G) = \{x \in \mathbb{R}^A \mid x \text{ is a feasible flow in } G\}$. If $X \subseteq A$ is a subset of arcs, then the *projection* of $P(G)$ onto X is $P(X) = \{y \in \mathbb{R}^X \mid \exists x \in P(G) \text{ s.t. } y_a = x_a\ \forall a \in X\}$, i.e., the linear algebraic projection of $P(G)$ onto the components in X.

(c) We would like to show that $Q = P(S)$, i.e., that the flow polyhedron projected onto the arcs with tail s is a polymatroid. The only thing left to prove is that every $q \in Q$ also belongs to $P(S)$, i.e., that if $q \in Q$ then there is a feasible flow x in G whose projection on S is q. Prove this.

Let $G(q)$ be the max flow network with u_{si} replaced by q_{si} for all $s \to i \in S$, let $x(q)$ be a max flow in $G(q)$, and define $S(q)$ to be a corresponding min cut. If $x(q)$ saturates all arcs in S we are done, so to get a contradiction assume that there is at least one $s \to i \in S$ with $x(q)_{si} < q_{si}$. This implies that the set $I = \{s \to i \in S \mid i \in S(q)\}$ is non-empty.

Now conformally decompose $x(q)$ into flows on s–t paths. Note that any path P whose first arc $s \to j$ is not in I cannot contain any other arc of $\delta^+(S(q))$ besides $s \to j$ (since it would then have to also contain an arc of $\delta^-(S(q))$, and $x(q)$ is zero on all such arcs). Thus when we subtract out the flow of all such paths from $x(q)$, we get a new flow $x(I)$ which still satisfies complementary slackness with $S(q)$, so that $x(I)$ is a max flow in the network corresponding to $v(I)$. But $\text{val}(x(I)) < \sum_{s\to i\in I} x(q)_{si} \leq \sum_{s\to i\in I} q_{si} \leq v(I)$ (by feasibility of q for Q), contradicting that $x(I)$ is a max flow for the $v(I)$ network.

(d) Note that $S = \delta^+(\{s\})$. This makes it tempting to conjecture that if $C = \delta^+(T)$ for some s–t cut T, then $P(C)$ is also a polymatroid. Prove that this is true or give a counterexample showing that this conjecture is false.

The conjecture is false: Consider the network in Figure 1 with $N = \{s, 1, 2, t\}$, $A = \{s \to 1, s \to 2, 1 \to t, 2 \to t, 1 \to 2\}$, and $u = (4, 10, 10, 4, 1)$. Let $T = \{s, 1\}$ so that $C = \{s \to 2, 1 \to 2, 1 \to t\}$. Put $X = \{s \to 2, 1 \to 2\}$ and $Y = \{1 \to 2, 1 \to t\}$, so that $X \cap Y = \{1 \to 2\}$ and $X \cup Y = \{s \to 2, 1 \to 2, 1 \to t\}$. Then $v(X) = v(Y) = 4$, $v(X \cap Y) = 1$, and $v(X \cup Y) = 8$, so that $v(X) + v(Y) = 4 + 4 \not\geq 1 + 8 = v(X \cap Y) + v(X \cup Y)$. Thus v is not submodular, so $P(C)$ is not a polymatroid.

Question 7. Suppose that f is a submodular set function on E with polyhedron $P(f)$. A *base* of P is a point $x \in P$ with $\mathbb{1}^T x = f(E)$. (This specializes to the usual definition that a base in a matroid $(E, \mathcal{I})$ is a subset $B \subseteq E$ such that $B \in \mathcal{I}$ and $|B| = f(E)$.) Prove that if x and y are two bases of $P(f)$ and $x_e > y_e$, then there exists $g \in E$ with $x_g < y_g$, and $\varepsilon > 0$, such that $x + \varepsilon(\chi_g - \chi_e)$ and $y - \varepsilon(\chi_g - \chi_e)$ (the same g in both cases) are both also bases. This is called the *Base Exchange Property*.

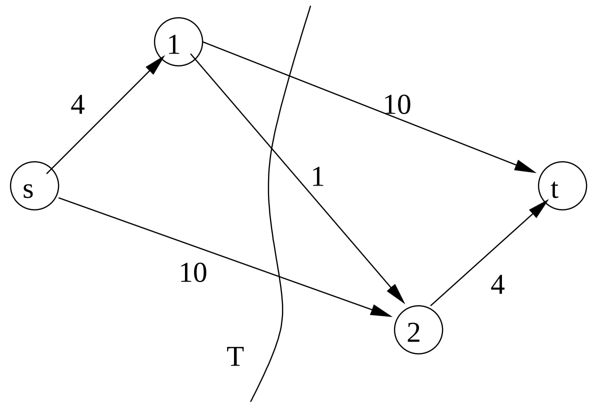

Figure 1: Counterexample

Define $E^+ = \{g \in E \mid x_g > y_g\}$ and $E^- = \{g \in E \mid x_g < y_g\}$. Now fix our attention on the given element $e \in E^+$. Note that $x + \varepsilon(\chi_g - \chi_e) \in P(f)$ for sufficiently small $\varepsilon > 0$ iff there is no x-tight set containing g but not e. Define $X_{\text{bad}} = \{g \in E^- \mid \exists\ x\text{-tight } G \text{ containing } g \text{ but not } e\}$ (the set of elements of E^- that cannot be swapped with e in x). Similarly, $y - \varepsilon(\chi_g - \chi_e) \in P(f)$ for sufficiently small $\varepsilon > 0$ iff there is no y-tight set containing e but not g. Define $Y_{\text{bad}} = \{g \in E^- \mid \exists\ y\text{-tight } G \text{ containing } e \text{ but not } g\}$ (the set of elements of E^- that cannot be swapped with e in y). These definitions imply that if there is an $g \in E^- - (X_{\text{bad}} \cup Y_{\text{bad}})$, then we are done.

Assume to the contrary that $X_{\text{bad}} \cup Y_{\text{bad}} = E^-$. For each $g \in X_{\text{bad}}$ there is an x-tight set X_g containing g but not e, and for each $g \in Y_{\text{bad}}$ there is a y-tight set Y_g containing e but not g. Then by #21 (a) $X^* = \cup_{g \in X_{\text{bad}}} X_g$ is an x-tight set containing X_{bad} but not e, and $Y^* = \cap_{g \in Y_{\text{bad}}} Y_g$ is a y-tight set containing e but $Y^* \cap Y_{\text{bad}} = \emptyset$. Now tightness of X^* and Y^*, submodularity of f, and feasibility of x and y imply that

$$\begin{aligned} x(X^*) + y(Y^*) &= f(X^*) + f(Y^*) \\ &\geq f(X^* \cap Y^*) + f(X^* \cup Y^*) \\ &\geq y(X^* \cap Y^*) + x(X^* \cup Y^*). \end{aligned}$$

This plus modularity of x and y implies that $x(Y^* - X^*) \leq y(Y^* - X^*)$. Now $X_{\text{bad}} \cup Y_{\text{bad}} = E^-$ and $Y^* \cap Y_{\text{bad}} = \emptyset$ imply that $Y^* \cap E^- \subseteq X_{\text{bad}} - Y_{\text{bad}} \subseteq X_{\text{bad}} \subseteq X^*$, or $Y^* \cap E^- \subseteq X^*$, implying that $(Y^* - X^*) \cap E^- = \emptyset$. But $x_g \geq y_g$ on $E - E^-$, $e \in Y^* - X^*$, and $x_e > y_e$ imply that $x(Y^* - X^*) > y(Y^* - X^*)$, a contradiction.

Question 8. Suppose that in a Max Flow / Min Cut network, instead of computing a Min Cut, we wanted to compute a cut S solving $\min_{\emptyset \neq S \subseteq N} \text{cap}(S)/|S|$, a *min ratio* cut. A standard way of dealing with such problems is to have a parameter ρ representing the value of the ratio $\text{cap}(S)/|S|$. Here we consider an extension of this that we call *Submodular Function Mean Minimization (SFMMin)*: We are given a submodular function f on E and want to solve

$$\min_{\emptyset \neq S \subseteq E} f(S)/|S|$$

(note that the optimal value of this might be positive, negative, or zero). Consider these LPs:

$$\begin{array}{rcll} \min \sum_{S\subseteq E} f(S)\pi_S & & & \\ \sum_{S\ni e} \pi_S - \sigma_e & = & 0 & \text{for all } e \in E \\ \sum_{e\in E} \sigma_e & = & 1 & \\ \sigma_e & \ge & 0 & \text{for all } e \in E \\ \pi_S & \ge & 0 & \text{for all } S \subseteq E. \end{array} \qquad \begin{array}{rcll} \max \rho & & & \\ y(S) & \le & f(S) & \text{for all } S \subseteq E \\ \rho & \le & y_e & \text{for all } e \in E, \\ y_e & & \text{free} & \text{for all } e \in E \\ \rho & & \text{free} & \end{array}$$

(a) Argue that these dual linear programs formulate SFMMin.

By standard arguments we can see that there must be an optimal solution to the primal with $\pi_S = 1$ for exactly one $S \subseteq E$, call it S^*, and then $\sigma = \chi(S^*)$ so that $\sum_{e\in E} \sigma_e = |S^*|$. By the usual argument we could then move the normalizing constraint $\sum_{e\in E} \sigma_e = 1$ into the denominator of the objective to get the objective $\min_{S\subseteq E} f(S)/|S|$, which is what we want.

(b) Use complementary slackness between your two LPs to get necessary and sufficient conditions for optimal solutions.

Let S^*, π^*, σ^*, y^*, and ρ^* be optimal. Since $\pi^*_{S^*} > 0$ we have $y^*(S^*)$ must equal $f(S^*)$. If $\rho^* < y^*_e$ then we must have $\sigma^*_e = 0$, i.e., $e \notin S^*$. If $y^*(S) < f(S)$ then we must have $\pi^*_S = 0$, i.e., $S \neq S^*$. If $\sigma^*_e > 0$ (i.e., $e \in S^*$) then we must have $\rho^* = y^*_e$. [Notice that $f(S^*) = y^*(S^*)$ and $y^*_e = \rho^*$ for all $e \in S^*$ implies that $y^*(S^*) = \rho^*|S^*| = f(S^*)$, or $\rho^* = f(S^*)/|S^*|$, the optimal objective value.]

(c) Suppose that we are running some hypothetical SFMin-like algorithm to solve this problem where we represent our current point $y \in B(f)$ as $y = \sum_i \lambda_i v^i$ with $\sum_i \lambda_i = 1$, where each v^i is a vertex associated with linear order $\prec_i$. How would we recognize optimality in such an algorithm?

Let $\rho = \min_e y_e$, and $S = \{e \mid y_e = \rho\}$. Suppose that in each v^i all elements of S come before all elements of $E - S$. Then by how Greedy works to generate v^i, we have that $v^i(S) = f(S)$, and so $y(S) = f(S)$. Since $y_e > \rho$ implies that $e \notin S$ and $e \in S$ implies that $y_e = \rho$, this y, ρ, and S satisfy the complementary slackness from (b), and so are optimal. [This give a rudimentary idea for an algorithm: whenever there exists some j, $k \in E$ such that $\rho = y_j < y_k$ and for some i we have $k \prec_i j$, then move k rightwards (which tends to decrease y_k) and j leftwards (which tends to increase y_j) in $\prec_i$ until we can find no such pair, and then we must be optimal.]

(d) Suppose that both S and T solve the min ratio problem. Prove that $S \cup T$ also solves it, and if $S \cap T \neq \emptyset$, $S \cap T$ also solves it.

Denote $\rho = f(S_1)/|S_1| = f(S_2)/|S_2|$. Submodularity implies that $f(S_1 \cup S_2) \le f(S_1) + f(S_2) - f(S_1 \cap S_2)$, and modularity that $|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2|$. If $S_1 \cap S_2 = \emptyset$, then $f(S_1 \cup S_2)/|S_1 \cup S_2| \le (f(S_1) + f(S_2))/(|S_1| + |S_2|) = (\rho|S_1| + \rho|S_2|)/(|S_1| + |S_2|) = \rho$. Since ρ is the minimum possible ratio, we must have that $f(S_1 \cup S_2)/|S_1 \cup S_2| = \rho$, and so $S_1 \cup S_2$ is also optimal for SFMMin.

Now assume that $S_1 \cap S_2 \neq \emptyset$. To get a contradiction, assume that $S_1 \cap S_2$ is not optimal for SFMMin, i.e., that $f(S_1 \cap S_2)/|S_1 \cap S_2| > \rho$. Then $f(S_1 \cup S_2)/|S_1 \cup S_2| \le (f(S_1) + f(S_2) - f(S_1 \cap S_2))/(|S_1| + |S_2| - |S_1 \cap S_2|) < (\rho|S_1| + \rho|S_2| - \rho|S_1 \cap S_2|)/(|S_1| + |S_2| - |S_1 \cap S_2|) = \rho$, contradicting that ρ is the minimum possible ratio. Thus we must have that $f(S_1 \cap S_2)/|S_1 \cap S_2| = f(S_1 \cup S_2)/|S_1 \cup S_2| = \rho$, and so both $S_1 \cup S_2$ and $S_1 \cap S_2$ are also optimal for SFMMin.

Question 9. All existing combinatorial SFMin algorithms prove that the current point x belongs to the base polytope $B(f)$ via finding vertices $v^j \in B(f)$ and expressing x as the convex combination $x = \sum_j \lambda_j v^j$ with $\sum_j \lambda_j = 1$ and $\lambda \geq 0$. This is unpleasant because even if f is integer-valued, the λ_j are typically quite fractional, and it makes the SFMin algorithms have to do linear algebra to keep the number of v^j small.

Here is the start of a different idea (due to Fujishige) for proving that $x \in B(f)$ called *combinatorial hull*. We use tildes to represent the simple projection obtained by dropping the first coordinate, so that $\tilde{E} = E - \{1\}$ and $\tilde{x} = (x_2, x_3, \ldots, x_n)$. Suppose that we know that $x(E) = f(E)$ (which is easily checkable), and that we have points $y, z \in B(f)$ (e.g., perhaps y and z are vertices of $B(f)$) such that $\tilde{y} \leq \tilde{x} \leq \tilde{z}$, i.e., (the projection of) x is contained in the box defined by (the projections of) y and z.

(a) Prove that this implies that $x \in B(f)$. (Note that when x and f are integral, this representation involves only integers, and only addition, subtraction, and comparison.)

It suffices to prove that $x(S) \leq f(S)$ for all $S \subset E$. Suppose that $1 \notin S$, so that $S = \tilde{S}$. Then $x(S) = \tilde{x}(\tilde{S}) \leq \tilde{z}(\tilde{S}) = z(S) \leq f(S)$.

Now suppose that $1 \in S$ so that $\tilde{S} = S - \{1\}$. Note that $x(E) = f(E)$ implies that $x_1 = f(E) - \tilde{x}(\tilde{E})$. Thus $x(S) = \tilde{x}(\tilde{S}) + x_1 = \tilde{x}(\tilde{S}) + f(E) - \tilde{x}(\tilde{E}) = f(E) - \tilde{x}(\tilde{E} - \tilde{S}) \leq f(E) - \tilde{y}(\tilde{E} - \tilde{S}) = \tilde{y}(\tilde{S}) + f(E) - \tilde{y}(\tilde{E}) = y(\tilde{S}) + y_1 = y(S) \leq f(S)$.

Open problem: This procedure can be iterated to find more complicated proofs that a point belongs to $B(f)$ involving more than two vertices of $B(f)$ (and possibly involving projecting out other coordinates). What is the "Carathéodory number" of such a representation, i.e., the smallest k such that any $x \in B(f)$ has a combinatorial hull representation using at most k vertices of $B(f)$? There is a rather trivial bound one can get of 2^{n-1} via greedily increasing x along some coordinate until we hit the boundary, which is one dimension smaller to get two points; for each of these, repeat to get two more points in one smaller dimension, etc. The obvious conjecture here is that some polynomial bound suffices. One way to go at this: what if we have a "too-large" such representation? How can we go about reducing the number of vertices used, as we do with convex combinations?

Suppose that we could solve this open problem and had an SFMin algorithm using combinatorial hull instead of convex hull. An important part of an SFMin algorithm is the fact that if a point x is represented via the convex combination of vertices v^j of a base polytope as $x = \sum_{j \in J} \lambda_j v^j$ with $\sum_{j \in J} \lambda_j = 1$ and $\lambda_j > 0$ for all $j \in J$ (often we'd instead write $\lambda_j \geq 0$, but we can trivially drop any j with $\lambda_j = 0$ from J), then $S \subseteq E$ is tight for x iff S is tight for each v^j.

(b) Prove that when x is in the combinatorial hull of y and z, any S tight for both y and z is tight also for x.

Suppose that S is tight for y and z, so that $f(S) = y(S) = z(S)$. If $1 \notin S$, then $f(S) = y(S) = \tilde{y}(\tilde{S}) \leq \tilde{x}(\tilde{S}) = x(S) = \tilde{x}(\tilde{S}) \leq \tilde{z}(\tilde{S}) = z(S) = f(S)$. Hence we have equality everywhere, and so $x(S) = f(S)$. If instead $1 \in S$, then $f(S) = z(S) = z(\tilde{S}) + z_1 = \tilde{z}(\tilde{S}) + f(E) - \tilde{z}(\tilde{E}) = f(E) - \tilde{z}(\tilde{E} - \tilde{S}) \leq f(E) - \tilde{x}(\tilde{E} - \tilde{S}) = \tilde{x}(\tilde{S}) + f(E) - \tilde{x}(\tilde{E}) = \tilde{x}(\tilde{S}) + x_1 = x(S) = \tilde{x}(\tilde{S}) + x_1 = \tilde{x}(\tilde{S}) + f(E) - \tilde{x}(\tilde{E}) = f(E) - \tilde{x}(\tilde{E} - \tilde{S}) \leq f(E) - \tilde{y}(\tilde{E} - \tilde{S}) = \tilde{y}(\tilde{S}) + f(E) - \tilde{y}(\tilde{E}) = y(\tilde{S}) + y_1 = y(S) = f(S)$. Again we thus have equality everywhere, and so $x(S) = f(S)$.

(c) Construct a counterexample showing that we could have that S is tight for x and z, but not tight for y.

Put $E = \{1, 2\}$, $f(\{1\}) = f(\{2\}) = f(\{1, 2\}) = 1$, and note that f is submodular. Then set $x = z = (0, 1)$ and $y = (1, 0)$, so that y and z are vertices of $B(f)$ and $\tilde{y} = (0) \leq \tilde{x} = (1) \leq \tilde{z} = (1)$, so that x is in the combinatorial hull of y and z. Then $S = \{2\}$ is tight for x and z, but not for y.

This attempt to show that combinatorial hull has the same property as convex hull does not have any condition equivalent to $\lambda_j > 0$ for $j \in J$. A reasonable equivalent is to insist that x be in the *relative interior* of y and z. Let $F = \{e \in E - \{1\} \mid y_e = z_e\}$. Then we say that x is in the relative interior of the combinatorial hull of y and z if $y_e < x_e < z_e$ for all $e \notin F - \{1\}$.

(d) Prove that when x is in the relative interior of the combinatorial hull of y and z and S is tight for x, then S is tight for both y and z.

Suppose that S is tight for x and that $S \not\subseteq F$. Suppose that $1 \notin S$, so that $\tilde{x}(\tilde{S}) < \tilde{z}(\tilde{S})$. Then $f(S) = x(S) = \tilde{x}(\tilde{S}) < \tilde{z}(\tilde{S}) = z(S) \leq f(S)$, a contradiction. Thus $S \subseteq F$ and so $y(S) = x(S) = z(S) = f(S)$, and S is tight for y and z.

Suppose instead that $1 \in S$ and $E - F \not\subseteq S$, so that $\tilde{y}(\tilde{E} - \tilde{S}) < \tilde{x}(\tilde{E} - \tilde{S})$. Then $f(S) = x(S) = \tilde{x}(\tilde{S}) + x_1 = \tilde{x}(\tilde{S}) + f(E) - \tilde{x}(\tilde{E}) = f(E) - \tilde{x}(\tilde{E} - \tilde{S}) < f(E) - \tilde{y}(\tilde{E} - \tilde{S}) = \tilde{y}(\tilde{S}) + f(E) - \tilde{y}(\tilde{E}) = \tilde{y}(\tilde{S}) + y_1 = y(S) \leq f(S)$, again a contradiction. Thus $E - F \subseteq S$, and so $y(E - S) = \tilde{y}(\tilde{E} - \tilde{S}) = \tilde{x}(\tilde{E} - \tilde{S}) = z(E - S)$. Since, e.g., $z(S) = z(E) - z(E - S)$, this implies that $y(S) = x(S) = z(S) = f(S)$, and so S is tight for y and z.

Question 10. Let's consider parametric submodular minimization. Suppose that E is a finite ground set and that $g(S, \lambda)$ is a function where $S \subseteq E$ and λ is a scalar parameter. One example would be where E is $N - \{s, t\}$ in a parametric max flow network with capacities $u_{ij}(\lambda)$, and $g(S, \lambda)$ is the value of cut $S + \{s\}$ w.r.t. λ.

We suppose that $g(S, \lambda)$ is submodular in S for each fixed value of λ, and that it satisfies the following *Decreasing Differences* property for each $S \subseteq T$ and $\lambda' \geq \lambda$:

$$g(T, \lambda) - g(S, \lambda) \geq g(T, \lambda') - g(S, \lambda'). \tag{3}$$

(a) Prove that the following weaker version of (3) implies (3): For all $e \in E$, $S \subseteq E$, and $\lambda' \geq \lambda$,

$$g(S + e, \lambda) - g(S, \lambda) \geq g(S + e, \lambda') - g(S, \lambda'). \tag{4}$$

Enumerate $T - S$ as $\{e_1, e_2, \ldots, e_k\}$. Then using (4) repeatedly we get $g(T, \lambda) - g(S, \lambda) = (g(S + \{e_1, e_2, \ldots, e_k\}, \lambda) - g(S + \{e_1, e_2, \ldots, e_{k-1}\}, \lambda)) + (g(S + \{e_1, e_2, \ldots, e_{k-1}\}, \lambda) - g(S + \{e_1, e_2, \ldots, e_{k-2}\}, \lambda)) + \cdots + (g(S + \{e_1\}, \lambda) - g(S, \lambda)) \geq (g(S + \{e_1, e_2, \ldots, e_k\}, \lambda') - g(S + \{e_1, e_2, \ldots, e_{k-1}\}, \lambda')) + (g(S + \{e_1, e_2, \ldots, e_{k-1}\}, \lambda') - g(S + \{e_1, e_2, \ldots, e_{k-2}\}, \lambda')) + \cdots + (g(S + \{e_1\}, \lambda') - g(S, \lambda')) = g(T, \lambda') - g(S, \lambda')$.

(b) Suppose that Q minimizes g at λ and Q' minimizes g at λ'. Prove that $Q \cap Q'$ also minimizes g at λ, and $Q \cup Q'$ minimizes g at λ'.

Using respectively the optimality of Q, submodularity of g, (3), and optimality of Q' we get $0 \geq g(Q, \lambda) - g(Q \cap Q', \lambda) \geq g(Q \cup Q', \lambda) - g(Q', \lambda) \geq g(Q \cup Q', \lambda') - g(Q', \lambda') \geq 0$. Thus we get equality everywhere, and so we have that $g(Q, \lambda) = g(Q \cap Q', \lambda)$ (i.e., $Q \cap Q'$ is optimal for λ), and $g(Q, \lambda') = g(Q \cup Q', \lambda')$ (i.e., $Q \cup Q'$ is optimal for λ').

(c) Consider now the following *strict* version of (3): For $S \subset T$ and $\lambda < \lambda'$

$$g(T, \lambda) - g(S, \lambda) > g(T, \lambda') - g(S, \lambda'). \tag{5}$$

Prove that when (5) is true that if Q is a min cut for λ and Q' is a min cut for λ' with $\lambda' > \lambda$, that $Q \subseteq Q'$. (Thus with (5), *every* min cut for λ is nested with *every* min cut for λ'.)

If $Q \not\subseteq Q'$, then $Q' \subset Q' \cup Q$ and so (5) applies to $S = Q'$ and $T = Q \cup Q'$. Then as in (b) we get $0 \geq g(Q, \lambda) - g(Q \cap Q', \lambda) \geq g(Q \cup Q', \lambda) - g(Q', \lambda) > g(Q \cup Q', \lambda') - g(Q', \lambda') \geq 0$, a contradiction. Hence we must have that $Q \subseteq Q'$.

Question 11. Consider the base polytope $B(f)$ of a submodular function f defined on ground set E. For $S \subseteq E$ define $M(S)$ to be the maximum value of $x(S)$ over $x \in B(f)$, and $m(S)$ to be the minimum value of $x(S)$ over $x \in B(f)$.

(a) Give closed-form expressions for $M(S)$ and $m(S)$ (in terms of f and S and E), and show how to compute some $x' \in B(f)$ with $x'(S) = M(S)$ and some $x'' \in B(f)$ with $x''(S) = m(S)$.

Clearly for any $x \in B(f)$ we must have that $x(S) \leq f(S)$. Let $\prec$ be a linear order where all elements of S come before all elements of $E - S$, with corresponding Greedy vertex $x' = v^{\prec}$. Then we can compute that $x'(S) = f(S) - f(\emptyset) = f(S)$, and so $M(S) = f(S)$, and this is attained by any vertex coming from a linear order with S coming before $E - S$.

For any $x \in B(f)$ we have $x(E - S) = x(E) - x(S) = f(E) - x(S)$, and so $x(S) = f(E) - x(E-S) \geq f(E) - f(E-S)$, and so $f(E) - f(E-S)$ is a lower bound on $m(S)$. Let $\prec$ be a linear order where all elements of S come after all elements of $E - S$, with corresponding Greedy vertex $x'' = v^{\prec}$. Then we can compute that $x''(S) = f(E) - f(E-S)$, and so $m(S) = f(E) - f(E-S)$, and this is attained by any vertex coming from a linear order with S coming after $E - S$.

(b) Let's extend part (a) a bit. Let $n = |E|$ and suppose that v is the Greedy vertex corresponding to linear order $\prec$. Assume w.l.o.g. that $\prec = 123 \cdots n$. For $j \in E$ define $B([1,j), \prec) = \{x \in B(f) \mid x_1 = v_1,\ x_2 = v_2, \ldots, x_{j-1} = v_{j-1}\}$ and $B((j,n], \prec) = \{x \in B(f) \mid x_n = v_n,\ x_{n-1} = v_{n-1}, \ldots, x_{j+1} = v_{j+1}\}$; notice that both $B([1,j), \prec)$ and $B((j,n], \prec)$ are non-empty since v belongs to both. Prove that $v_j = \max\{x_j \mid x \in B([1,j), \prec)\}$ and $v_j = \min\{x_j \mid x \in B((j,n], \prec)\}$.

We prove the first part, as the second part is similar. Suppose that $x \in B([1,j,\prec)$. Define $S_j = 123 \cdots j$ and $S_{j-1} = 123 \cdots j-1$. Since $v(S_{j-1}) = f(S_{j-1})$ by Greedy, and $x(S_{j-1}) = v(S_{j-1})$ by $x \in B([1,j), \prec)$, we have $x(S_{j-1}) = f(S_{j-1})$. By Greedy $v_j = f(S_j) - f(S_{j-1}) = f(S_j) - x(S_{j-1})$. Now $f(S_j) \geq x(S_j) = x_j + x(S_{j-1}) = x_j + f(S_{j-1})$, or $x_j \leq f(S_j) - f(S_{j-1}) = v_j$. Thus v_j is indeed the max value of the j-th component among members of $B([1,j), \prec)$.

(c) If $S, T \subseteq E$ with $S \subseteq T$, then the *interval* $[S,T] = \{R \subseteq E \mid S \subseteq R \subseteq T\}$. Let $\alpha \geq 0$ be a scalar, $T \subseteq E$, and define $f_{T,-\alpha}(S)$ to be $f(S) - \alpha$ for $S \in [T,E]$, and $f(S)$ for $S \notin [T,E]$, and define $f_{T,+\alpha}(S)$ to be $f(S)$ for $S \in [\emptyset, T]$, and $f(S) + \alpha$ for $S \notin [\emptyset, T]$. Prove that $f_{T,\pm\alpha}$ are again submodular with $f_{T,\pm\alpha}(\emptyset) = 0$.

First consider $f_{T,+\alpha}$ (the other case is similar). Since $\emptyset \in [\emptyset, T]$, we have $f_{T,+\alpha}(\emptyset) = f(\emptyset) = 0$.

Consider the submodular inequality $f(S) + f(R) \geq f(S \cup R) + f(S \cap R)$. If both S and R belong to $[\emptyset, T]$, then all four terms have f equal to $f_{T,+\alpha}$, and so it is preserved. If neither S nor R belongs to $[\emptyset, T]$, then the LHS gains $2\alpha \geq 0$, and the RHS gains at most 2α (really, at most α since $S \cup R \not\subseteq T$), so it is preserved again. If $S \not\subseteq T$ but $R \subseteq T$, then $S \cap R \subseteq T$ but $S \cup R \not\subseteq T$, and so both sides gain α, so it is again preserved. This is related to the concept of #-duality from Fujishige's book pp. 43–44.

(d) The *membership problem* for $B(f)$ is this: Given some point $x \in \mathbb{R}^E$ with $x(E) = f(E)$, either prove that $x \in B(f)$ or find some $S \subset E$ such that $x(S) > f(S)$.

Show how to reduce the membership problem for a general x and submodular f to the membership problem for 0 and an associated $\hat{f}$ with $\hat{f}(E) = \hat{f}(\emptyset) = 0$, i.e., determining whether $0 \in B(\hat{f})$ is equivalent to determining if $x \in B(f)$.

Define $\hat{f}(S) = f(S) - x(S)$. Clearly we have that $\hat{f}(E) = \hat{f}(\emptyset) = 0$. Since $x(S)$ is modular, $\hat{f}(S)$ is submodular. Now $x \notin B(f)$ iff there is some $S \subset E$ such that $x(S) > f(S)$, which translates into $0 > f(S) - x(S) = \hat{f}(S)$ and so $0 \notin B(\hat{f})$.

Question 12. Two useful variants of SFMin are these: Given T with $\emptyset \subset T \subset E$, solve (1) $\min_{S \subseteq T} f(S)$, and (2) $\min_{S \supseteq T} f(S)$. Define auxiliary set functions like this: For $S \subseteq T$ define $f_{\subseteq T}(S) = f(S)$; for $S \subseteq E - T$ define $f_{\supseteq T}(S) = f(S \cup T) - f(T)$.

(a) Prove that both $f_{\subseteq T}$ and $f_{\supseteq T}$ are submodular and equal to zero on the empty set.

Since $f(T)$ is just a constant, clearly both are submodular. Also, $f_{\subseteq T}(\emptyset) = 0 = f(T) - f(T) = f_{\supseteq T}(\emptyset)$.

(b) Prove that S solves $\min_{S \subseteq T} f(S)$ iff S solves SFMin for $f_{\subseteq T}$, and S solves $\min_{S \supseteq T} f(S)$ iff $S - T$ solves SFMin for $f_{\supseteq T}$.

The first follows from the fact that $f_{\subseteq T} = f$ on subsets of T. For the second, suppose that S solves $\min_{S \supseteq T} f(S)$. Then $R \equiv S - T \subseteq E - T$. If R does not minimize $f_{\supseteq T}$, then there is some $R' \subseteq E - T$ with $f_{\supseteq T}(R') < f_{\supseteq T}(R)$, or $f(R' \cup T) - f(T) < f(R \cup T) - f(T)$, or $f(R' \cup T) < f(S)$. But then $R' \cup T$ contains T and contradicts that S solves $\min_{S \supseteq T} f(S)$. Thus R minimizes $f_{\supseteq T}$. Conversely, suppose that $S - T$ solves SFMin for $f_{\supseteq T}$. If S doesn't solve $\min_{S \supseteq T} f(S)$ then there is some $S' \supseteq T$ with $f(S') < f(S)$. But then $f_{\supseteq T}(S' - T) = f(S') - f(T) < f(S) - f(T) = f_{\supseteq T}(S - T)$, contradicting that $S - T$ solves SFMin for $f_{\supseteq T}$.

(c) Suppose that $x \in \mathbb{R}^E$. For $T \subseteq E$ define $x^{|T} \in \mathbb{R}^T$ by $x^{|T}_e = x_e$, i.e., vector x restricted to the components in T. If $x \in \mathbb{R}^T$ and $y \in \mathbb{R}^{E-T}$, define $x \oplus y$ by $(x \oplus y)_e = x_e$ if $e \in T$, $(x \oplus y)_e = y_e$ if $e \in E - T$. Prove that if $x \in B(f_{\subseteq T})$ and $y \in B(f_{\supseteq T})$, then $x \oplus y \in B(f)$.

Define $z = x \oplus y$. Let $S \subseteq E$. Then $z(S) = z(S \cap T) + z(S - T) = x(S \cap T) + y(S - T) \leq f_{\subseteq T}(S \cap T) + f_{\supseteq T}(S - T) = f(S \cap T) + (f(S \cup T) - f(T)) \leq f(S)$. When $S = E$ this specializes to $z(E) = z(T) + z(E - T) = x(T) + y(E - T) = f(T) + (f(E) - f(T)) = f(E)$, and so $z \in B(f)$.

(d) As a partial converse to (c), suppose that $z \in B(f)$ and T is such that T is z-tight. Prove that $x \equiv z^{|T} \in B(f_{\subseteq T})$ and $y \equiv z^{|(E-T)} \in B(f_{\supseteq T})$.

For $S \subseteq T$ we have $x(S) = z(S) \leq f(S) = f_{\subseteq T}(S)$, and by hypothesis $x(T) = f(T)$, and so $x \in B(f_{\subseteq T})$. For $S \subseteq E - T$ we have $y(S) = z(S) = z(S) + (z(T) - f(T)) = z(S \cup T) - f(T) \leq f(S \cup T) - f(T) = f_{\supseteq T}(S)$. Specializing to $S = E - T$ we get $y(E - T) = z(E - T) = z(E - T) + (z(T) - f(T)) = z(E) - f(T) = f(E) - f(T) = f_{\supseteq T}(E - T)$, and so $y \in B(f_{\supseteq T})$.

(e) Assume that $\prec$ is a linear order which has the elements of T before all other elements. For $e \in T$ and linear order $\prec$ we have two ways to generate $v_e^{\prec \subseteq T}$: we can do Greedy w.r.t. f and $\prec$ and take component e, or we can do Greedy w.r.t. $f_{\subseteq T}$ and take component e. For $e \in E - T$ and linear order $\prec$ we have two ways to generate $v_e^{\prec \supseteq T}$: we can do Greedy w.r.t. f and $\prec$ and take component e, or we can do Greedy w.r.t. $f_{\supseteq T}$ and take component e. In each case prove that we get the same answer either way.

The $\prec_{\subseteq T}$ case: Here $e^{\prec_{\subseteq T}} = e^{\prec}$, so Greedy gives the same answer. The $\prec_{\supseteq T}$ case: Here $e^{\prec_{\supseteq T}} \cup T = e^{\prec}$, and so for component e we get $v_e^{\prec_{\supseteq T}} = f_{\supseteq T}(e^{\prec_{\supseteq T}} + e) - f_{\supseteq T}(e_e^{\prec_{\supseteq T}}) = (f((e^{\prec_{\supseteq T}} + e) \cup T) - f(T)) - (f(e^{\prec_{\supseteq T}} \cup T) - f(T)) = f(e^{\prec} + e) - f(e^{\prec}) = v_e^{\prec}$.

(f) (1) Suppose that $\emptyset \subset T_1 \subset T_2 \subset E$. Prove that $(f_{\subseteq T_2})_{\subseteq T_1} = f_{\subseteq T_1}$. (2) Suppose that $\emptyset \subset T_1 \subset E$ and $\emptyset \subset T_2 \subset E - T_1$. Prove that $(f_{\supseteq T_1})_{\supseteq T_2} = f_{\supseteq (T_1 \cup T_2)}$. (3) Suppose that $\emptyset \subset T_1 \subset T_2 \subset E$. Prove that $(f_{\supseteq T_1})_{\subseteq T_2 - T_1} = (f_{\subseteq T_2})_{\supseteq T_1}$. (Therefore we can apply these operations repeatedly and in any order, and we know that the resulting function will depend only on the largest set we have to contain, and the smallest set we have to be contained in.)

(1) For $S \subseteq T_1$, $(f_{\subseteq T_2})_{\subseteq T_1}(S) = f_{\subseteq T_1}(S) = f(S)$.

(2) For $S \subseteq E - (T_1 \cup T_2)$, $(f_{\supseteq T_1})_{\supseteq T_2}(S) = f_{\supseteq T_1}(S \cup T_2) - f_{\supseteq T_1}(T_2) = (f(S \cup T_2 \cup T_1) - f(T_1)) - (f(T_2 \cup T_1) - f(T_1)) = f(S \cup (T_1 \cup T_2)) - f(T_1 \cup T_2) = f_{\supseteq (T_1 \cup T_2)}(S)$.

(3) For R s.t. $T_1 \subseteq R \subseteq T_2$ define $S = R - T_1$. Then $(f_{\supseteq T_1})_{\subseteq T_2 - T_1}(S) = f_{\supseteq T_1}(S) = f(S \cup T_1) - f(T_1) = f_{\subseteq T_2}(S \cup T_1) - f_{\subseteq T_2}(T_1) = (f_{\subseteq T_2})_{\supseteq T_1}$.

Question 13 Suppose that we have a submodular function f with polyhedron $P(f)$, and a scalar σ. Define the hyperplane H_σ as $\{x \in \mathbb{R}^E \mid x(E) = \sigma\}$.

(a) How can we determine whether $P(f)|_\sigma \equiv P(f) \cap H_\sigma$ is empty or not?

If $\sigma > f(E)$ then clearly $P(f)|_\sigma = \emptyset$, as all $x \in P(f)$ satisfy $x(E) \le f(E) < \sigma$.

On the other side, suppose that $\sigma \le f(E)$. Then we know that there are points $x \in B(f) \subset P(f)$ satisfying $x(E) = f(E)$. Also, any $y \le x$ is also in $P(f)$, and thus e.g. for some e $y \equiv x - (f(E) - \sigma)\chi_e \in P(f)$ and has $y(E) = x(E) - (f(E) - \sigma) = \sigma$, and so $P(f)|_\sigma \ne \emptyset$.

(b) Suppose that $P(f)|_\sigma \ne \emptyset$. Given weight vector $w \in \mathbb{R}^E$, how can we adapt Greedy to solve $\max w^T x$ s.t. $x \in P(f)|_\sigma$?

One way to go at this is to define $f|_\sigma(S) = f(S)$ for $S \ne E$, $f|_\sigma(E) = \sigma$, and prove that $f|_\sigma(S)$ is submodular. This follows since the only significant change to $f(S) + f(T) \ge f(S \cup T) + f(S \cap T)$ could be when the union is E, and then the RHS goes down while the LHS stays the same. Then Greedy adapted to this is the same as ordinary Greedy, except that in the last step it sets x_{e_n} to $\sigma - f(E - e_n)$ instead of $f(E) - f(E - e_n)$. But since the proof that Greedy works doesn't depend on the sign of π_E, the same proof still shows that this adapted Greedy produces the optimal solution.

References

[1] Annotations in this sans-serif font

[2] R. K. Ahuja, T. L. Magnanti and J. B. Orlin (1993). *Network Flows: Theory, Algorithms, and Applications*. Prentice-Hall, Englewood Cliffs.
Good reference on network flow.

[3] R. E. Bixby, W. H. Cunningham, and D. M. Topkis (1985). The Partial Order of a Polymatroid Extreme Point. *Math. of OR*, **10**, 367–378.
How to compute some exchange capacities.

[4] W. H. Cunningham (1984). Testing Membership in Matroid Polyhedra. *JCT Series B*, **36**, 161–188.
Early try at SFMin.

[5] W. H. Cunningham (1985). On Submodular Function Minimization. *Combinatorica*, **3**, 185–192.
Another early try at SFMin.

[6] J. Edmonds (1970). Submodular Functions, Matroids, and Certain Polyhedra. In *Combinatorial Structures and their Applications*, R. Guy, H. Hanani, N. Sauer, and J. Schönheim, eds., Gordon and Breach, 69–87.
Source of SFMin duality LPs.

[7] U. Feige, V. Mirrokni, and J. Vondrák (2007). Maximizing Non-monotone Submodular Functions. *Proceedings of 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 461–471.
Approximation algorithm for SFMax.

[8] L. K. Fleischer (2000). Recent Progress in Submodular Function Minimization. *Optima*, September 2000, 1–11.
A shorter survey of early algorithms for SFMin.

[9] L. K. Fleischer and S. Iwata (2003). A Push-Relabel Framework for Submodular Function Minimization and Applications to Parametric Optimization. "Submodularity" special issue of *Discrete Applied Mathematics*, S. Fujishige ed., **131**, 311-322.
Has Push-Relabel version of Schrijver, parametric SFMin.

[10] S. Fujishige (2005). *Submodular Functions and Optimization*. Second Edition. North-Holland.
Standard reference book for submodular optimization.

[11] M. R. Garey and D. S. Johnson (1979). *Computers and Intractability, A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company, New York.
Basic reference on NP Completeness.

[12] M. Grötschel, L. Lovász, and A. Schrijver (1988). *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag.
Standard text on how the Ellipsoid Algorithm shows the equivalence of Separation and Optimization.

[13] S. Iwata (2002). A Fully Combinatorial Algorithm for Submodular Function Minimization. *J. Combin. Theory Ser. B*, **84**, 203–212; a corrected version is available at `http://www.sr3.t.u-tokyo.ac.jp/~iwata/`.
First fully combinatorial algorithm for SFMin.

[14] S. Iwata (2003). A Faster Scaling Algorithm for Minimizing Submodular Functions. *SIAM J. on Computing*, **32**, 833–840.
One of the fastest known SFMin algorithms.

[15] S. Iwata (2008). Submodular Function Minimization. *Mathematical Programming*, **112**, 45–64.
Another SFMin survey.

[16] S. Iwata, L. Fleischer, and S. Fujishige (2001). A Combinatorial, Strongly Polynomial-Time Algorithm for Minimizing Submodular Functions. *J. ACM*, **48**, 761–777.
The original IFF SFMin algorithm paper.

[17] S. Iwata and J. B. Orlin (2009). A Simple Combinatorial Algorithm for Submodular Function Minimization. Technical report; an extended abstract appears in *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA 2009), pp. 1230–1237.
The most recent and sometimes faster SFMin algorithm.

[18] A. Krause and D. Golovin (2013). Submodular Function Maximization. Chapter in *Tractability: Practical Approaches to Hard Problems* (to appear), Cambridge University Press.
Survey of SFMax.

[19] A. Krause and C. Guestrin (2011). Submodularity and its Applications in Optimized Information Gathering. *ACM Transactions on Intelligent Systems and Technology*, **2**, Article 32.
Survey of how SFMax applies to problems such as sensor location.

[20] L. Lovász (1983). Submodular Functions and Convexity. In *Mathematical Programming — The State of the Art*, A. Bachem, M. Grötschel, B. Korte eds., Springer, Berlin, 235–257.
Gives result showing that submodular functions are convex via the Lovász extension.

[21] S. T. McCormick (2006). Submodular Function Minimization. Chapter 7 in the *Handbook on Discrete Optimization*, Elsevier, K. Aardal, G. Nemhauser, and R. Weismantel, eds., 321–391. See `http://www.elsevier.com/wps/find/bookdescription.cws_home/699541/description`
Published version of Tom's SFMin survey.

[22] K. Murota (2003). *Discrete Convex Analysis.* SIAM Monographs on Discrete Mathematics and Applications, Society for Industrial and Applied Mathematics, Philadelphia.
Standard reference for discrete convexity.

[23] Nemhauser, George L., Wolsey, Laurence A., and Fisher, Marshall L. (1978). An analysis of approximations for maximizing submodular set functions - I. *Mathematical Programming*, **14** 265-294.
Contains the first approximation algorithm for SFMax.

[24] J.B. Orlin (2007). A Faster Strongly Polynomial Algorithm for Submodular Function Minimization. *Proceedings of IPCO 12*, M. Fischetti and D. Williamson, eds., Ithaca, NY, 240–251.
The current fastest SFMin algorithm in most cases.

[25] M. N. Queyranne (1998). Minimizing Symmetric Submodular Functions. *Math. Prog.*, **82**, 3–12.
Queyranne's Algorithm for Symmetric SFMin.

[26] A. Schrijver (2000). A Combinatorial Algorithm Minimizing Submodular Functions in Strongly Polynomial Time. *J. Combin. Theory Ser. B* **80**, 346–355.
Schrijver's Algorithm for SFMin.

[27] A. Schrijver (2003). *Combinatorial Optimization: Polyhedra and Efficiency.* Springer, Berlin.
Standard textbook for basics of combinatorial optimization.