5.3.1 Span

The following definition is also motivated by the linear algebra setting.

Definition 5.1 *Given a matroid $M = (E, \mathcal{I})$ and given $S \subseteq E$, let*

$$\text{span}(S) = \{e \in E : r(S \cup \{e\}) = r(S)\}.$$

The span of a set is also called the closure. Observe that $S \subseteq \text{span}(S)$. We claim that $r(S) = r(\text{span}(S))$; in other words, if adding an element to S does not increase the rank, adding many such elements also does not increase the rank. Indeed, take a maximal independent subset of S, say J. If $r(\text{span}(S)) > |J|$ then there exists $e \in \text{span}(S) \setminus J$ such that $J + e \in \mathcal{I}$. Thus $r(S + e) \geq r(J + e) = |J| + 1 > |J| = r(S)$ contradicting the fact that $e \in \text{span}(S)$.

Definition 5.2 *A set S is said to be closed if $S = \text{span}(S)$.*

Exercise 5-9. Given a matroid M with rank function r and given an integer $k \in \mathbb{N}$, what is the rank function of the truncated matroid M_k (see Exercise 5-6 for a definition).

Exercise 5-10. What is the rank function of a laminar matroid, see exercise 5-7?

5.4 Matroid Polytope

Let

$$X = \{\chi(S) \in \{0, 1\}^{|E|} : S \in \mathcal{I}\}$$

denote the incidence (or characteristic) vectors of all independent sets of a matroid $M = (E, \mathcal{I})$, and let the *matroid polytope* be defined as $\text{conv}(X)$. In this section, we provide a complete characterization of $\text{conv}(X)$ in terms of linear inequalities. In addition, we illustrate the different techniques proposed in the polyhedral chapter for proving a complete description of a polytope.

Theorem 5.4 *Let*

$$P = \{x \in \mathbb{R}^{|E|} : \begin{array}{ll} x(S) \leq r(S) & \forall S \subseteq E \\ x_e \geq 0 & \forall e \in E\} \end{array}$$

where $x(S) := \sum_{e \in S} x_e$. Then $\text{conv}(X) = P$.

It is clear that $\text{conv}(X) \subseteq P$ since $X \subseteq P$. The harder part is to show that $P \subseteq \text{conv}(X)$. In the next three subsections, we provide three different proofs based on the three techniques to prove complete polyhedral descriptions.

5.4.1 Algorithmic Proof

Here we provide an algorithmic proof based on the greedy algorithm. From $\text{conv}(X) \subseteq P$, we know that

$$\max\{c^T x : x \in X\} = \max\{c^T x : x \in \text{conv}(X)\} \le \max\{c^T x : \begin{array}{ll} x(S) \le r(S) & S \subseteq E \\ x_e \ge 0 & e \in E\}. \end{array}$$

Using LP duality, we get that this last expression equals:

$$\min\{\textstyle\sum_S r(S) y_S : \begin{array}{ll} \sum_{S:e\in S} y_S \ge c(e) & \forall e \in E \\ y_S \ge 0 & S \subseteq E\}. \end{array}$$

Our goal now is, for any cost function c, to get an independent set S and a dual feasible solution y such that $c^T\chi(S) = \sum_S r(S)y_S$ which proves that $\text{conv}(X) = P$.

Consider any cost function c. We know that the maximum cost independent set can be obtained by the greedy algorithm. More precisely, it is the last set S_k returned by the greedy algorithm when we consider only those elements up to e_q where $c(e_q) \ge 0 \ge c(e_{q+1})$. We need now to exhibit a dual solution of the same value as S_k. There are exponentially many variables in the dual, but this is not a problem. In fact, we will set most of them to 0.

For any index $j \le k$, we have $S_j = \{s_1, s_2, \cdots, s_j\}$, and we define U_j to be all elements in our ordering up to and excluding s_{j+1}, i.e. $U_j = \{e_1, e_2, \cdots, e_l\}$ where $e_{l+1} = s_{j+1}$. In other words, U_j is all the elements in the ordering just before s_{j+1}. One important property of U_j is that

$$r(U_j) = r(S_j) = j.$$

Indeed, by independence $r(S_j) = |S_j| = j$, and by (R_2), $r(U_j) \ge r(S_j)$. If $r(U_j) > r(S_j)$, there would be an element say $e_p \in U_j \setminus S_j$ such that $S_j \cup \{e_p\} \in \mathcal{I}$. But the greedy algorithm would have selected that element (by (I_1)) contradicting the fact that $e_p \in U_j \setminus S_j$.

Set the non-zero entries of y_S in the following way. For $j = 1, \cdots, k$, let

$$y_{U_j} = c(s_j) - c(s_{j+1}),$$

where it is understood that $c(s_{k+1}) = 0$. By the ordering of the $c(\cdot)$, we have that $y_S \ge 0$ for all S. In addition, for any $e \in E$, we have that

$$\sum_{S:e\in S} y_S = \sum_{j=t}^{k} y_{U_j} = c(s_t) \ge c(e),$$

where t is the least index such that $e \in U_t$ (implying that e does not come before s_t in the ordering). This shows that y is a feasible solution to the dual. Moreover, its dual value is:

$$\sum_S r(S)y_S = \sum_{j=1}^{k} r(U_j)y_{U_j} = \sum_{j=1}^{k} j(c(s_j) - c(s_{j+1})) = \sum_{j=1}^{k} (j-(j-1))c(s_j) = \sum_{j=1}^{k} c(s_j) = c(S_k).$$

This shows that the dual solution has the same value as the independent set output by the greedy algorithm, and this is true for all cost functions. This completes the algorithmic proof.

5.4.2 Vertex Proof

Here we will focus on any vertex x of

$$P = \{x \in \mathbb{R}^{|E|} : \begin{array}{ll} x(S) \le r(S) & \forall S \subseteq E \\ x_e \ge 0 & \forall e \in E\} \end{array}$$

and show that x is an integral vector. Since $x(\{e\}) \le r(\{e\}) \le 1$, we get that $x \in \{0,1\}^{|E|}$ and thus it is the incidence vector of an independent set.

Given any $x \in P$, consider the *tight* sets S, i.e. those sets for which $x(S) = r(S)$. The next lemma shows that these tight sets are closed under taking intersections or unions. This lemma is really central, and follows from submodularity.

Lemma 5.5 *Let $x \in P$. Let*

$$\mathcal{F} = \{S \subseteq E : x(S) = r(S)\}.$$

Then

$$S \in \mathcal{F}, T \in \mathcal{F} \Rightarrow S \cap T \in \mathcal{F}, S \cup T \in \mathcal{F}.$$

Observe that the lemma applies even if S and T are disjoint. In that case, it says that $\emptyset \in \mathcal{F}$ (which is always the case as $x(\emptyset) = 0 = r(\emptyset)$) and $S \cup T \in \mathcal{F}$.

Proof: The fact that $S, T \in \mathcal{F}$ means that:

$$r(S) + r(T) = x(S) + x(T). \tag{2}$$

Since $x(S) = \sum_{e \in S} x_e$, we have that

$$x(S) + x(T) = x(S \cap T) + x(S \cup T), \tag{3}$$

i.e. that the function $x(\cdot)$ is modular (both x and $-x$ are submodular). Since $x \in P$, we know that $x(S \cap T) \le r(S \cap T)$ (this is true even if $S \cap T = \emptyset$) and similarly $x(S \cup T) \le r(S \cup T)$; this implies that

$$x(S \cap T) + x(S \cup T) \le r(S \cap) + r(S \cup T). \tag{4}$$

By submodularity, we have that

$$r(S \cap T) + r(S \cup T) \le r(S) + r(T). \tag{5}$$

Combining (2)–(5), we get

$$r(S) + r(T) = x(S) + x(T) = x(S \cap T) + x(S \cup T) \le r(S \cap T) + r(S \cup T) \le r(S) + r(T),$$

and therefore we have equality throughout. This implies that $x(S \cap T) = r(S \cap T)$ and $x(S \cup T) = r(S \cup T)$, i.e. $S \cap T$ and $S \cup T$ in $\mathcal{F}$. $\triangle$

To prove that any vertex or extreme point of P is integral, we first characterize any face of P. A *chain* $\mathcal{C}$ is a family of sets such that for all $S, T \in \mathcal{C}$ we have that either $S \subseteq T$ or $T \subseteq S$ (or both if $S = T$).

Theorem 5.6 *Consider any face F of P. Then there exists a chain $\mathcal{C}$ and a subset $J \subseteq E$ such that:*

$$F = \{x \in \mathbb{R}^{|E|} : \begin{array}{ll} x(S) \leq r(S) & \forall S \subseteq E \\ x(C) = r(C) & \forall C \in \mathcal{C} \\ x_e \geq 0 & \forall e \in E \setminus J \\ x_e = 0 & \forall e \in J.\} \end{array}$$

Proof: By Theorem 3.5 of the polyhedral notes, we know that any face is characterized by setting some of the inequalities of P by equalities. In particular, F can be expressed as

$$F = \{x \in \mathbb{R}^{|E|} : \begin{array}{ll} x(S) \leq r(S) & \forall S \subseteq E \\ x(C) = r(C) & \forall C \in \mathcal{F} \\ x_e \geq 0 & \forall e \in E \setminus J \\ x_e = 0 & \forall e \in J.\} \end{array}$$

where $J = \{e : x_e = 0 \text{ for all } x \in F\}$ and $\mathcal{F} = \{S : x(S) = r(S) \text{ for all } x \in F\}$. To prove the theorem, we need to argue that the system of equations:

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

can be replaced by an equivalent (sub)system in which $\mathcal{F}$ is replaced by a chain $\mathcal{C}$. To be equivalent, we need that

$$\text{span}(\mathcal{F}) = \text{span}(\mathcal{C})$$

where by span($\mathcal{L}$) we mean

$$\text{span}(\mathcal{L}) := \text{span}\{\chi(C) : C \in \mathcal{L}\}.$$

Let $\mathcal{C}$ be a maximal subchain of $\mathcal{F}$, i.e. $\mathcal{C} \subseteq \mathcal{F}$, $\mathcal{C}$ is a chain and for all $S \in \mathcal{F} \setminus \mathcal{C}$, there exists $C \in \mathcal{C}$ such that $S \not\subseteq C$ and $C \not\subseteq S$. We claim that $\text{span}(\mathcal{C}) = \text{span}(\mathcal{F})$.

Suppose not, i.e. $H \neq \text{span}(\mathcal{F})$ where $H := \text{span}(\mathcal{C})$. This means that there exists $S \in \mathcal{F} \setminus \mathcal{C}$ such that $\chi(S) \notin H$ but S cannot be added to $\mathcal{C}$ without destroying the chain structure. In other words, for any such S, the set of 'chain violations'

$$V(S) := \{C \in \mathcal{C} : C \not\subseteq S \text{ and } S \not\subseteq C\}$$

is non-empty. Among all such sets S, choose one for which $|V(S)|$ is as small as possible ($|V(S)|$ cannot be 0 since we are assuming that $V(S) \neq \emptyset$ for all possible S). Now fix some set $C \in V(S)$. By Lemma 5.5, we know that both $C \cap S \in \mathcal{F}$ and $C \cup S \in \mathcal{F}$. Observe that there is a linear dependence between $\chi(C), \chi(S), \chi(C \cup T), \chi(C \cap T)$:

$$\chi(C) + \chi(S) = \chi(C \cup S) + \chi(C \cap S).$$

This means that, since $\chi(C) \in H$ and $\chi(S) \notin H$, we must have that either $\chi(C \cup S) \notin H$ or $\chi(C \cap S) \notin H$ (otherwise $\chi(S)$ would be in H). Say that $\chi(B) \notin H$ where B is either $C \cup S$ or $C \cap S$. This is a contradiction since $|V(B)| < |V(S)|$, contradicting our choice of S. Indeed, one can see that $V(B) \subset V(S)$ and $C \in V(S) \setminus V(B)$. △

As a corollary, we can also obtain a similar property for an extreme point, starting from Theorem 3.6.

Corollary 5.7 *Let x be any extreme point of P. Then there exists a chain $\mathcal{C}$ and a subset $J \subseteq E$ such that x is the unique solution to:*

$$\begin{array}{ll} x(C) = r(C) & \forall C \in \mathcal{C} \\ x_e = 0 & \forall e \in J. \end{array}$$

From this corollary, the integrality of every extreme point follows easily. Indeed, if the chain given in the corollary consists of $C_1 \subset C_2 \subset C_p$ the the system reduces to

$$\begin{array}{ll} x(C_i \setminus C_{i-1}) = r(C_i) - r(C_{i-1}) & i = 1, \cdots, p \\ x_e = 0 & \forall e \in J, \end{array}$$

where $C_0 = \emptyset$. For this to have a unique solution, we'd better have $|C_i \setminus C_{i-1} \setminus J| \leq 1$ for all i and the values for the resulting x_e's will be integral. Since $0 \leq x_e \leq r(\{e\}) \leq 1$, we have that x is a $0-1$ vector and thus $x = \chi(S)$ for some set S. As $|S| \leq r(S) \leq |S|$, we have $|S| = r(S)$ and thus $S \in \mathcal{I}$ and therefore x is the incidence vector of an independent set. This completes the proof.

5.4.3 Facet Proof

Our last proof of Theorem 5.4 focuses on the facets of $\text{conv}(X)$.

First we need to argue that we are missing any equalities. Let's focus on the (interesting) case in which any singleton set is independent: $\{e\} \in \mathcal{I}$ for every $e \in E$. In that case $dim(\text{conv}(X)) = |E|$ since we can exhibit $|E| + 1$ affinely independent points in X: the 0 vector and all unit vectors $\chi(\{e\})$ for $e \in E$. Thus we do not need any equalities. See exercise 5-11 if we are not assuming that every singleton set is independent.

Now consider any facet F of $\text{conv}(X)$. This facet is induced by a valid inequality $\alpha^T x \leq \beta$ where $\beta = \max\{\sum_{e \in I} \alpha_e : I \in \mathcal{I}\}$. Let

$$\mathcal{O} = \{I \in \mathcal{I} : \sum_{e \in I} \alpha_e = \beta\},$$

i.e. $\mathcal{O}$ is the set of all independent sets whose incidence vectors belong to the face. We'll show that there exists an inequality in our description of P which is satisfied at equality by the incidence vectors of all sets $I \in \mathcal{O}$.

We consider two cases. If there exists $e \in M$ such that $\alpha_e < 0$ then $I \in \mathcal{O}$ implies that $e \notin I$, implying that our face F is included in the face induced by $x_e \geq 0$ (which is in our description of P).

For the other case, we assume that for all $e \in E$, we have $\alpha_e \geq 0$. We can further assume that $\alpha_{\max} := \max_{e \in E} \alpha_e > 0$ since otherwise F is trivial. Now, define S as

$$S = \{e \in E : \alpha_e = \alpha_{\max}\}.$$

Claim 5.8 *For any $I \in \mathcal{O}$, we have $|I \cap S| = r(S)$.*

This means that the face F is contained in the face induced by the inequality $x(S) \le r(S)$ and therefore we have in our description of P one inequality inducing each facet of $\text{conv}(X)$. Thus we have a complete description of $\text{conv}(X)$.

To prove the claim, suppose that $|I \cap S| < r(S)$. Thus $I \cap S$ can be extended to an independent set $X \in \mathcal{I}$ where $X \subseteq S$ and $|X| > |I \cap S|$. Let $e \in X \setminus (I \cap S)$; observe that $e \in S$ by our choice of X. Since $\alpha_e > 0$ we have that $I + e \notin \mathcal{I}$, thus there is a circuit $C \subseteq I + e$. By the unique circuit property (see Theorem 5.1), for any $f \in C$ we have $I + e - f \in \mathcal{I}$. But $C \setminus S \neq \emptyset$ since $(I \cap S) + e \in \mathcal{I}$, and thus we can choose $f \in C \setminus S$. The cost of $I + e - f$ satisfies:

$$c(I + e - f) = c(I) + c(e) - c(f) > c(I),$$

contradicting the definition of $\mathcal{O}$.

5.5 Facets?

Now that we have a description of the matroid polytope in terms of linear inequalities, one may wonder which of these (exponentially many) inequalities define facets of $\text{conv}(X)$.

For simplicity, let's assume that $r(\{e\}) = 1$ for all $e \in E$ (e belongs to some independent set). Then, every nonnegativity constraint defines a facet of $P = \text{conv}(X)$. Indeed, the 0 vector and all unit vectors except $\chi(\{e\})$ constitute $|E|$ affinely independent points satisfying $x_e = 0$. This mean that the corresponding face has dimension at least $|E| - 1$ and since the dimension of P itself is $|E|$, the face is a facet.

We now consider the constraint $x(S) \le r(S)$ for some set $S \subseteq E$. If S is not closed (see Definition 5.2) then $x(S) \le r(S)$ definitely does not define a facet of $P = \text{conv}(X)$ since it is implied by the constraints $x(\text{span}(S)) \le r(S)$ and $x_e \ge 0$ for $e \in \text{span}(S) \setminus S$.

Another situation in which $x(S) \le r(S)$ does not define a facet is if S can be expressed as the disjoint union of $U \neq \emptyset$ and $S \setminus U \neq \emptyset$ and $r(U) + r(S \setminus U) = r(S)$. In this case, the inequality for S is implied by those for U and for $S \setminus U$.

Definition 5.3 *S is said to be inseparable if there is no U with $\emptyset \neq U \subset S$ such that $r(S) = r(U) + r(S \setminus U)$.*

From what we have just argued, a necessary condition for $x(S) \le r(S)$ to define a facet of $P = \text{conv}(X)$ is that S is *closed and inseparable*. This can be shown to be sufficient as well, although the proof is omitted.

As an example, consider a partition matroid with $M = (E, \mathcal{I})$ where

$$\mathcal{I} = \{X \subseteq E : |X \cap E_i| \le k_i \text{ for all } i = 1, \cdots, l\},$$

for disjoint E_i's. Assume that $k_i \ge 1$ for all i. The rank function for this matroid is:

$$r(S) = \sum_{i=1}^{l} \min(k_i, |S \cap E_i|).$$