FACET-DEFINING VALID INEQUALITIES OF THE CLUSTERING POLYTOPES

YOSHITSUGU YAMAMOTO

Abstract. We give a short proof of the theorem in [1] that the connectivity inequality and the nonnegativity are facet-defining valid inequalities of the *Clustering Polytope*. For the linear ordering polytope see [2] and [3].

1. Notation

- n : number of items (≥ 3)
- $D := n(n-1)$
- $d := \binom{n}{2} = D/2$
- $N := \{1, 2, \ldots, n\}$
- $N^2_{\neq} := \{(i,j) \mid i,j \in N : i \neq j\}$, ordered distinct pairs
- $N^3_{\neq} := \{(i,j,k) \mid i,j,k \in N : i \neq j, j \neq k, k \neq i\}$, ordered distinct triples

2. Definitions

Definition 2.1 (Incidence Vector)**.** A vector $\boldsymbol{x} \in \mathbb{R}^D$ is said to be the *incidence vector* of the clustering $\mathcal{C} := \{C_1, \ldots, C_l, \ldots, C_s\}$ when

$$x_{ij} = \begin{cases} 1 & \exists C_l \in \mathcal{C} \text{ such that } i,j \in C_l \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Note that $x_{ij} = x_{ji}$ for all $(i,j) \in N^2_{\neq}$.

Definition 2.2 (Clustering Polytope)**.** We refer to the convex hull of the incidence vectors of all the clusterings of N as the *clustering polytope* and denote it by P^n, i.e.,

$$P^n := \mathrm{co}\left\{ \boldsymbol{x} \in \mathbb{R}^D \mid \boldsymbol{x} \text{ is the incidence vector of a clustering of } N \right\}, \tag{2.2}$$

where co means the convex hull.

Definition 2.3 (Valid Inequality)**.** An inequality $\boldsymbol{a}^\top \boldsymbol{x} \leq a_0$ is said to be a *valid inequality of* P^n when $\boldsymbol{a}^\top \boldsymbol{x} \leq a_0$ holds for all $\boldsymbol{x} \in P^n$ or equivalently it holds for the incidence vectors of all clusterings of N.

Definition 2.4 (Facet-defining Valid Inequality)**.** A valid inequality $\boldsymbol{a}^\top \boldsymbol{x} \leq a_0$ of P^n is said to be a *facet-defining valid inequality of* P^n or simply to be *facet-defining* when there are as many as $\dim P^n$ affinely independent incidence vectors satisfying $\boldsymbol{a}^\top \boldsymbol{x} = a_0$.

Date: 08+09+10+15 Aug 2009.

3. Linear System of Incidence Vectors

Lemma 3.1.

(1) *Every incidence vector* $\boldsymbol{x}$ *of a clustering satisfies the following linear system:*

$$x_{ij} - x_{ji} = 0 \qquad \forall (i,j) \in N^2_{\neq} \qquad \textit{(symmetry)} \tag{3.1}$$

$$0 \leq x_{ij} \leq 1 \qquad \forall (i,j) \in N^2_{\neq} \tag{3.2}$$

$$x_{ij} + x_{jk} - x_{ki} \leq 1 \quad \forall (i,j,k) \in N^3_{\neq} \qquad \textit{(connectivity)} \tag{3.3}$$

(2) *Every binary vector satisfying* (3.1), (3.2) *and* (3.3) *is the incidence vector of a clustering.*

Proof. It is clear from definition that an incidence vector satisfies (3.1) and (3.2). If at least one of x_{ij} and x_{jk} is 0, (3.3) is met whether x_{ki} may take 0 or 1. Suppose $x_{ij} = x_{jk} = 1$. Then $i, j \in C_l$ for some l and also $j, k \in C_{l'}$ for some l'. Since clustering is a partition of N, we have $l = l'$ and $i, j, k \in C_l$. This implies $x_{ki} = 1$. Therefore $x_{ij} + x_{jk} - x_{ki} \leq 1$.

Concerning the second statement, define the binary relation $i \sim j$ when either $i = j$ or $i \neq j$ and $x_{ij} = 1$. Suppose that i, j and k are distinct and $i \sim j$ and $j \sim k$, i.e., $x_{ij} = x_{jk} = 1$. Then by (3.3) $x_{ki} \geq x_{ij} + x_{jk} - 1 = 1$, meaning $k \sim i$. Therefore the relation $\sim$ is an equivalence relation and yields a partition of N. □

We call (3.1) the *symmetry equality* and (3.3) the *connectivity inequality.*

Lemma 3.2.

$$P^n \subseteq \left\{ \boldsymbol{x} \in \mathbb{R}^D \;\middle|\; (3.1), (3.2), (3.3) \right\}, \tag{3.4}$$

$$\dim P^n = d. \tag{3.5}$$

Proof. The first statement is clear from the second statement of Lemma 3.1. Note that exactly half the equations (3.1) are linearly independent, implying $\dim P \leq d$. For $i, j \in N$ with $i < j$ let $\boldsymbol{x}(\{i,j\})$ be the incidence vector of clustering consisting of cluster $\{i,j\}$ and clusters of a single item $\{k\}$ for $k \neq i, j$. They amount to d and are linearly independent. Therefore P^n contains $d+1$ affinely independent points: $\mathbf{0}$, $x(\{i,j\})$ for $i, j \in N$ with $i < j$. This proves $\dim P \geq d$. □

4. Facet-Defining Valid Inequality

Lemma 4.1. *The connectivity inequality* (3.3) *is a facet-defining valid inequality of* P^3.

Proof. Note that $D = 6$ and $d = 3$. Consider a connectivity inequality, e.g.,

$$x_{12} + x_{23} - x_{31} \leq 1. \tag{4.1}$$

The incidence vectors of the three clusterings $\mathcal{C}_1 := \{\{1,2\},\{3\}\}$, $\mathcal{C}_2 := \{\{1\},\{2,3\}\}$ and $\mathcal{C}_3 := \{\{1,2,3\}\}$ are given by

$$\boldsymbol{x}_1 = (1,0,0,1,0,0)^\top$$
$$\boldsymbol{x}_2 = (0,1,0,0,1,0)^\top$$
$$\boldsymbol{x}_3 = (1,1,1,1,1,1)^\top$$

when the components are indexed as $(1,2),(2,3),(3,1),(2,1),(3,2),(1,3)$. They are linearly independent, hence affinely independent, and satisfy (4.1) by equality. This proves that the connectivity inequality is facet-defining when $n = 3$. □

Fig.1 shows $P^3 \subseteq \mathbb{R}^6$, where the axes correspond to x_{12}, x_{23} and x_{31}, respectively, and x_{21}, x_{32} and x_{13} at the same time.

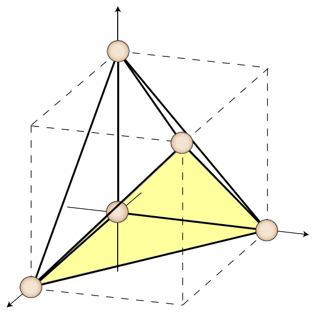

FIGURE 1. Clustering polytope $P^3 \subseteq \mathbb{R}^6$

Theorem 4.2. *The connectivity inequality* (3.3) *is a facet-defining valid inequality of* P^n *for any* $n \geq 3$.

Proof. We prove the assertion by induction over the number of items. Take a connectivity inequality $x_{ij} + x_{jk} - x_{ki} \leq 1$ for the clusterings of $n+1$ items. By renumbering the items if necessary, we assume $(i, j, k) = (1, 2, 3)$ without loss of generality, i.e., the connectivity inequality under consideration is

$$x_{12} + x_{23} - x_{31} \leq 1. \tag{4.2}$$

As the induction hypothesis, we suppose that it is a facet-defining valid inequality of P^n. Then there are $d = \binom{n}{2}$ clusterings $\mathcal{C}_1, \ldots, \mathcal{C}_p, \ldots, \mathcal{C}_d$ of N such that their incidence vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p, \ldots, \boldsymbol{x}_d \in \mathbb{R}^D$ are affinely independent and satisfy (4.2) by equality. Noticing that the right hand side of this equality is nonzero, we easily see that these incidence vectors are linearly independent.

For each of these clusterings let $\mathcal{C}'_p$ be defined as

$$\mathcal{C}'_p := \mathcal{C}_p \cup \{\{n+1\}\},$$

Note that item $n+1$ alone forms a cluster. Then the incidence vector $\boldsymbol{x}'_p \in \mathbb{R}^{(n+1)n}$ of $\mathcal{C}'_p$ is

$$\boldsymbol{x}'_p = \begin{pmatrix} \boldsymbol{x}_p \\ \mathbf{0}^n \\ \mathbf{0}^n \end{pmatrix},$$

where $\mathbf{0}^n$ is a zero vector of $\mathbb{R}^n$, and the last $2n$ components are indexed as $(1, n+1), (2, n+1), \ldots, (n, n+1), (n+1, 1), (n+1, 2), \ldots, (n+1, n)$. Clearly they satisfy (4.2) by equality.

For each $l \in N$ Let $\mathcal{E}_l$ be the clustering of $N \cup \{n+1\}$ defined by

$$\mathcal{E}_l := \begin{cases} \{\{1, 2, n+1\}\} \cup \{\{h\} \mid h \in N \setminus \{1, 2\}\} & l = 1 \\ \{\{2, 3, n+1\}\} \cup \{\{h\} \mid h \in N \setminus \{2, 3\}\} & l = 2 \\ \{\{1, 2\}, \{l, n+1\}\} \cup \{\{h\} \mid h \in N \setminus \{1, 2, l\}\} & l \neq 1 \text{ or } 2 \end{cases}$$

The incidence vector of $\mathcal{E}_l$ is

$$\boldsymbol{y}_l = \begin{pmatrix} \boldsymbol{e}^D_{(1,2)} + \boldsymbol{e}^D_{(2,1)} \\ \boldsymbol{e}^n_{(1,n+1)} + \boldsymbol{e}^n_{(2,n+1)} \\ \boldsymbol{e}^n_{(n+1,1)} + \boldsymbol{e}^n_{(n+1,2)} \end{pmatrix}$$

when $l = 1$,

$$\boldsymbol{y}_l = \begin{pmatrix} \boldsymbol{e}^D_{(2,3)} + \boldsymbol{e}^D_{(3,2)} \\ \boldsymbol{e}^n_{(2,n+1)} + \boldsymbol{e}^n_{(3,n+1)} \\ \boldsymbol{e}^n_{(n+1,2)} + \boldsymbol{e}^n_{(n+1,3)} \end{pmatrix}$$

when $l = 2$, and

$$\boldsymbol{y}_l = \begin{pmatrix} \boldsymbol{e}^D_{(1,2)} + \boldsymbol{e}^D_{(2,1)} \\ \boldsymbol{e}^n_{(l,n+1)} \\ \boldsymbol{e}^n_{(n+1,l)} \end{pmatrix}$$

when $l \neq 1, 2$, where $\boldsymbol{e}^D_{(i,j)}$ is a unit vector of $\mathbb{R}^D$ with 1 at the (i, j) component and $\boldsymbol{e}^n_{(i,n+1)}$ is a unit vector of $\mathbb{R}^n$ with 1 at the $(i, n+1)$ component. It is readily seen that all of these incidence vectors $\boldsymbol{y}_l$ satisfy (4.2).

The matrix of columns $\boldsymbol{x}'_p$'s and $\boldsymbol{y}_l$'s is shown below.

$$\left[\begin{array}{ccc|ccc} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_d & & & \\ & & & \boldsymbol{y}_1 & \cdots & \boldsymbol{y}_n \\ \hline \boldsymbol{0}^n & \cdots & \boldsymbol{0}^n & & & \\ \cline{1-3} & & & & A & \\ \boldsymbol{0}^n & \cdots & \boldsymbol{0}^n & & & \end{array}\right]$$

The right bottom matrix A is of the form

$$A = \left[\begin{array}{ccc|ccc} 1 & & & & & \\ 1 & 1 & & & & \\ & 1 & 1 & & & \\ \hline & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \\ \hline 1 & & & & & \\ 1 & 1 & & & & \\ & 1 & 1 & & & \\ \hline & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{array}\right].$$

Clearly the columns of A are linearly independent. Since $\boldsymbol{x}'_p$'s have zeros in their last $2n$ components, we will see that the columns of the whole matrix, i.e., $\boldsymbol{x}'_p$ $(p = 1, \ldots, d)$ and $\boldsymbol{y}_l$ $(l = 1, \ldots n)$ are linearly independent. We thus have $d + n = \binom{n}{2} + n = \binom{n+1}{2}$ linearly independent nonzero incident vectors satisfying (4.2) by equality. This shows that (4.2) is a facet-defining valid inequality of P^{n+1}.

Combining this with Lemma 4.1, we have proved the theorem. □

Theorem 4.3. *The nonnegativity inequality $x_{ij} \geq 0$ is a facet-defining valid inequality of P^n for any $n \geq 3$.*

Proof. We can assume $(i, j) = (1, 2)$ without loss of generality. For $k, l \in N$ with $k < l$ let $\mathcal{C}(\{k, l\})$ be the clustering $\{\{k, l\}\} \cup \{\{h\} \mid h \in N;\ h \neq k, l\}$. Then its incidence vector is $\boldsymbol{e}^D_{(k,l)} + \boldsymbol{e}^D_{(l,k)}$. Collect those incidence vectors for all possible choices of $k, l \in N$ such that $k < l$ and $\{k, l\} \neq \{1, 2\}$. They amount to $d - 1$, satisfy $x_{12} = 0$. Together with the zero vector, we have d affinely independent incidence vectors satisfying the nonnegativity inequality $x_{12} \geq 0$ by equality. This proves that $x_{12} \geq 0$ is a facet-defining valid inequality. □

5. Other Valid Inequalities of P^n

Let $\{i_1, i_2, \ldots, i_k\}$ be a set of k items and suppose $x_{i_1 i_2} = \cdots = x_{i_{k-1} i_k} = 1$. Then for $\boldsymbol{x}$ to be an incidence vector of a clustering $x_{i_k i_1} = 1$. This requirement is formulated as

$$x_{i_k i_1} \geq x_{i_1 i_2} + \cdots + x_{i_{k-1} i_k} - (k - 2).$$

Clearly this is a valid inequality of P^n.

References

[1] M. Grötschel and Y. Wakabayashi, “Facets of the clique partitioning polytope, ”*Mathematical Programming* **47** (1990) 367–387.

[2] M. Grötschel, M. Jünger and G. Reinelt, “On the acyclic subgraph polytope,” *Mathematical Programming* **33** (1985) 28–42.

[3] M. Grötschel, M. Jünger and G. Reinelt, “Facets of the linear ordering polytope,” *Mathematical Programming* **33** (1985) 43–60.

Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan

E-mail address: yamamoto@sk.tsukuba.ac.jp