

Ruriko Yoshida

# Fundamental holes and saturation points of a commutative semigroup and their applications to contingency tables

Ruriko Yoshida  
Dept. of Mathematics Duke University

Joint work with A. Takemura

[www.math.duke.edu/~ruriko](http://www.math.duke.edu/~ruriko)

# Puzzle

Is there a table satisfying these given marginals?

		1	1	1	1	1	1
	1	1	1	1	1	1	1
0	1	0	1	1	0	1	1
1	0	1	0	1	0	1	1
1	0	0	1	0	1	0	1
0	1	1	0	0	1	1	1
1	1	0	0	1	1	1	0
0	0	1	1	1	1	1	0

Each cell has nonnegative integral value.

# Answer

		1	1	1	1	
	1	1	1	1	1	
0	1	0	1	1	0	1
1	0	1	0	1	0	1
1	0	0	1	0	1	1
0	1	1	0	0	1	1
1	1	0	0	1	1	0
0	0	1	1	1	1	0

There does not exist such a table, although the marginals are consistent.

## Problem

Suppose we have a given set of marginals for contingency tables.

**Want:** decide whether there exists a table satisfying the given marginals.

This is called the **multi-dimensional integer planar transportation problem**.

In terms of Optimization, we can rewrite this problem as an **integral feasibility problem**, that is:

Decide whether there exists an integral solution in the system

$$Ax = b, x \geq 0,$$

where  $A \in \mathbb{Z}^{d \times n}$  and  $b \in \mathbb{Z}^d$ .

## Observation

Assume the lattice  $L$  generated by the columns of  $A$  is  $\mathbb{Z}^d$ . Let  $\text{cone}(A)$  be the cone generated by the columns of  $A$  and  $P_b = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ .

$$P_b \neq \emptyset \Leftrightarrow b \in \text{cone}(A).$$

Let  $Q$  be the semigroup generated by the columns  $\mathbf{a}_i$  of  $A$ , i.e.  $Q = \{x \in \mathbb{R}^d : \sum_{i=1}^n \alpha_i \mathbf{a}_i, \alpha_i \in \mathbb{Z}_+\} \subset \text{cone}(A) \cap \mathbb{Z}^d$ .

$$P_b \cap \mathbb{Z}^n \neq \emptyset \Leftrightarrow b \in Q.$$

$$(P_b \neq \emptyset) \wedge (P_b \cap \mathbb{Z}^n = \emptyset) \Leftrightarrow b \in (\text{cone}(A) \cap \mathbb{Z}^d - Q).$$

We study on the set of **holes** of  $Q$ ,  $H := \text{cone}(A) \cap \mathbb{Z}^d - Q$ .

### **Motivation:**

- (Algebra): Almost all focus in the algebraic literature on this topic is on the normal case (i.e. there are no holes).
- (Statistics): This is significant for statistics because many affine semigroups with statistical connections are not normal.

**Note:**  $Q$  is normal iff the Hilbert basis of  $\text{cone}(A)$  is in  $Q$ .

**Problem:** Find **the necessary and sufficient conditions for  $H$ 's finiteness.**

## Notation and definitions

**Def.** The semigroup  $Q_{\text{sat}} = \text{cone}(A) \cap L$  is called the **saturation** of  $Q$ .

$$Q = AZ_+^n = \{\lambda_1 \mathbf{a}_1 + \cdots + \lambda_n \mathbf{a}_n : \lambda_1, \dots, \lambda_n \in \mathbb{Z}_+\}$$

$$K = A\mathbb{R}_+^n = \{\lambda_1 \mathbf{a}_1 + \cdots + \lambda_n \mathbf{a}_n : \lambda_1, \dots, \lambda_n \in \mathbb{R}_+\}$$

$$L = AZ^n = \{\lambda_1 \mathbf{a}_1 + \cdots + \lambda_n \mathbf{a}_n : \lambda_1, \dots, \lambda_n \in \mathbb{Z}\}$$

$$Q_{\text{sat}} = K \cap L = \text{saturation of } A \supset Q$$

$$H = Q_{\text{sat}} \setminus Q = \text{holes in } Q_{\text{sat}}$$

$$S = \{\mathbf{a} \in Q : \mathbf{a} + Q_{\text{sat}} \subset Q\} = \text{saturation points of } Q$$

$$\bar{S} = Q \setminus S = \text{non-saturation points of } Q$$

Under the assumption above  $K$  and  $Q$  are **pointed** and  $S$  is non-empty by Problem 7.15 of [Miller and Sturmfels, 2004].

## Minimal saturation points

We now consider minimal points of  $S$  with respect to  $S$ ,  $Q$  and  $Q_{\text{sat}}$ . We call  $\mathbf{a} \in S$  an  $S$ -minimal (a  $Q$ -minimal, a  $Q_{\text{sat}}$ -minimal, resp.) if there exists no other  $\mathbf{b} \in S$ ,  $\mathbf{b} \neq \mathbf{a}$ , such that  $\mathbf{a} - \mathbf{b} \in S$  ( $Q$ ,  $Q_{\text{sat}}$ , resp.). More formally  $\mathbf{a} \in S$  is

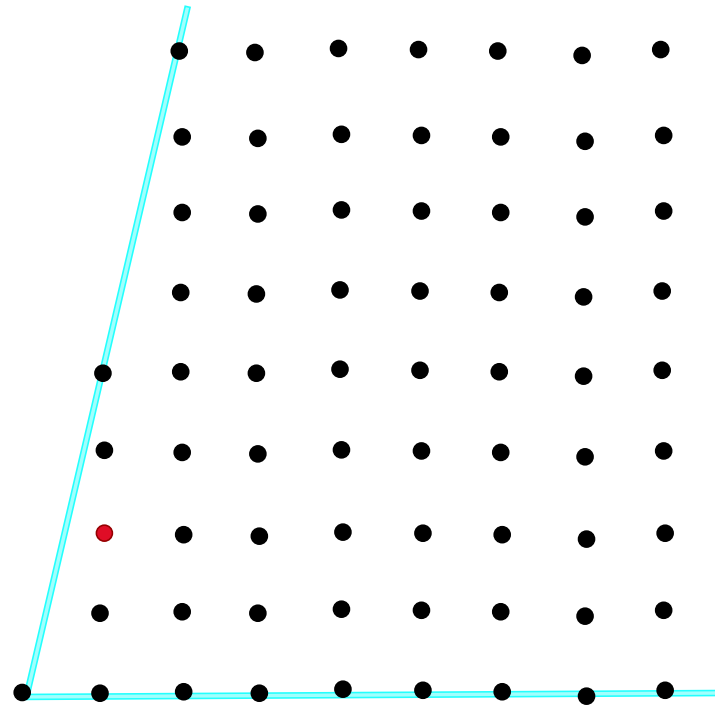
- a) an  **$S$ -minimal saturation point** if  $(\mathbf{a} + (-(S \cup \{0\}))) \cap S = \{\mathbf{a}\}$ ,
- b) a  **$Q$ -minimal saturation point** if  $(\mathbf{a} + (-Q)) \cap S = \{\mathbf{a}\}$ ,
- c) a  **$Q_{\text{sat}}$ -minimal saturation point** if  $(\mathbf{a} + (-Q_{\text{sat}})) \cap S = \{\mathbf{a}\}$ .

Let  $\min(S; S)$  denote the set of  $S$ -minimal saturation points,  $\min(S; Q)$  the set of  $Q$ -minimal saturation points, and  $\min(S; Q_{\text{sat}})$  the set of  $Q_{\text{sat}}$ -minimal saturation points.

**Note.**  $\min(S; Q_{\text{sat}}) \subset \min(S; Q) \subset \min(S; S)$ .



# Example



$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{pmatrix}.$$

## Example

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{pmatrix}.$$

$$H = \{(1, 2)^t\}.$$

$$\bar{S} = \{(0, 0)^t, (1, 0)^t, (1, 1)^t\}.$$

$$\min(S; S) = \{(1, 3)^t, (1, 4)^t, (2, 0)^t, (2, 1)^t, (2, 2)^t, (2, 3)^t, (2, 4)^t, (2, 5)^t, (3, 0)^t, (3, 1)^t, (3, 2)^t\}.$$

Thus,  $H$ ,  $\bar{S}$ , and  $\min(S; S)$  are all finite.

## Fundamental holes

**Def.** We call  $a \in H \subset Q_{\text{sat}}$ ,  $a \neq 0$ , a **fundamental hole** if

$$Q_{\text{sat}} \cap (a + (-Q)) = \{a\}.$$

Let  $H_0$  be the set of fundamental holes.

**Ex.**  $A = (3 \ 5 \ 7)$ .  $Q_{\text{sat}} = \{0, 1, \dots\}$ ,  $Q = \{0, 3, 5, 6, 7, \dots\}$ ,  $-Q = \{0, -3, -5, -6, -7, \dots\}$ .  $H = \{1, 2, 4\}$ . Among the 3 holes, 1 and 2 are fundamental. For example,  $2 \in H$  is fundamental because

$$\{0, 1, \dots\} \cap \{2, -1, -3, -4, -5, \dots\} = \{2\}.$$

On the other hand  $4 \in H$  is not fundamental because

$$\{0, 1, \dots\} \cap \{4, 1, -1, -2, -3, \dots\} = \{4, 1\}.$$

## Fundamental holes

**Lemma.** [Takemura and Y., 2006]

$H_0$  is finite.

Let  $H_0 = \{\mathbf{y}_1, \dots, \mathbf{y}_M\}$ . For each  $\mathbf{y}_h \in H_0$  and each  $\mathbf{a}_i$ , if there exists some  $\lambda \in \mathbb{Z}$  such that  $\mathbf{y}_h + \lambda \mathbf{a}_i \in Q$ , let

$$\bar{\lambda}_{hi} = \min\{\lambda \in \mathbb{Z} \mid \mathbf{y}_h + \lambda \mathbf{a}_i \in Q\}.$$

Otherwise define  $\bar{\lambda}_{hi} = \infty$ .

**Thm.** [Takemura and Y., 2006]

$H$  is finite if and only if  $\bar{\lambda}_{hi} < \infty$  for all  $h = 1, \dots, M$  and all  $i = 1, \dots, n$ .

**Thm.** [Takemura and Y., 2006]

Let  $B = \{\mathbf{b}_1, \dots, \mathbf{b}_L\}$  denote the Hilbert basis of  $Q_{\text{sat}}$ . If  $\mathbf{b}_l + \lambda \mathbf{a}_i \in Q$  for some  $\lambda \in \mathbb{Z}$ , let

$$\bar{\mu}_{li} = \min\{\lambda \in \mathbb{Z} \mid \mathbf{b}_l + \lambda \mathbf{a}_i \in Q\}$$

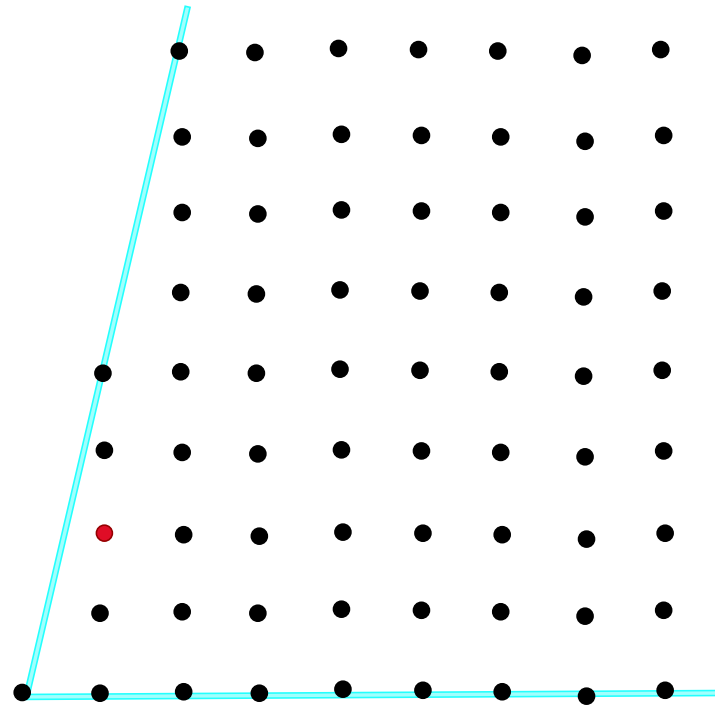
and  $\bar{\mu}_{li} = \infty$  otherwise.

Then  $H$  is finite if and only if  $\bar{\mu}_{li} < \infty$  for all  $l = 1, \dots, L$  and all  $i = 1, \dots, n$ .

**Remark.** For each  $1 \leq i \leq n$ , let  $\tilde{Q}_{(i)} = \{\sum_{j \neq i} \lambda_j \mathbf{a}_j \mid \lambda_j \in \mathbb{Z}_+, j \neq i\}$  be the semigroup spanned by  $\mathbf{a}_j, j \neq i$ . For each extreme  $\mathbf{a}_i$  and for each  $\mathbf{b}_l \notin Q$ , we only have to check

$$\mathbf{b}_l \in (-\mathbb{Z}_+ \mathbf{a}_i) + \tilde{Q}_{(i)}, \text{ for } l = 1, \dots, L.$$

# Example



$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{pmatrix}.$$

## Example

$$B = \{\mathbf{b}_1 = (1, 0)^t, \mathbf{b}_2 = (1, 1)^t, \mathbf{b}_3 = (1, 2)^t, \mathbf{b}_4 = (1, 3)^t, \mathbf{b}_5 = (1, 4)^t\}.$$

Then we can write  $\mathbf{b}_3$  as the following:

$$\begin{aligned} (1, 2)^t &= -(1, 0)^t + 2 \cdot (1, 1)^t \\ &= (1, 0)^t - (1, 1)^t + (1, 3)^t \\ &= (1, 1)^t - (1, 3)^t + (1, 4)^t \\ &= 2 \cdot (1, 3)^t - (1, 4)^t. \end{aligned}$$

We have  $\bar{\mu}_{3i} = 1$  for each  $i = 1, \dots, 4$  and  $\bar{\mu}_{li} = 0$ , where  $l \neq 3$  for each  $i = 1, \dots, 4$ . Thus by Theorem above, the number of elements in  $H$  is finite. Note that  $H$  consists of only one elements  $\{\mathbf{b}_3 = (1, 2)^t\}$ .

Ruriko Yoshida

**Thm.** [Takemura and Y., 2006]

The following statements are equivalent.

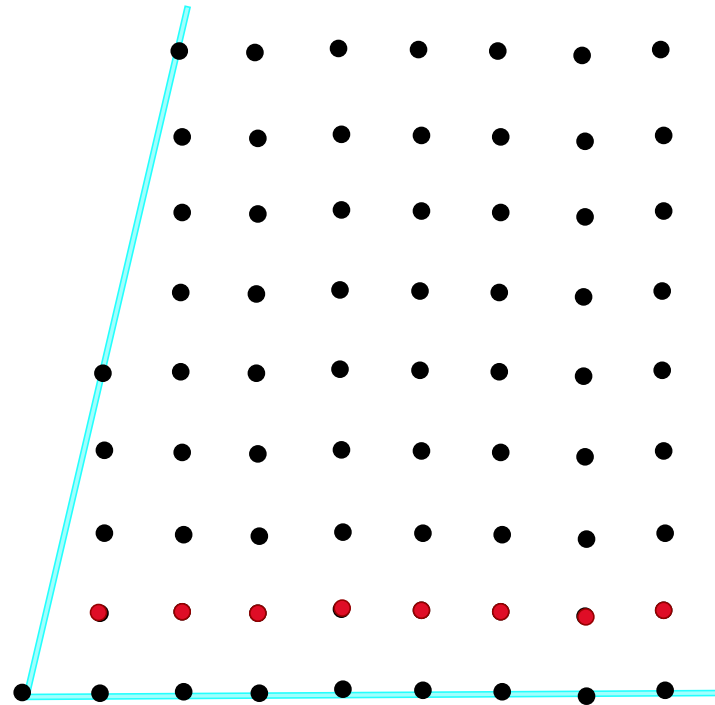
1.  $\min(S; S)$  is finite.
2.  $\text{cone}(S)$  is a rational polyhedral cone.
3. There is some  $s \in S$  on every extreme ray of  $K$ .
4.  $H$  is finite.
5.  $\bar{S}$  is finite.

**Prop.** [Takemura and Y., 2006]

$\min(S; Q)$  is finite.



# Example



$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{pmatrix}.$$

## Example

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & 4 \end{pmatrix}.$$

$H$  consists of elements  $\{(k, 1) : k \in \mathbb{Z}, k \geq 1\}$ .

$$\bar{S} = \{(i, 0)^t : i \in \mathbb{Z}, i \geq 0\},$$

$$\min(S; S) = \{(k, j)^t : k \in \mathbb{Z}, k \geq 1, 2 \leq j \leq 3\} \cup \{(1, 4)\}.$$

Thus,  $H$ ,  $\bar{S}$ , and  $\min(S; S)$  are all infinite. However,  $\min(S; Q) = \{(1, 2)^t, (1, 3)^t, (1, 4)^t\}$  is finite.

# Applications to contingency tables

$2 \times 2 \times 2 \times 2$  tables with 2-marginals.

The semigroup has 16 generators  $\mathbf{a}_1, \dots, \mathbf{a}_{16}$  in  $\mathbb{Z}^{24}$ .

The Hilbert basis of the cone generated by these 16 vectors contains 17 vectors  $\mathbf{b}_1, \dots, \mathbf{b}_{17}$ . The first 16 vectors are the same as  $\mathbf{a}_i$ , i.e.  $\mathbf{b}_i = \mathbf{a}_i$ ,  $i = 1, \dots, 16$ . The 17-th vector  $\mathbf{b}_{17}$  is

$$\mathbf{b}_{17} = (1 \ 1 \ \dots \ 1)^t$$

consisting of all 1's.

Thus,  $\mathbf{b}_{17} \notin Q$ . Then we set the 16 systems of linear equations such that:

$$P_j : \quad \mathbf{b}_1x_1 + \mathbf{b}_2x_2 + \cdots + \mathbf{b}_{16}x_{16} = \mathbf{b}_{17}$$
$$x_j \in \mathbb{Z}_-, \quad x_i \in \mathbb{Z}_+, \quad \text{for } i \neq j,$$

for  $j = 1, 2, \dots, 16$ .

Using LattE, we showed that the 16 systems of linear equations have integral solutions.

Thus by theorems above,  $H$ ,  $\bar{S}$ , and  $\min(S; S)$  are finite.

$2 \times 2 \times 2 \times 2$  tables with 2-marginals and 3-marginal i.e. [12][13][14][123] and with levels of 2 on each node.

The semigroup is generated by 16 vectors in  $\mathbb{Z}^{12}$ .

The Hilbert basis consists of these 16 vectors and two additional vectors

$$\mathbf{b}_{17} = (1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0)^t, \quad \mathbf{b}_{18} = (1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 1 \ 1)^t.$$

Thus,  $\mathbf{b}_{17}, \mathbf{b}_{18} \notin Q$ .

Then we set the system of linear equations such that:

$$\mathbf{b}_1x_1 + \mathbf{b}_2x_2 + \cdots + \mathbf{b}_{16}x_{16} = \mathbf{b}_{17}$$
$$x_1 \in \mathbb{Z}_-, x_i \in \mathbb{Z}_+, \text{ for } i = 2, \cdots, 16.$$

We solved the system via lrs, CDD and LattE.

We noticed that this system has no real solution (infeasible).

Thus by theorems above,  $H$ ,  $\bar{S}$ , and  $\min(S; S)$  are infinite.

**Prop.** [Takemura and Y., 2006]

$3 \times 4 \times 7$  table with 2-marginals has infinite number of holes.

**Sketch of pf.**

					sum
	$c$	0	0	0	$c$
	0	0	0	0	0
	0	0	0	0	0
sum	$c$	0	0	0	$c$

Table 1: the 7-th  $3 \times 4$  slice is uniquely determined by its row and its column sums.  $c$  is an arbitrary positive integer. Thus for each choice of positive integer the beginning  $3 \times 4 \times 6$  part remains to be a hole. Since the positive integer is arbitrary,  $3 \times 4 \times 7$  table has infinite number of holes.



## Future work

**Known.** Results on the saturation of 3-DIPTP are summarized in Theorem 6.4 of a paper by Ohsugi and Hibi, (2006). They show that a normality (i.e.  $Q$  is saturated) or non-normality (i.e.  $Q$  is not saturated) of  $Q$  is not known only for the following three cases:

$$5 \times 5 \times 3, \quad 5 \times 4 \times 3, \quad 4 \times 4 \times 3.$$

We want to decide whether semigroups of these tables above are normal or not.

Also we want to decide whether  $3 \times 4 \times 6$  table with 2-marginals have a finite number of holes.

Ruriko Yoshida

# Questions?

A preprint is available at arxiv:

<http://arxiv.org/abs/math.ST/0603108>

Ruriko Yoshida

Thank you....