Applications of an algorithm to compute holes of semi-groups to contingency tables

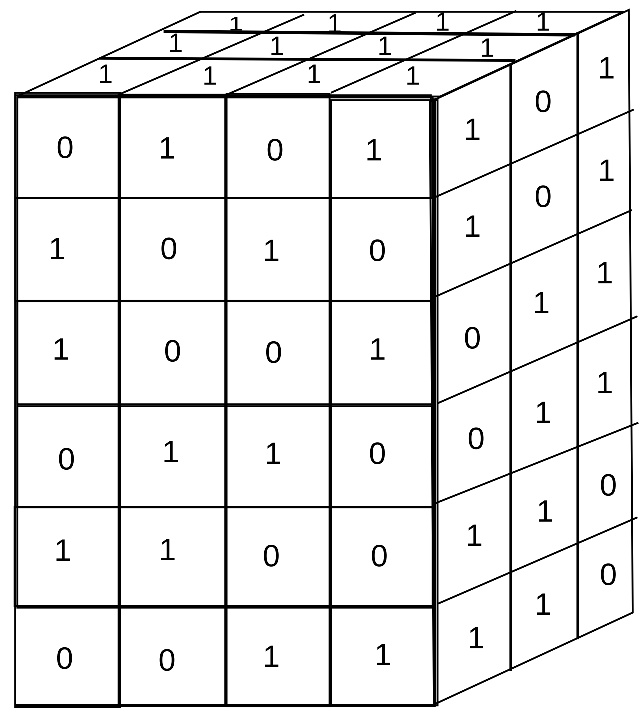
Ruriko Yoshida
Dept. of Statistics University of Kentucky

Joint work with A. Takemura and R. Hemmecke

www.ms.uky.edu/~ruriko

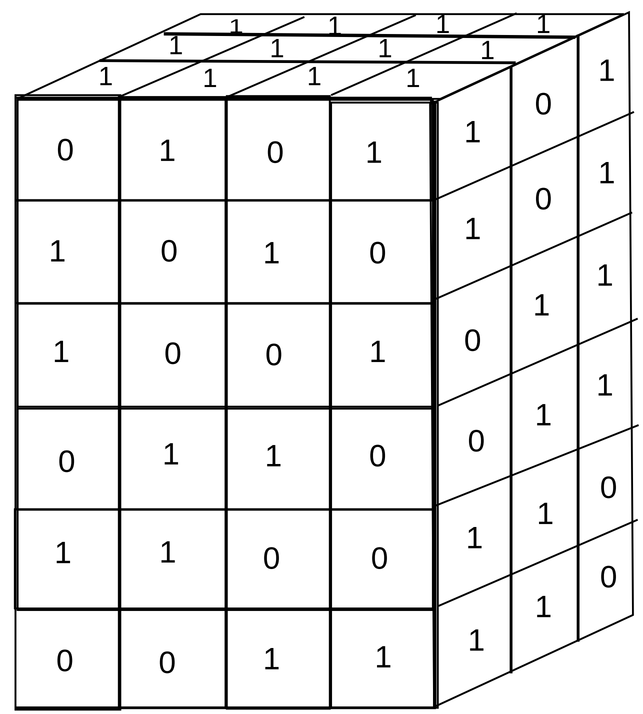
Puzzle

Is there a table satisfying these given margins?



Each cell has nonnegative integral value.

Answer



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

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

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

This is called the **multi-dimensional integer planar transportation problem** and it can be applied to **data sequrity 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 $P_b = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$. We assume that $\text{cone}(A)$ is pointed.

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

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

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

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

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

Motivation: Once we solve this problem, then we can solve an integer linear feasibility problem in a constant time if we vary the right-hand-side b.

Note: Q is normal (i.e. $H = \emptyset$) iff the Hilbert basis of cone(A) is in Q.

Note: Barvinok and Woods showed that: suppose we fix d and n. We can compute all holes of Q in polynomial time using **short rational functions** in polynomial time.

However: It is an **implicit representation** of H, and also their method cannot be implemented at this moment.

Problem: Find **an explicit representation of** H.

Fundamental holes

Def. We call $\boldsymbol{a} \in H \subset Q_{\mathrm{sat}}$, $\boldsymbol{a} \neq 0$, a **fundamental hole** if there is no other hole $h' \in H$ such that $h - h' \in Q$. Let F be the set of fundamental holes.

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

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

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

$$4 - 1 = 3 \in Q.$$

Def. The semigroup $Q_{\rm sat} = \text{cone}(A) \cap L$ is called the **saturation** of Q (i. e. $Q_{\rm sat} = Q + H$ or $H = Q_{\rm sat} - Q$).

We call $s \in Q$ a **saturation point** of Q, if $s + Q_{\rm sat} \subseteq Q$. The set of all saturation points of Q is denoted by S.

Note. If cone(A) is pointed then $S \neq \emptyset$.

$s \in S$ is called a **Q-minimal** of S there is no other $s' \in S$ with $s - s' \in Q$.

Let $f \in F$. $s \in (f + Q) \cap Q$ is called **Q-minimal** of $(f + Q) \cap Q$ if there is no other $s' \in (f + Q) \cap Q$ with $s - s' \in Q$.

$s \in S$ is called a **S-minimal** of S there is no other $s' \in S$ with $s - s' \in S$.

Example

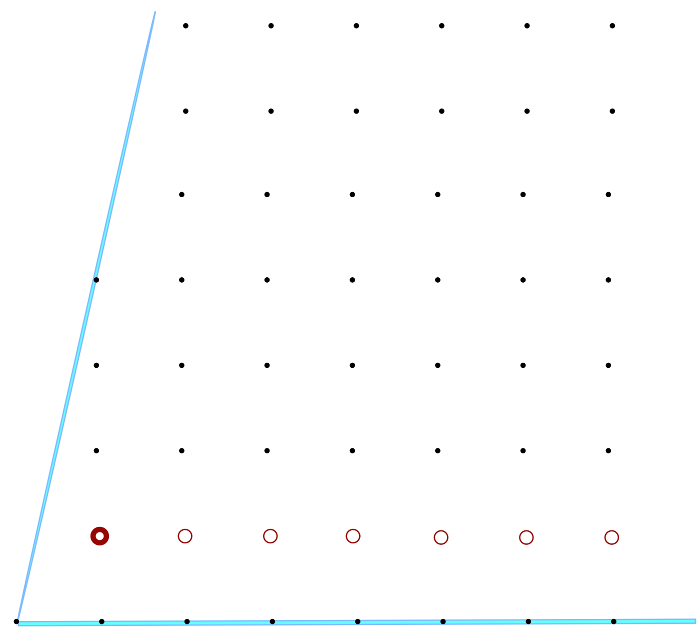


Figure 1: Non-holes, holes and fundamental hole for Example.

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

Example cont.

Q has infinitely many holes

$$H = \{(1,1)^\intercal + \alpha \cdot (1,0)^\intercal : \alpha \in \mathbb{Z}_+\},$$

out of which only $(1,1)^\intercal$ is fundamental,

Q has three Q-minimal saturation points: $(1,2)^\intercal$, $(1,3)^\intercal$, and $(1,4)^\intercal$.

The **output** from our algorithm looks like:

$$H = \{(1,1)^\intercal + \alpha \cdot (1,0)^\intercal : \alpha \in \mathbb{Z}_+\}.$$

Algorithm

Input: $A \in \mathbb{Z}^{d \times n}$.

Output: An explicit representation of H.

1. Compute the set F of fundamental holes.

2. For each of the finitely many $f \in F$, compute all $z = f + A\lambda$ where all minimal inhomogenous solutions (λ, μ) of

$$\{(\lambda, \mu) \in \mathbb{Z}_{+}^{2n} : f + A\lambda = A\mu\}. \tag{1}$$

3. From the minimal inhomogenous solutions (λ, μ) of (1), compute an explicit representation of the holes of Q in $f + Q$.

Computing fundamental holes

The set F of fundamental holes is finite, since it is a subset of

$$P := \left\{ \sum_{j=1}^{n} \lambda_j A_{.j} : 0 \leq \lambda_1, \ldots, \lambda_n < 1 \right\}.$$

Algorithm. (Computing fundamental holes)

- Compute the minimal integral generating set B of cone$(A) \cap L$.
- Check each $z \in B$ whether it is a fundamental hole or not, that is, compute $B \cap F$.
- Generate all nonnegative integer combinations of elements in $B \cap F$ that lie in P and check for each such z whether it is a fundamental hole or not.

Example cont

In our example, the lattice $L = \mathbb{Z}^2$. With this, the minimal Hilbert basis B of cone$(A) \cap L$ consists of 5 elements:

$$B = \{(1,0)^\intercal, (1,1)^\intercal, (1,2)^\intercal, (1,3)^\intercal, (1,4)^\intercal\},$$

out of which only $(1,1)^\intercal$ is a hole.

Being in B, $(1,1)^\intercal$ must be a fundamental hole. Thus, $B \cap F = \{(1,1)^\intercal\}$.

Note that $2 \cdot (1,1)^\intercal = (2,2)^\intercal \in Q$ and consequently, there is no other fundamental hole in Q, i.e. $F = \{(1,1)^\intercal\}$.

Computing minimal inhomogenous solutions

The (finitely many) minimal inhomogeneous solutions to the above linear system can be computed, for example, with `4ti2`.

Example cont. Let $f = (1,1)^\intercal$ and consider $(f+Q) \cap Q$. The linear system to solve is

$$\begin{array}{ccccccccccccccccc} 1 & + & \lambda_1 & + & \lambda_2 & + & \lambda_3 & + & \lambda_4 & = & \mu_1 & + & \mu_2 & + & \mu_3 & + & \mu_4 \\ 1 & & & + & 2\lambda_2 & + & 3\lambda_3 & + & 4\lambda_4 & = & & & 2\mu_2 & + & 3\mu_3 & + & 4\mu_4 \end{array}$$

with $\lambda_i, \mu_j \in \mathbb{Z}_+, i,j \in \{1,2,3,4\}$.

Example cont

`4ti2` gives the following 5 minimal inhomogeneous solutions (λ, μ) to system (1):

$$\begin{array}{rcl}
(\lambda,\mu) & \to & z = f + A\lambda \\
(0,0,0,2,0,0,3,0)^\intercal & \to & (3,9)^\intercal \\
(0,1,0,0,1,0,1,0)^\intercal & \to & (2,3)^\intercal \\
(0,0,1,0,1,0,0,1)^\intercal & \to & (2,4)^\intercal \\
(0,0,1,0,0,2,0,0)^\intercal & \to & (2,4)^\intercal \\
(0,0,0,1,0,1,1,0)^\intercal & \to & (2,5)^\intercal
\end{array}$$

Thus, we have $\{(2,3)^\intercal, (2,4)^\intercal, (2,5)^\intercal, (3,9)^\intercal\}$.

Computing the holes in $f + Q$

Let $I_{A,f} \in \mathbb{Q}[x_1, \ldots, x_n]$ be the monomial ideal generated by

$$I_{A,f} = \langle x^\lambda : \lambda \in \mathbb{Z}_+^n, f + A\lambda \in (f + Q) \cap Q \rangle.$$

Note. If cone(A) is pointed, there are only finitely many $\lambda \in \mathbb{Z}_+^n$ such that $f + A\lambda = z$ for each $z \in f + Q$. Thus, by solving $f + A\lambda = z, \lambda \in \mathbb{Z}_+^n$ for all minimal inhomogenuous solutions in $(f + Q) \cap Q$, we can find a finite generating set for $I_{A,f}$.

Note. While the monomial x^λ corresponds to $z = f + A\lambda \in f + Q$, we have $z \in (f + Q) \cap Q$ if and only if $x^\lambda \in I_{A,f}$. Thus, the set of holes in $f + Q$ corresponds to the set of standard monomials of the monomial ideal $I_{A,f}$.

Example cont

Construct the generators of the monomial ideal $I_{A,f}$ by finding all representations of the form $z = f + A\lambda, \lambda \in \mathbb{Z}_+^4$ for each z in $(f + Q) \cap Q$ for each $z \in \{(2,3)^\intercal, (2,4)^\intercal, (2,5)^\intercal, (3,9)^\intercal\}$.

$$
\begin{aligned}
z &= f + A\lambda \\
(2,3)^\intercal &= (1,1)^\intercal + A(0,1,0,0)^\intercal \\
(2,4)^\intercal &= (1,1)^\intercal + A(0,0,1,0)^\intercal \\
(2,5)^\intercal &= (1,1)^\intercal + A(0,0,0,1)^\intercal \\
(3,9)^\intercal &= (1,1)^\intercal + A(0,0,0,4)^\intercal
\end{aligned}
$$

Example

Thus, we get the monomial ideal

$$I_{A,f} = \langle x_2, x_3, x_4 \rangle,$$

whose set of standard monomials is $\{x_1^\alpha : \alpha \in \mathbb{Z}_+\}$.

Thus, the set of holes in $f + Q$ is

$$\{f + \alpha A_{.1} : \alpha \in \mathbb{Z}_+\} = \{(1,1)^\intercal + \alpha(1,0)^\intercal : \alpha \in \mathbb{Z}_+\}$$

Computing Q-minimal saturation points

Want to compute $\min(S; Q)$, the set of all Q-minimal saturation points of Q. Note that $\min(S; Q)$ is always finite.

We have the following equivalences:

$$
\begin{aligned}
s \in S \quad &\Leftrightarrow \quad s \in Q \text{ and } s + Q_{\mathrm{sat}} \subseteq Q \quad \text{(by definition)} \\
&\Leftrightarrow \quad s \in Q \text{ and } s + H \subseteq Q \quad (Q_{\mathrm{sat}} = Q \cup H \text{ and } s + Q \subseteq Q, \forall s \in Q) \\
&\Leftrightarrow \quad s \in Q \text{ and } s + F \subseteq Q \quad (H \subseteq F + Q) \\
&\Leftrightarrow \quad s + f \in f + Q \text{ and } s + f \subseteq Q \quad \forall f \in F \\
&\Leftrightarrow \quad s + f \in (f + Q) \cap Q \quad \forall f \in F.
\end{aligned}
$$

Consequently, we have

$$s \in S \Leftrightarrow s \in \bigcap_{f \in F} [((f + Q) \cap Q) - f]$$

and thus, with $s = A\lambda$ for some $\lambda \in \mathbb{Z}_+^n$ (as $s \in Q$), we get

$$s \in S \Leftrightarrow x^\lambda \in \bigcap_{f \in F} I_{A,f} =: I_A,$$

by definition of $I_{A,f}$. I_A is a monomial ideal and can be found algorithmically, for example with the help of Gröbner bases. The elements $s \in \min(S; Q)$ correspond exactly to the (finitely many!) ideal generators x^λ of I_A via the relation $s = A\lambda$.

Example cont

In our example, we have $I_A = I_{A,f} = \langle x_2, x_3, x_4 \rangle$, as there exists only one fundamental hole f. The three generators of I_A correspond to the three Q-minimal saturation points $(1,2)^\intercal$, $(1,3)^\intercal$, and $(1,4)^\intercal$.

Deciding whether H is finite or not

Thm. [Takemura and Y., 2006]

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

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

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

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

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

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

Thm. [Takemura and Y., 2006]

The following statements are equivalent.

1. $\min(S; S)$, the set of all S-minimal points in S, is finite.
2. 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.

Applications to contingency tables

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

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

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

$$\boldsymbol{b}_{17} = (1\ 1\ \ldots\ 1)^t$$

consisting of all 1's.

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

$$P_j : \quad \boldsymbol{b}_1 x_1 + \boldsymbol{b}_2 x_2 + \cdots + \boldsymbol{b}_{16} x_{16} = \boldsymbol{b}_{17}$$
$$x_j \in \mathbb{Z}_-, \; x_i \in \mathbb{Z}_+, \; \text{for } i \neq j,$$

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

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-margins and 3-margin 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

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

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

Then we set the system of linear equations such that:

$$\boldsymbol{b}_1 x_1 + \boldsymbol{b}_2 x_2 + \cdots + \boldsymbol{b}_{16} x_{16} = \boldsymbol{b}_{17}$$
$$x_1 \in \mathbb{Z}_-, \; x_i \in \mathbb{Z}_+, \; \text{for } i = 2, \cdots, 16.$$

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-margins 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 3×4 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:

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

Note. $4 \times 4 \times 3$ is solved! 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-margins have a finite number of holes.

Questions?

A preprint is available at arxiv:

http://arxiv.org/abs/math.CO/0607599

Thank you....