An algorithm to compute holes of semi-groups

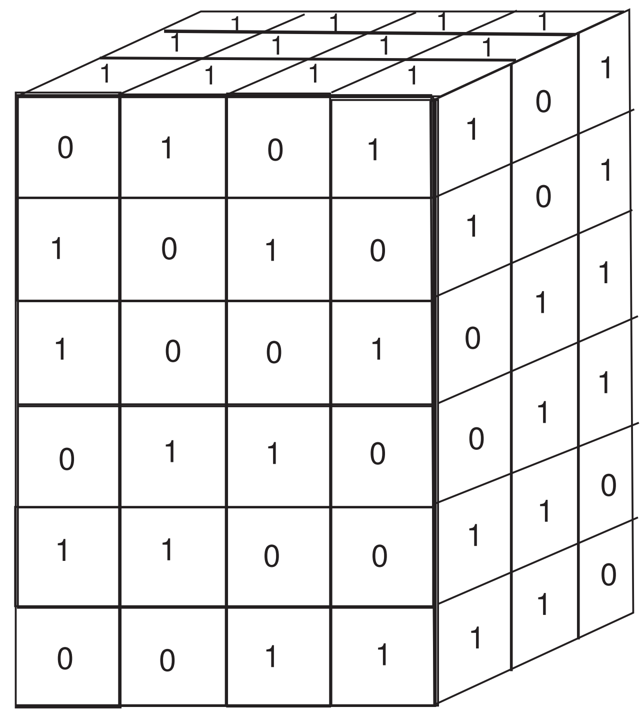
Ruriko Yoshida
Dept. of Statistics University of Kentucky

Joint work with A. Takemura and R. Hemmecke

www.ms.uky.edu/~ruriko

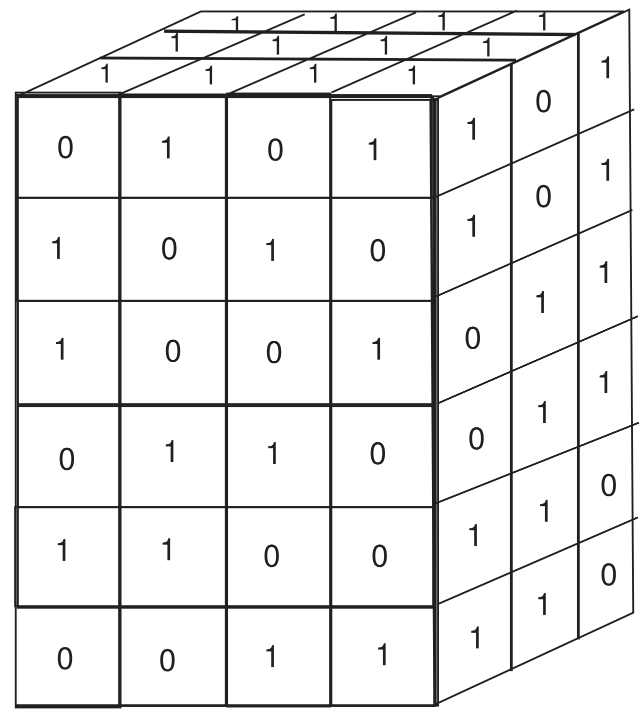
Puzzle

Is there a nonnegative integral valued table satisfying these given margins?



Each cell has nonnegative integral value.
Hint: There exists a nonnegative real valued table satisfying the constraints.

Answer



There does not exist such a nonnegative integral valued 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.

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

Observation

Let Q be the semigroup generated by the columns $\boldsymbol{a}_i$ of A, that is,

$Q = \{\sum_{i=1}^{n} \alpha_i \boldsymbol{a}_i : \alpha_i \in \mathbb{Z}_+\} \subset \text{cone}(A) \cap \mathbb{Z}^d$.

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

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

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

Motivation: One of motivations is that once we solve this problem, then we can solve an integer linear feasibility problem efficiently 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**.

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**.

Example

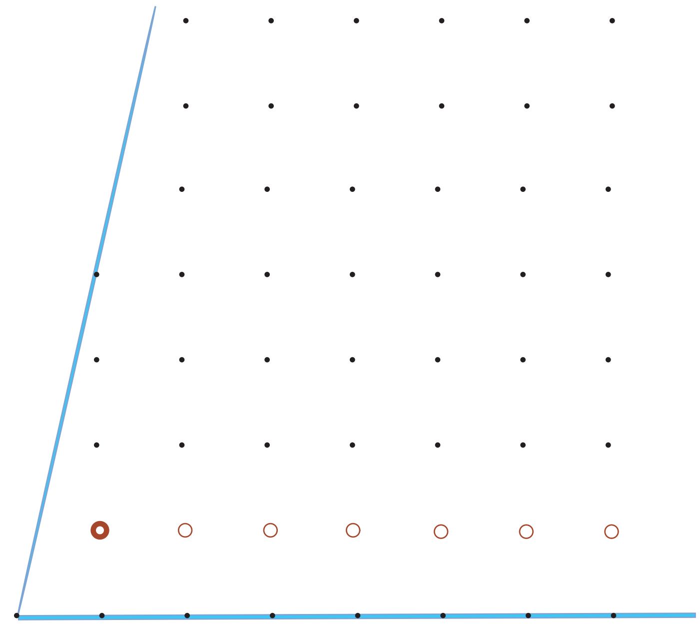


Figure 1: Red dots represent holes.

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

Fundamental holes

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 $H = Q_{\rm sat} - Q$).

Def. We call $\boldsymbol{a} \in H \subset Q_{\rm sat}$, $\boldsymbol{a} \neq 0$, 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. $A = (3\ 5\ 7)$. $Q_{\rm sat} = \{0, 1, \ldots\}$, $Q = \{0, 3, 5, 6, 7, \ldots\}$, $H = \{1, 2, 4\}$. 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

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

Example

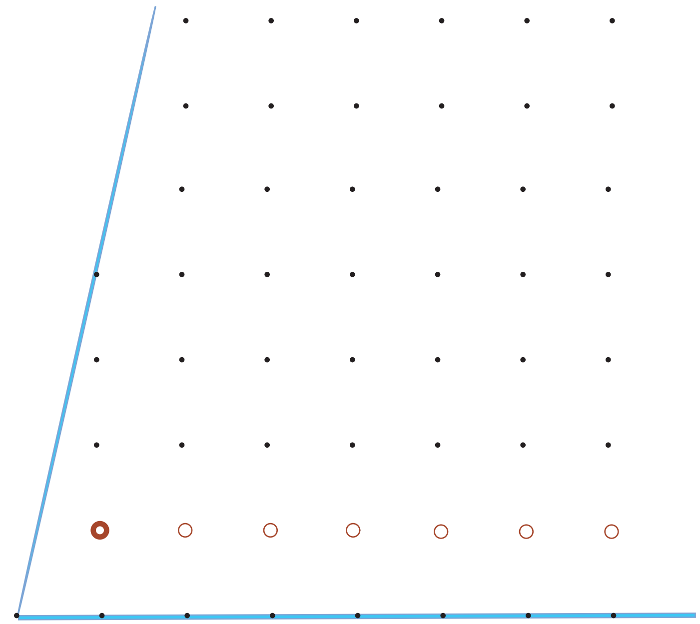


Figure 2: Non-holes, holes and fundamental hole for Example.

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

Example cont.

Q has infinitely many holes

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

out of which only $(1,1)^{\mathsf{T}}$ is fundamental,

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

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

Computing the holes in $f + Q$

Let $f \in F$ and $I_{A,f} \in \mathbb{Q}[x_1, \ldots, x_n]$ 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 $I_{A,f}$.

Theorem. (Hemmecke, Takemura, Y, 2006) 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 $I_{A,f}$.

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 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 the lattice points in

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

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_{sat}, 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):

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

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

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

$$\{f + \alpha A_1 : \alpha \in \mathbb{Z}_+\} = \{(1,1)^{\mathsf{T}} + \alpha(1,0)^{\mathsf{T}} : \alpha \in \mathbb{Z}_+\}$$

Applications to contingency tables

Sequencial Importance Sampling (SIS)

For a formal definition of SIS, see (Chen, 2001), (Chen, Diaconis, Holmes, Liu 2005), (Chen, Dinwoodie, Sullivant, 2006), etc., etc.

How does SIS work?
For example, suppose we have the following table.

7	5	1	13
5	10	6	21
2	6	8	16
14	21	15	50

Now we consider τ, all integral valued tables with the same column sums c_i and row sums r_i for $i = 1, 2, 3$.

Example cont...

We want to sample a table from τ. We pick an integer from $[0, \min\{8, 9\}]$ with some distribution (say a uniform distribution). For example, we picked 5.

5	?	?	13
?	?	?	21
?	?	?	16
14	21	15	50

Then update r_1 and c_1 as follows:

5	?	?	8
?	?	?	21
?	?	?	16
9	21	15	50

Example cont...

We do this process untill we fill up all cells. Then we get a table:

5	7	1	13
7	8	6	21
2	6	8	16
14	21	15	50

Questions: How can we choose a sample which does not end up a non-consistant table? Relations between holes and samples.

SIS and holes

Suppose $(T_{i_1,\cdots,i_m}) \in \tau$ be a $d_1 \times \cdots \times d_m$ table and we set:

$$b = Ax,$$

where $x = (T_{1,1,\cdots,1}, T_{1,1,\cdots,2}, \cdots, T_{d_1,d_2,\cdots,d_m})$ and
$b = (\sum_{i_1} T_{i_1,\cdots,i_m}, \cdots, \sum_{i_m} T_{i_1,\cdots,i_m})$.

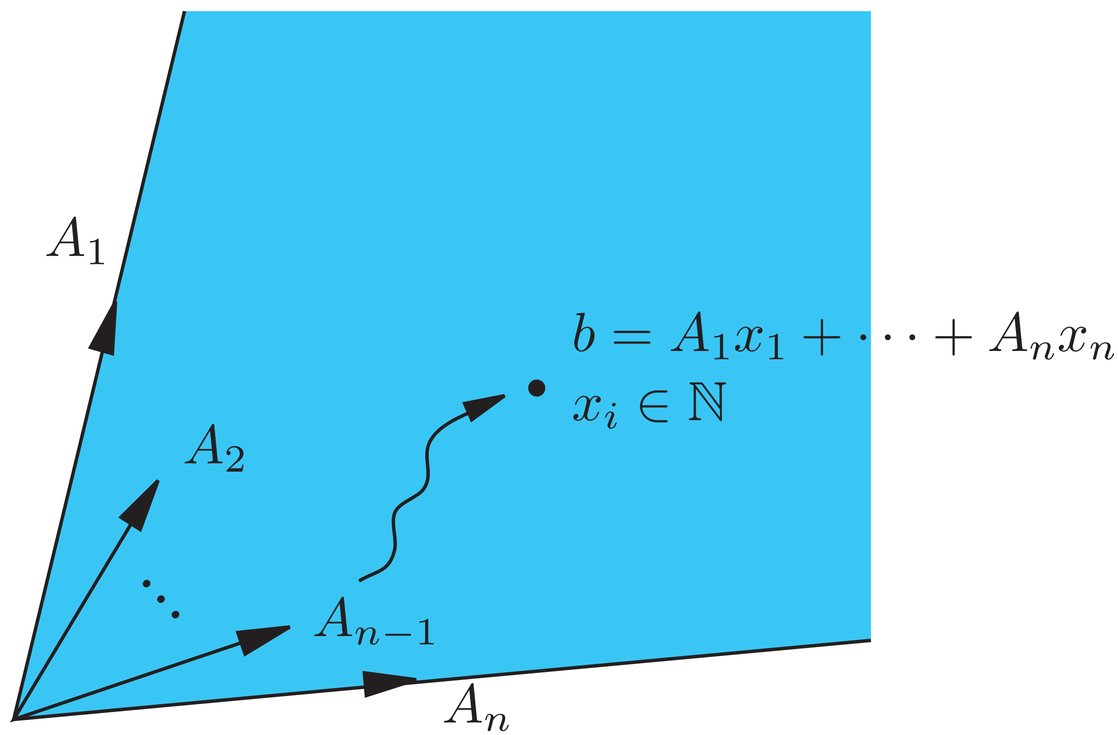
Thus we can rewrite:

$$b = A_1x_1 + A_2x_2 + \cdots + A_nx_n,\ x_i \in \mathbb{Z}_+.$$

To get a table satisfying the given marginals, we take a path

$$A_1x_1 \to A_1x_1 + A_2x_2 \to \cdots \to A_1x_1 + A_2x_2 + \cdots + A_nx_n.$$

From the view of Q



Taking a sample via SIS can be viewed as a path from the origin to $b \in Q$.

SIS and holes

Suppose Q is not normal (such as three-way tables).

To sample via SIS, we need to check if it is possible to reach b from the current point $s = A_1y_1 + A_2y_2 + \cdots + A_ny_n$.

To check there is a path from s to b by adding $z_i \in \mathbb{Z}_+$ to y_i for some i:

- If $b - s \in Q$, then there is a path.
- If $b - s \in H$, then we reject.

Thus knowing H, one might be speed up some computation of SIS. (we need to investigate how practical our algorithm is).

Finiteness of holes of Q

Theorem: (Takemura and Y, 2006): Suppose we fix d and n. Then, there is a polynomial time algorithmm to decide whether the set of holes H of Q for a matrix A is finite or not.

Examples: The matrix for defining $2 \times 2 \times 2 \times 2$ tables with 2-marginals has finitely many holes.

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

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 3×4 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....