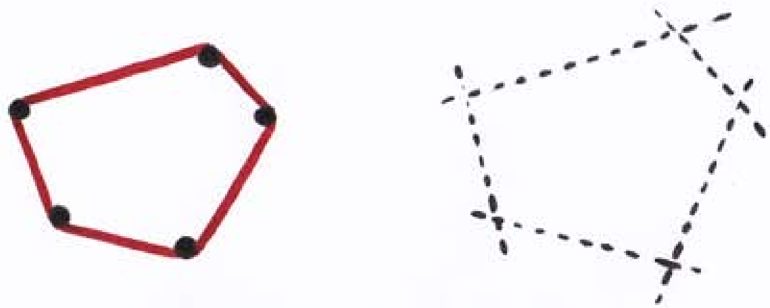
Polytopes

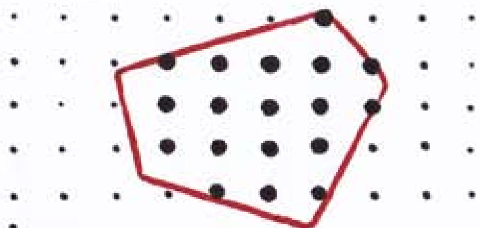
Generalization of polygon (2 dim.), polyhedron (3 dim.), etc. to $\mathbb{R}^n$.

Definition 1: Convex hull of finitely many points in $\mathbb{R}^n$.

Definition 2: Bounded region formed by intersection of finitely many half-spaces.



We want to count integer points:



Objects encoded by integer points: network flows, contingency tables, magic squares, Latin squares, knapsack packings, matchings...

Contingency tables

A contingency table is a nonnegative integer matrix.

EXAMPLE

$$
\begin{array}{c|cccc|c}
 & \text{Black} & \text{Brown} & \text{Red} & \text{Blond} & \Sigma \\
\hline
\text{Brown} & 68 & 119 & 26 & 7 & 220 \\
\text{Blue} & 20 & 84 & 17 & 94 & 215 \\
\text{Hazel} & 15 & 54 & 14 & 10 & 93 \\
\text{Green} & 5 & 29 & 14 & 16 & 64 \\
\hline
\Sigma & 108 & 286 & 71 & 127 & 592
\end{array}
$$

row margins

column margins

Cross-tab of eye, hair color.

$\Pi(R,C)$: polytope of nonneg. matrices with given margins $R=(r_1,\dots,r_m)$, $C=(c_1,\dots,c_n)$. This is the <u>transportation polytope</u>.

Applications:

- Cross-tabulation of categorical variables in a sample; significance testing.
- Program for moving goods from sites of supply to sites of demand; flow optimization.

Special- and generalizations:

- Magic squares
- Flows in arbitrary graphs
- Capacitated flows
- Higher-order tensors
- Latin squares/rook arrays

Counting integer points

Pick's Theorem (1899).

If P is a polygon with vertices in $\mathbb{Z}^2$, then

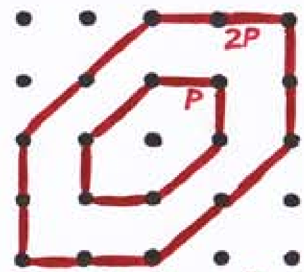
$$\mathrm{Area}(P) = I + \tfrac{1}{2}B - 1.$$

\# integer points in interior

\# integer points on boundary

Ehrhart's Theorem (1967).

If P is a polytope with vertices in $\mathbb{Z}^n$, then there is a polynomial $\ell_P(t)$ which gives $|tP \cap \mathbb{Z}^n|$ whenever t is a nonnegative integer.



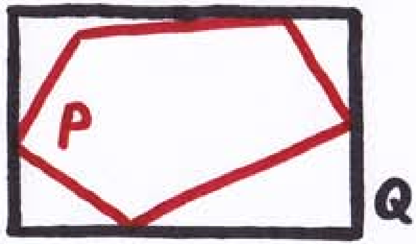
Ex.: $\ell_P(t) = 3t^2 + 3t + 1$

Implementation:

- Polynomial time when n fixed (Barvinok 1994; LattE)
- \#P-complete in general

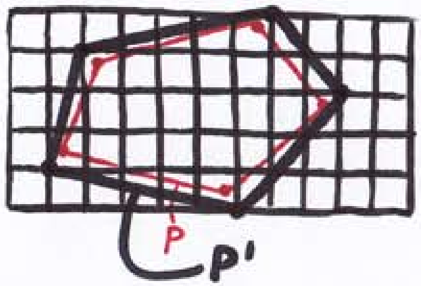
Coping in general dimension

Monte Carlo ("dart-throwing"):



- Choose points from $Q \cap \mathbb{Z}^n$ uniformly
- Observe $\frac{|P \cap \mathbb{Z}^n|}{|Q \cap \mathbb{Z}^n|}$

MC + dynamic programming:



- Replace P by a fixed-resolution simulacrum P' whose integer points can be counted iteratively in polynomial time
- Use iterates to throw darts at P'

MC + Markov chains:

- Sample $P \cap \mathbb{Z}^n$ by a random walk
- Must show good mixing
- Near-uniform sampling $\leftrightarrow$ appx. counting (Jerrum, Valiant, Vazirani)

Entropy methods ...

Independence models

Boxes are nice: If X is a uniform random integer point, then for $1 \leq j_1 < j_2 < \cdots < j_r \leq n$ and $a_1, a_2, \ldots, a_r \in \mathbb{Z}$,

$$\Pr\left[\bigwedge_{i=1}^{r} X_{j_i} = a_i\right] = \prod_{i=1}^{r} \Pr[X_{j_i} = a_i]. \qquad (\heartsuit)$$

Might $(\heartsuit)$ hold <u>approximately</u> for other polytopes, when $r \ll \dim P$?

(Inspiration: Phenomena in high-dimensional convex geometry...)

Definition.

An <u>independence model</u> is a random vector $X = (X_1, \ldots, X_n)$ supported on $\mathbb{Z}^n$ which satisfies $(\heartsuit)$. In other words, the coordinates are independent.

An independence model for a simplex truncated by a cube

Let $P = \{(x_1, \ldots, x_n) \in [0,1]^n : \quad x_1 + \cdots + x_n = r\}$.

If Y is uniformly distributed on $P \cap \mathbb{Z}^n$, then Y_j is 0–1 Bernoulli and has mean r/n.

Let X be the independence model with $X_j \sim Y_j$, $1 \le j \le n$.

Then X hits every point of $P \cap \mathbb{Z}^n$ at rate $e^{-nh(r/n)}$, where $h(p) = -p \ln p - (1-p) \ln (1-p)$.

Also, $\Pr(X \in P) \approx \sqrt{\dfrac{n}{2\pi r(n-r)}}$, so

$$|P \cap \mathbb{Z}^n| \approx e^{nh(r/n)} \sqrt{\frac{n}{2\pi r(n-r)}}.$$

We recover the entropy formula for $\binom{n}{r}$:

Proposition. $\ln \binom{sn}{sr} = sn \cdot h\left(\frac{r}{n}\right) - \Theta(\ln s)$.

Entropy

Entropy is a statistic associated to a random variable:

$$H[X] = \sum_{x \in \operatorname{supp} X} \Pr[X=x] \ln \frac{1}{\Pr[X=x]}.$$

(Assume X discrete, but $H[X]$ can still be ∞.)

Basic properties:

- $H[X]$ is strictly concave w.r.t. the p.m.f. of X.
  (Thus $H[X] \le \ln |\operatorname{supp} X|$, achieved by the uniform distr.)
- $H[X,Y] \le H[X] + H[Y]$,
  with equality iff. X, Y independent.

Maximum entropy principle:

Given partial information about a random variable X, Jaynes proposed the maximum-entropy distribution among all candidates as the *likeliest*...

The maximum-entropy independence model

Let $P = \{x \in \mathbb{R}^n : \quad x \geq 0, \quad \underset{A \in \mathbb{R}^{m\times n} \text{ of rank } m}{Ax} = \underset{b\in\mathbb{R}^m}{b}, \quad (x \leq \underset{k \in (\mathbb{R}\cup\{\infty\})^n}{k})\}.$

Let $X = (X_1, \ldots, X_n)$ be a random vector with max. entropy subject to:

- $\operatorname{supp} X \subseteq \mathbb{Z}^n_{\geq 0}$
- $\mathbb{E}[AX] = b \qquad$ (equiv.: $\mathbb{E}[X] \in P$)

Properties:

- X exists and has unique distribution
- X is an independence model (called the MEIM)
- X has constant mass $e^{-H[X]}$ on points of $P \cap \mathbb{Z}^n$
- Coordinates of X are (truncated) geometric:

$$\Pr[X_j = t] = p_j q_j^t \qquad (t = 0, 1, 2, \ldots, k_j)$$

for some parameters $p_j \in [0,1], \quad q_j \in [0, \infty]$

(Note: $q_j = \infty$ requires special interpretation; also, $q_j \in [0,1)$ if $k_j = \infty$)

Truncated geometric distribution

For $\kappa \in \mathbb{Z}_{>0} \cup \{\infty\}$, $0 \le x \le \kappa$, we say $X \sim TG(x;\kappa)$ if

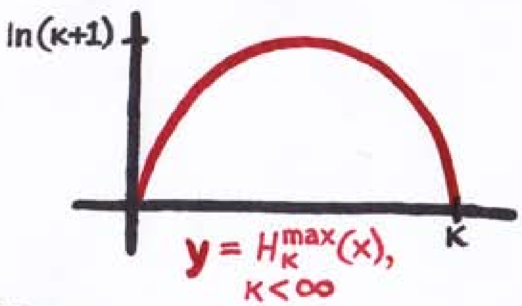
$$\Pr(X=t) = pq^t \quad \text{for} \quad t = 0,1,\ldots(,\kappa),$$

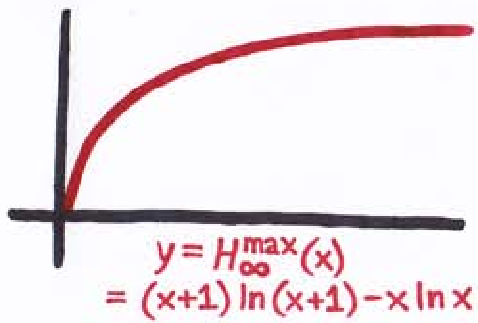
where $p = p(x;\kappa)$, $q = q(x;\kappa)$ satisfy

$$\begin{cases} 1 = p(1+q+q^2+\cdots+q^\kappa) \\ x = p(q+2q^2+\cdots+\kappa q^\kappa) \end{cases}.$$

This is the max. ent. distr. on $\{0,1,\ldots,\kappa\}$ with mean x.

Let $H_\kappa^{\max}(x)$ denote its entropy.





Properties:

- Nice formulas for $p, q, H_\kappa^{\max}$ when $\kappa = 1$ or $\kappa = \infty$
- $H_\kappa^{\max}(x)$ is strictly concave in x
- $H_\kappa^{\max}(x) = -[\ln p + x \ln q]$
- $(H_\kappa^{\max})'(x) = -\ln q$

Using the MEIM to count integer points

Let X be the MEIM for P.

Then $|P \cap \mathbb{Z}^n| = \underbrace{e^{H[X]}}_{\text{This is the max. of a strictly concave function on a polytope, so it can be computed in polynomial time.}} \underbrace{\Pr[X \in P]}_{\text{We want to estimate this! Equivalently, to estimate } \Pr[AX=b].}.$

The "I-bound":

Theorem. Let $P = \{x \in \mathbb{R}^n : x \geq 0,\ Ax = b\}$, where $A = \left[\vec{a}_1 \middle| \vec{a}_2 \middle| \cdots \middle| \vec{a}_n\right] \in \mathbb{R}^{m \times n}$ is of rank m and $b \in \mathbb{R}^m$.

Let X be the MEIM associated to P, and
$q_j = q(\mathbb{E}[X_j];\ \kappa)$, $1 \leq j \leq n$.

Then
$$|P \cap \mathbb{Z}^n| \leq e^{H[X]} \min_{\substack{a_{j_1}, \ldots, a_{j_m} \\ \text{lin. indep.}}} \prod_{i=1}^{m} \underbrace{(1 - q_{j_i})}_{\text{or } \frac{1}{1+\mathbb{E}(X_{j_i})}}.$$

The Littlewood-Offord problem

L., O. asked: How many subsets of $\{a_1, \dots, a_n\}$ can have equal sum?

When $a_1, \dots, a_n \in \mathbb{R} \setminus \{0\}$, the answer is $\binom{n}{\lfloor n/2 \rfloor}$ (Erdős, 1945).

The case of $a_1, \dots, a_n \in \mathbb{R}^m \setminus \{0\}$ was addressed by Halász:

Theorem (Halász, 1977).

Let $a_1, \dots, a_n \in \mathbb{R}^m$. Let $\varepsilon_1, \dots, \varepsilon_m$ be independent symmetric Bernoulli random variables, and let $S := \varepsilon_1 a_1 + \cdots + \varepsilon_n a_n$.

Define $\operatorname{conc}_{1}(S) := \max\limits_{y \in \mathbb{R}^m} \Pr\left(|S - y| < 1 \right)$.

Suppose there exists $\delta > 0$, such that for any $|e| = 1$ one can select at least δn vectors a_k with $|\langle a_k, e \rangle| \geq 1$. Then

$$\operatorname{conc}_{1}(S) \leq c(\delta, m)\, n^{-m/2},$$

where $c(\delta, m)$ depends only on δ and m.

By a scaling limit argument, we can replace each 1 by an arbitrarily small $\varepsilon > 0$ and pass to a "point concentration" version.

The "E-bound"

Theorem. Let $P = \{x \in \mathbb{R}^n : \; x \geq 0, \; Ax = b\}$, where

$A = [\vec{a}_1 | \vec{a}_2 | \cdots | \vec{a}_n] \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Let X be the MEIM,

and suppose $2 \leq \mathbb{E}[X_j] < N$ for all $j = 1, \ldots, n$.

Suppose $n = pm$ and $\hat{a}_i = \hat{a}_{m+i} = \hat{a}_{2m+i} = \cdots = \hat{a}_{(p-1)m+i}$ for $i = 1, \ldots, m$, where $\{a_1, \ldots, a_m\}$ is a basis for $\mathbb{R}^m$.

Then for fixed N, m, we have

$$|P \cap \mathbb{Z}^n| \leq (1 + o(1)) e^{H[X]} \prod_{i=1}^{m} \left[\frac{\pi}{6} \sum_{t=1}^{p} \left(\lfloor \mathbb{E}[X_{(t-1)m+i}] + 1 \rfloor^2 - 1 \right) \right]^{-1/2}$$

as $p \to \infty$.

$\mathrm{conc}(Y) := \max_{y \in \mathrm{supp}\, Y} \Pr[Y = y]$

Key lemmas:

- If $X_1, \ldots, X_p$ are indep. discrete RVs s.th. $\mathrm{conc}(X_j) \leq \frac{1}{N_j} \; \forall j$, then $\mathrm{conc}(X_1 + \cdots + X_p)$ is maximized by X_j uniformly distributed on $\{0, 1, \ldots, N_j - 1\}$.

- Bender's local limit theorem for RVs with log-concave p.m.f.

The "H-bound"

Theorem (executive summary).

For every integer $m \geq 1$ and real $\varepsilon > 0$, there exists $\delta = \delta(m, \varepsilon) > 0$, such that the following holds:

Let P be a polytope in the form $P = \{x \in \mathbb{R}^n : x \geq 0,\ Ax = b\}$, where $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{R}^m$, and the columns of A span $\mathbb{R}^m$ p times $(p \geq 1)$. Let X be the MEIM for P. Then if $\mathbb{E}[X_j] \geq \varepsilon$ for $1 \leq j \leq n$, we have

$$|P \cap \mathbb{Z}^n| \leq \delta p^{-m/2} e^{H[X]}.$$

The full result includes explicit constants:

$$|P \cap \mathbb{Z}^n| \leq \left(p^{-m/2} \cdot C + (C')^p\right) e^{H[X]},$$

where C, C' are given in terms of $\min_{1 \leq t \leq p} \mathbb{E}[X_{(t-1)m+i}]$, $i = 1, \ldots, m$ (where the first p blocks of m columns of A are assumed to be bases).

Numerical examples

- Contingency tables with $R=(220, 215, 93, 64)$, $C=(108, 286, 71, 127)$:

  H-bound: 8.01×10^{26} I-bound: 7.14×10^{18}
  Actual count: 1.23×10^{15}

- $3\times 3\times 3$ tables with layer sums equal to 20:

  H-bound: 3.66×10^{20} I-bound: 7.00×10^{19}
  Actual count: 6.43×10^{14}

- Integer points of a simplex $2x_1+11x_2+18x_3+\cdots+14x_{25}=5000$:

  H-bound: 2.00×10^{44} I-bound: 1.07×10^{44}
  E-bound: 1.04×10^{44} Actual count: 8.57×10^{42}

- Integer points of the simplex $x_1+x_2+\cdots+x_{10000}=100$:

  H-bound: 1.774×10^{242}
  Actual count: 1.755×10^{242}

Bounded contingency tables

The truncated transportation polytope:

$$\Pi_K(R,C) := \Big\{ X \in \mathbb{R}_{\geq 0}^{m\times n} : \sum_{j=1}^{n} x_{ij} = r_i \quad (i=1,\dots,m),$$
$$\sum_{i=1}^{m} x_{ij} = c_j \quad (j=1,\dots,n),$$
$$x_{ij} \leq k_{ij} \quad \text{for all } i,j \Big\},$$

where

$$R = (r_1,\dots,r_m) \in \mathbb{Z}_{\geq 0}^{m}, \qquad C = (c_1,\dots,c_n) \in \mathbb{Z}_{\geq 0}^{n},$$
$$N = r_1 + \cdots + r_m = c_1 + \cdots + c_n, \quad \text{and} \quad K = (k_{ij}) \in \left(\mathbb{Z}_{\geq 0} \cup \{\infty\}\right)^{m\times n}.$$

The integer points of $\Pi_K(R,C)$ are K-bounded contingency tables.
Let $T_K(R,C)$ be the # of such tables.
Let $I_K(R,C)$ be the <u>independence estimate</u>

$$\begin{pmatrix} \text{\# tables w/ sum} \\ \text{of entries} = N \end{pmatrix} \cdot \begin{pmatrix} \text{fraction of those} \\ \text{tables having row} \\ \text{sums } r_1,\dots,r_m \end{pmatrix} \cdot \begin{pmatrix} \text{fraction having} \\ \text{column sums} \\ c_1,\dots,c_n \end{pmatrix}.$$

Counting contingency tables via permanents

Theorem (Barvinok).

Let $R \in \mathbb{Z}^m_{\geq 0}$, $C \in \mathbb{Z}^n_{\geq 0}$, $|R| = |C| = N$ as before.
Let $W = (w_{ij}) \in \mathbf{R}^{m \times n}$. Let $\Gamma = (\gamma_{ij})$ be an $m \times n$ matrix whose entries are independent exponential random variables of mean 1. Let $A = A(\Gamma)$ be the $N \times N$ matrix obtained by replacing each entry (i,j) of Γ by an $r_i \times c_j$ block with all entries equal to $w_{ij}\gamma_{ij}$.

Let each contingency table X with margins (R, C) be counted with weight $w(X) = \prod_i \prod_j w_{ij}^{x_{ij}}$.

Then the total weight of all such tables is

$$T(R, C; W) := \frac{\mathbb{E}[\operatorname{per} A]}{r_1! \cdots r_m! \, c_1! \cdots c_n!}.$$

Using matrix scaling and bounds on the permanent of a doubly stochastic matrix, we can compute $\mathbb{E}[\operatorname{per} A]$ to within a factor $N^{O(m+n)}$ in polynomial time.

Adapting Barvinok's theorem to bounded tables

$\boxed{\begin{array}{l}\text{Assume} \\ k_{ij} < \infty \\ \forall i,j.\end{array}}$

$$\begin{array}{cc}
\left[\begin{array}{cccc}
\le k_{11} & \le k_{12} & & \\
\le k_{21} & & & \\
& & \ddots & \\
& & & \le k_{mn}
\end{array}\right] & \begin{array}{c} r_1 \\ r_2 \\ \vdots \\ r_m \end{array} \\
\begin{array}{cccc} c_1 & c_2 & \cdots & c_n \end{array} &
\end{array}
\quad \xleftrightarrow{\text{bijection}} \quad
\begin{array}{cc}
\left[\begin{array}{ccccccccccccc}
\bullet & \bullet & \cdots & \bullet & & & & & & & & & \\
& & & & \bullet & \bullet & \cdots & \bullet & & & & & \\
& & & & & & & & \ddots & & & & \\
& & & & & & & & & \bullet & \bullet & \cdots & \bullet \\
\bullet & & & & \bullet & & & & & \bullet & & & \\
& \bullet & & & & \bullet & & & \cdots & & \bullet & & \\
& & \ddots & & & & \ddots & & & & & \ddots & \\
& & & \bullet & & & & \bullet & & & & & \bullet
\end{array}\right] & \begin{array}{c} r_1 \\ r_2 \\ \vdots \\ r_m \\ \tilde{c}_1 \\ \tilde{c}_2 \\ \vdots \\ \tilde{c}_n \end{array} \\
\begin{array}{ccccccccccccc} k_{11} & k_{12} & \cdots & k_{1n} & k_{21} & k_{22} & \cdots & k_{2n} & \cdots & k_{m1} & k_{m2} & \cdots & k_{mn} \end{array} &
\end{array}$$

($m \times n$ bounded table)

($(m+n) \times (mn)$ table with entries $\bullet$ counted at weight 1, others at weight 0)

$$T_K(R,C) = T(\mathcal{R}, \mathcal{C}; W)$$

where $\mathcal{R} = (r_1, \ldots, r_m, \tilde{c}_1, \ldots, \tilde{c}_n)$, $\mathcal{C} = (k_{11}, \ldots, k_{1n}, k_{21}, \ldots, k_{2n}, \ldots, k_{m1}, \ldots, k_{mn})$,

$$\tilde{c}_j = \sum_{i=1}^{m} k_{ij} - c_j .$$

$T_K(R,C)$ is *approximately* log-concave w.r.t. R, C; that is, $\ln T_K(R,C)$ can be written as the sum of a concave function $f(R,C)$ and a function bounded by $O\left(m + n + \sum_i \ln r_i + \sum_j \ln c_j\right)$.

Exact and approximate generating functions

From the last slide: $\ln T_K(R,C) = \underbrace{f(R,C)}_{\substack{\text{concave}\\ \text{approximation}\\ \text{from above}}} + \underbrace{\text{remainder}}_{O(m+n+\sum \ln r_i + \sum \ln c_j)}$.

Let

$$G(\vec{x},\vec{y}) = G_K(\vec{x},\vec{y}) := \prod_{i=1}^{m}\prod_{j=1}^{n}\left[1 + x_i y_j + (x_i y_j)^2 + \cdots + (x_i y_j)^{k_{ij}}\right]$$
$$= \sum_R \sum_C T_K(R,C)\, x^R y^C.$$

Let $\tilde{G}(\vec{x},\vec{y}) = \sum_R \sum_C e^{f(R,C)} x^R y^C$.

Upper bound: $T_K(R,C) \le \inf_{x_i, y_j > 0} \dfrac{G(\vec{x},\vec{y})}{x^R y^C}$.

Lower bound:

Given R,C, may choose x_*, y_* so that $e^{f(R,C)} x_*^R y_*^C$ is the largest term in the expansion of $\tilde{G}(x_*, y_*)$.

Thus $\dfrac{\tilde{G}(x_*, y_*)}{x_*^R y_*^C} \le e^{f(R,C)} \cdot (\#\text{ nonzero terms of } \tilde{G})$,

so $\inf_{x_i, y_j > 0} \dfrac{G(\vec{x},\vec{y})}{x^R y^C} \le T_K(R,C) \cdot \underbrace{\left(\#\substack{\text{nonzero}\\ \text{terms of } \tilde{G}}\right)}_{\le \prod_i (1+\sum_j k_{ij}) \cdot \prod_j (1+\sum_i k_{ij})} \cdot \underbrace{e^{\text{remainder}}}_{e^{O(m+n+\sum \ln r_i + \sum \ln c_j)}}$.

Asymptotic formulas for $\ln T_K(R,C)$, part 1

Theorem. Let d be fixed, and let k_{ij} vary between fixed positive bounds. Then

$$\ln T_K(R,C) = \ln\left(\inf_{x_i,y_j>0} \frac{G(x,y)}{x^R y^C}\right) + O(\max\{m,n\} \ln \max\{m,n\}) \qquad (\blacklozenge)$$

uniformly for R,C satisfying

$$\max\{r_1,\dots,r_m\} \le \frac{d}{m}(r_1+\cdots+r_m),$$
$$\max\{c_1,\dots,c_n\} \le \frac{d}{n}(c_1+\cdots+c_n).$$

Moreover, the first term in $(\blacklozenge)$ is $e^{\Omega(mn)}$, so

$$\ln T_K(R,C) \sim \ln \inf_{x_i,y_j>0} \frac{G(x,y)}{x^R y^C}$$

as $m,n \longrightarrow \infty$.

Asymptotic formulas for $\ln T_K(R,C)$, part 2

The MEIM for $\Pi_K(R,C)$ is a matrix $X = (X_{ij})$ of indep. TG random variables with mean $Z \in \Pi_K(R,C)$ maximizing $\sum_i \sum_j H^{max}_{k_{ij}}(z_{ij})$. $(= H[X])$

We know that X is uniform on the int. pts. of $\Pi_K(R,C)$, so

$$H[X] \geq \ln T_K(R,C).$$

This upper bound on $\ln T_K(R,C)$ turns out to be asymptotic to the true value, thanks to

Lemma. $$\ln\left(\inf_{x_i, y_j > 0} \frac{G(x,y)}{x^R y^C} \right) = \max_{Z \in \Pi_K(R,C)} \sum_i \sum_j H^{max}_{k_{ij}}(z_{ij}).$$

Hence

Theorem. $$\ln T_K(R,C) \sim \max_{Z \in \Pi_K(R,C)} \sum_i \sum_j H^{max}_{k_{ij}}(z_{ij})$$

in the asymptotic régime of the last theorem.

Detecting correlation of margins, part 1

We say that margins R, C are positively (or negatively) correlated if $T_K(R,C) > I_K(R,C)$ (or $T_K(R,C) < I_K(R,C)$).

We study the case of K constant: $k_{ij} = \kappa$ for all i, j, where $\kappa \in \mathbb{Z}_{>0} \cup \{\infty\}$. Suppose neither R nor C is a constant vector.

Barvinok observed that, as $s \to \infty$, the margins

$$R^{(s)} = \underbrace{(sr_1, \ldots, sr_m, sr_1, \ldots, sr_m, \ldots, sr_1, \ldots, sr_m)}_{s \text{ repetitions}},$$

$$C^{(s)} = \overbrace{(sc_1, \ldots, sc_n, sc_1, \ldots, sc_n, \ldots, sc_1, \ldots, sc_n)}^{s \text{ repetitions}}$$

are asymptotically correlated with positive sign and

$$\frac{T_K(R^{(s)}, C^{(s)})}{I_K(R^{(s)}, C^{(s)})} = e^{\Omega(s^2)} \quad \text{when } \kappa = \infty, \quad \text{but}$$

with negative sign and

$$\frac{I_K(R^{(s)}, C^{(s)})}{T_K(R^{(s)}, C^{(s)})} = e^{\Omega(s^2)} \quad \text{when } \kappa = 1.$$

What does the transition look like?

Detecting correlation of margins, part 2

There are positively correlated margins whenever $\kappa \geq 2$:

Theorem. Let $R \in \mathbb{Z}^m_{>0}$, $C \in \mathbb{Z}^n_{>0}$, and $\kappa \in \{2,3,4,\dots\}$.
Then there exists $\delta = \delta(\kappa) \in (0,1)$, such that if (R,C) satisfy

$$\left(\max_i r_i\right)\left(\max_j c_j\right) < \delta \kappa N,$$

then

$$\lim_{s\to\infty} \frac{1}{s^2} \ln T_\kappa(R^{(s)}, C^{(s)}) \geq \lim_{s\to\infty} \frac{1}{s^2} \ln I_\kappa(R^{(s)}, C^{(s)}),$$

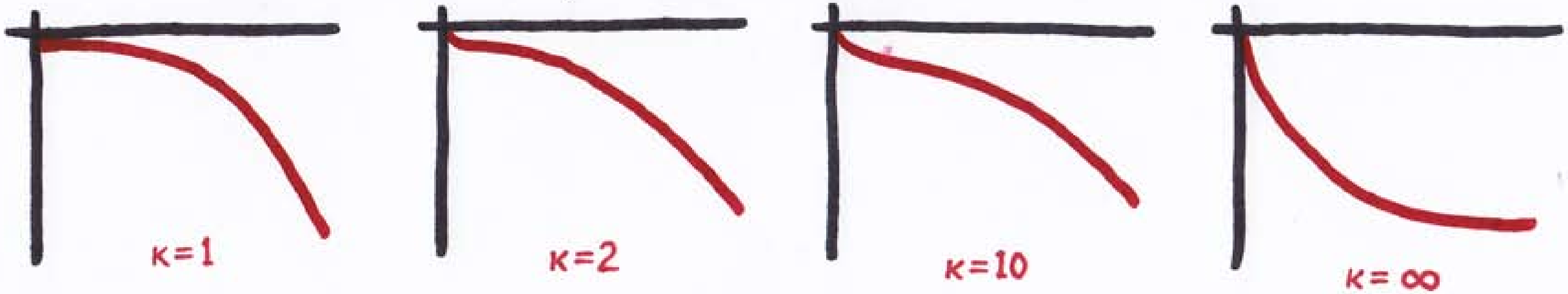
with strict inequality if neither R nor C is a constant vector.

Proof idea:

- Both $T_\kappa(R,C)$, $I_\kappa(R,C)$ can be estimated via entropy formulas.
- Compare four independence models with TG entries satisfying $\mathbf{E}[X_{ij}] = \frac{N}{mn}, \frac{r_i}{n}, \frac{c_j}{m}, \frac{r_i c_j}{N}$. The entropies of the first two differ by more than the entropies of the last two.
  Thus (informally) learning R causes less surprise when we know C.
- Actual proof relies heavily on convexity of $\phi(x) := x^2 (H_\kappa^{\max})''(x)$.

Detecting correlation of margins, part 3

The decisive function $\phi(x) = x^2 (H_\kappa^{max})''(x)$:



Conjecture: Let $R \in \mathbb{Z}_{>0}^m$, $C \in \mathbb{Z}_{>0}^n$, and $\kappa \in \{2,3,4,\dots\}$. Let δ_{cr} be the largest value of δ possible in the previous theorem. Then if (R,C) satisfy

$$\delta_{cr}\,\kappa n < r_i < (1-\delta_{cr})\,\kappa n \quad \text{for } i = 1,\dots,m,$$
$$\delta_{cr}\,\kappa m < c_j < (1-\delta_{cr})\,\kappa m \quad \text{for } j = 1,\dots,n,$$

the conclusion of the previous theorem holds with reversed sign.

Numerical evidence suggests:

$\kappa =$	2	4	6	8	10
$\delta_{cr} \approx$	.05	.11	.14	.15	.16