GAUSSIAN FORMULAS FOR THE NUMBER OF INTEGER POINTS AND VOLUMES OF POLYTOPES

Alexander Barvinok

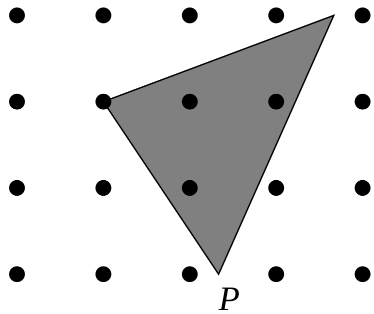
Papers are available at

http://www.math.lsa.umich.edu/∼barvinok/papers.html

Based on a joint work with John Hartigan (Cornell)

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Let $P \subset \mathbb{R}^n$ be a polytope. We want to compute (exactly or approximately) the number $|P \cap \mathbb{Z}^n|$ of integer points in P.



We assume that P is defined by a system of linear equations

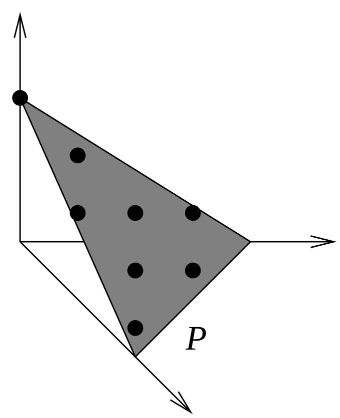
$$Ax = b$$

and inequalities

$$x \geq 0.$$

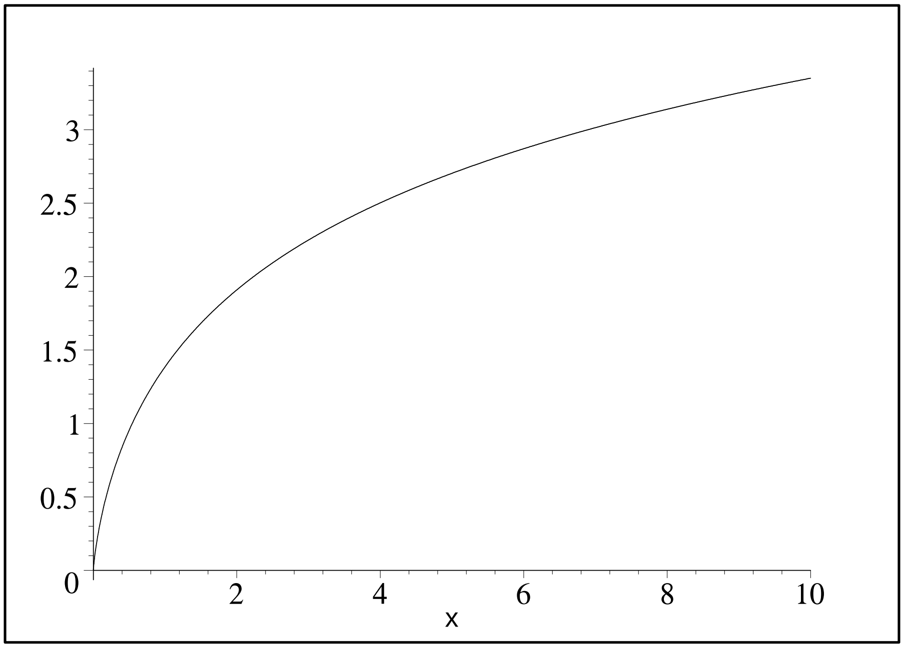
Here A is an integer $d \times n$ matrix of rank $d < n$ and b is an integer n-vector.

So the picture looks more like this



Let us consider a function

$$g(x) = (x+1)\ln(x+1) - x\ln x \quad \text{for} \quad x \geq 0$$



Let us solve the optimization problem:

$$\text{Find} \qquad \max \sum_{j=1}^{n} g\left(x_j\right)$$

$$\text{Subject to:} \quad x = (x_1, \ldots, x_n) \in P.$$

Since g is strictly concave, the maximum point

$$z = (z_1, \ldots, z_n)$$

is unique and can be found efficiently by interior point methods, for example.

Recall that

$$P = \left\{ x \in \mathbb{R}^n : \quad Ax = b \quad \text{and} \quad x \geq 0 \right\},$$

where A is a $d \times n$ integer matrix

integer entries

$$d \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \qquad A$$

n

Let $A = (a_{ij})$.

Let $\Lambda = Ax : \quad x \in \mathbb{Z}^n$ be the lattice in $\mathbb{Z}^d$. Unless $b \in \Lambda$, we have $P \cap \mathbb{Z}^n = \emptyset$.

Let $z = (z_1, \ldots, z_n)$ be the point maximizing

$$g(x) = \sum_{j=1}^{n} \Big((x_j + 1) \ln (x_j + 1) - x_j \ln x_j \Big)$$

for $x = (x_1, \ldots, x_n)$ in P.

Let us compute a $d \times d$ matrix $Q = (q_{ij})$ by

$$q_{ij} = \sum_{k=1}^{n} a_{ik} a_{jk} \left(z_j^2 + z_j \right).$$

We approximate the number of integer points in P by

$$|P \cap \mathbb{Z}^n| \;\approx\; \frac{e^{g(z)} \det \Lambda}{(2\pi)^{d/2} \left(\det Q\right)^{1/2}}.$$

An example

Let us compute the number of 4×4 non-negative integer matrices with row sums 220, 215, 93 and 64 and column sums 108, 286, 71 and 127.

	108	286	71	127
220	*	*	*	*
215	*	*	*	*
93	*	*	*	*
64	*	*	*	*

The number of such matrices is

$$1225914276768514 \approx 1.23 \times 10^{15}.$$

We have a system of 8 equations (row/column sums), they are dependent, however. Let us throw one equation away. The Gaussian formula gives

$$1.3 \times 10^{15}$$

(an overestimate by about 6%).

J. De Loera computed more examples.

A random variable x is *geometric* if

$$\mathbf{Pr}\{x=k\} = pq^k \quad \text{for} \quad k=0,1,\ldots$$

where $p+q=1$ and $p,q>0$. We have

$$\mathbf{E}\,x = \frac{q}{p} \quad \text{and} \quad \mathbf{var}\,x = \frac{q}{p^2}.$$

Conversely,

$$\text{if} \quad \mathbf{E}\,x = z \qquad \text{then} \qquad p=\frac{1}{1+z}, \quad q=\frac{z}{1+z} \quad \text{and} \quad \mathbf{var}\,x = z^2+z.$$

Theorem. *Let $P \subset \mathbb{R}^n$ be a polytope that is the intersection of an affine subspace in $\mathbb{R}^n$ and the non-negative orthant $\mathbb{R}^n_+$. Suppose that P has a non-empty-interior (contains a point with strictly positive coordinates).*

Then the strictly concave function

$$g(x) = \sum_{j=1}^{n} \Big((x_j+1)\ln(x_j+1) - x_j \ln x_j \Big)$$

attains its maximum on P at a unique point $z=(z_1,\ldots,z_n)$ with positive coordinates.

Suppose that $x_1,\ldots,x_n$ are independent geometric random variables with expectations $z_1,\ldots,z_n$ and let $X=(x_1,\ldots,x_n)$. Then the probability mass function of X is constant on $P\cap\mathbb{Z}^n$ and equal to $e^{-g(z)}$ at every $x \in P\cap\mathbb{Z}^n$. In particular,

$$|P\cap\mathbb{Z}^n| = e^{g(z)}\mathbf{Pr}\{X\in P\}.$$

Now, suppose that

$$P = \Big\{ x : \quad Ax = b, \quad x \geq 0 \Big\}.$$

Let $Y = AX$, that is, $Y = (y_1, \ldots, y_d)$, where

$$y_i = \sum_{j=1}^{n} a_{ij} x_j.$$

Theorem implies that

$$|P \cap \mathbb{Z}^n| = e^{g(z)} \mathbf{Pr} \Big\{ Y = b \Big\}.$$

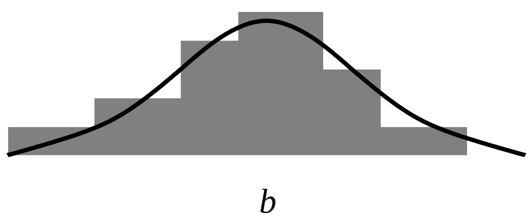
We note that

$$\mathbf{E}\, Y = b$$

and that

$$\mathbf{cov}\,(y_i, y_j) = \sum_{k=1}^{n} a_{ik} a_{jk} \mathbf{var}\, x_k = \sum_{k=1}^{n} a_{ik} a_{jk} \left(z_k^2 + z_k \right).$$

Now, we observe that Y is the sum of n independent random vectors $x_j A_j$, where A_j is the j-th column of A, so we make a leap of faith and assume that Y is close to the Gaussian Y^* with the same expectation b and the covariance matrix Q.



As we estimate

$$\mathbf{Pr}\, \{ Y = b \}$$

by approximating a discrete random variable Y with a Gaussian random variable Y^*, it is crucial (and indeed very helpful) that $\mathbf{E}\, Y = \mathbf{E}\, Y^* = b$, so the Local Central Limit Theorem arguments apply.

More intuition from statistics and statistical physics

In 1957, E.T. Jaynes formulated a general principle. Let Ω be a large but finite probability space with an unknown measure μ, let $f_1, \ldots, f_d : \Omega \longrightarrow \mathbb{R}$ be random variables with known expectations

$$\mathbf{E}\, f_i = \alpha_i \quad \text{for} \quad i = 1, \ldots, d$$

and let $g : \Omega \longrightarrow \mathbb{R}$ be yet another random variable. Then to compute or estimate $\mathbf{E}\, g$ one should assume that μ is the probability measure on Ω of the largest entropy such that that $\mathbf{E}\, f_i = \alpha_i$ for $i = 1, \ldots, d$.

In 1963, I.J. Good argued that the "null hypothesis" concerning an unknown probability distribution from a given class should be the one stating that the distribution is the maximum entropy distribution in the class.

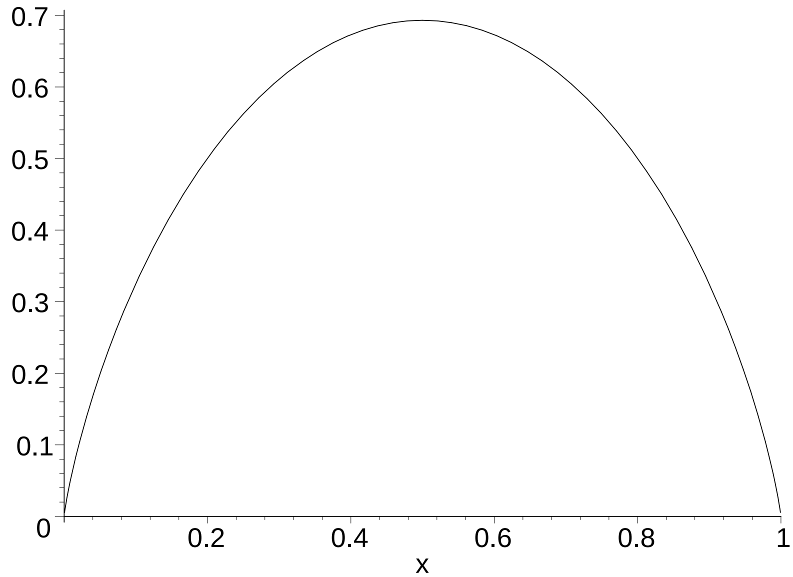
In our case, Ω is the set $\mathbb{Z}_+^n$ of non-negative integer vectors, f_i are the linear equations defining polytope P, and μ is the counting probability measure on $P \cap \mathbb{Z}_+^n$. We approximate μ by the maximum entropy distribution on $\mathbb{Z}_+^n$ subject to the constraints $\mathbf{E}\, f_i = \alpha_i$, where f_i are the linear equations defining P.

Fact: Among all distributions on $\mathbb{Z}_+$ with a given expectation, the geometric distribution has the maximum entropy. The entropy of a geometric distribution with expectation x is

$$g(x) = (x+1)\ln(x+1) - x \ln x.$$

Let us consider a function

$$h(x) = x \ln \frac{1}{x} + (1-x) \ln \frac{1}{1-x} \quad \text{for} \quad 0 \leq x \leq 1$$



Let us solve the optimization problem:

$$\begin{aligned} \text{Find} \qquad & \max \sum_{j=1}^{n} h(x_j) \\ \text{Subject to:} \quad & x = (x_1, \ldots, x_n) \in P \quad \text{and} \\ & 0 \leq x_j \leq 1 \quad \text{for} \quad j = 1, \ldots, n. \end{aligned}$$

Since h is strictly concave, the maximum point

$$z = (z_1, \ldots, z_n)$$

is unique and can be found efficiently by interior point methods, for example.

Recall that

$$P = \Big\{ x \in \mathbb{R}^n : \quad Ax = b \quad \text{and} \quad x \geq 0 \Big\},$$

where A is a $d \times n$ integer matrix

integer entries

$$d \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \qquad A$$

n

Let $A = (a_{ij})$.
Let $\Lambda = Ax : \quad x \in \mathbb{Z}^n$ be the lattice in $\mathbb{Z}^d$.
Let $z = (z_1, \ldots, z_n)$ be the point maximizing

$$h(x) = \sum_{j=1}^{n} \Big(x_j \ln \frac{1}{x_j} + (1 - x_j) \ln \frac{1}{1 - x_j} \Big)$$

for $x = (x_1, \ldots, x_n)$ in P.
Let us compute a $d \times d$ matrix $Q = (q_{ij})$ by

$$q_{ij} = \sum_{k=1}^{n} a_{ik} a_{jk} \left(z_j - z_j^2 \right).$$

We approximate the number of 0-1 points in P by

$$|P \cap \{0, 1\}^n| \; \approx \; \frac{e^{h(z)} \det \Lambda}{(2\pi)^{d/2} \left(\det Q\right)^{1/2}}.$$

A random variable x is *Bernoulli* if

$$\mathbf{Pr}\left\{x=0\right\}=p \quad \text{and} \quad \mathbf{Pr}\left\{x=1\right\}=q$$

where $p+q=1$ and $p,q>0$. We have

$$\mathbf{E}\,x=q \quad \text{and} \quad \mathbf{var}\,x=pq.$$

Conversely,

$$\text{if} \quad \mathbf{E}\,x=z \qquad \text{then} \qquad p=1-z, \quad q=z \quad \text{and} \quad \mathbf{var}\,x=z-z^2.$$

Theorem. *Let $P \subset \mathbb{R}^n$ be a polytope that is the intersection of an affine subspace in $\mathbb{R}^n$ and the unit cube $0 \leq x_j \leq 1$ for $j=1,\ldots,n$. Suppose that P has a non-empty-interior (contains a point with the coordinates strictly between 0 and 1).*

Then the strictly concave function

$$h(x)=\sum_{j=1}^{n}\Big(x_j \ln\frac{1}{x_j}+(1-x_j)\ln\frac{1}{1-x_j}\Big)$$

attains its maximum on P at a unique point $z=(z_1,\ldots,z_n)$ with the coordinates strictly between 0 and 1.

Suppose that $x_1,\ldots,x_n$ are independent Bernoulli random variables with expectations $z_1,\ldots,z_n$ and let $X=(x_1,\ldots,x_n)$. Then the probability mass function of X is constant on $P\cap\{0,1\}^n$ and equal to $e^{-h(z)}$ at every $x\in P\cap\{0,1\}^n$. In particular,

$$|P\cap\{0,1\}^n|=e^{h(z)}\mathbf{Pr}\left\{X\in P\right\}.$$

Now, approximate $Y=AX$ by a Gaussian random variable with the same expectation and the covariance matrix.

For a polytope $P \subset \mathbb{R}^n$ defined by the system $Ax = b$, $x \geq 0$, where A is a $d \times n$ matrix of rank $d < n$, we want to compute $\operatorname{vol} P$ (with respect to the Lebesgue measure in the affine span of P induced from the Euclidean structure in $\mathbb{R}^n$).

Let us consider a function

$$f(x) = n + \sum_{j=1}^{n} \ln x_j$$

and find the point $z \in P$, $z = (z_1, \ldots, z_n)$, maximizing the value of P (point z is called the *analytical center* of P).

Compute $d \times d$ matrix $Q = (q_{ij})$ by

$$q_{ij} = \sum_{k=1}^{n} a_{ik} a_{jk} z_k^2$$

and approximate

$$\operatorname{vol} P \;\approx\; \frac{e^{f(z)} \left(\det AA^T\right)^{1/2}}{(2\pi)^{d/2} \left(\det Q\right)^{1/2}}.$$

A random variable x is exponential with expectation $c > 0$

$$\mathbf{Pr}\left\{x > t\right\} = \begin{cases} e^{-t/c} & \text{if } \; t \geq 0 \\ 1 & \text{if } t < 0. \end{cases}$$

Theorem. *Let $x_1, \ldots, x_n$ be independent exponential random variables, where $\mathbf{E}\, x_j = z_j$, and let $X = (x_1, \ldots, x_n)$. Then the density of X is constant on P and equal to $e^{-f(z)}$ at every point of P.*

The maximum entropy distribution on $\mathbb{R}_+$ with expectation $c > 0$ is the exponential distribution with the entropy $1 + \ln c$.

We consider a random variable $Y = AX$. Hence $\mathbf{E}\, Y = b$ and the covariance matrix of Y is Q. The density of Y at b is

$$e^{-f(z)} \left(\det AA^T\right)^{-1/2} \operatorname{vol} P.$$

We approximate Y by the Gaussian random variable Y^* with expectation b and the covariance matrix Q.

Examples: multi-index transportation polytopes and multi-way contingency tables

Transportation polytopes. Let us choose positive $r_1, \ldots, r_m$ and $c_1, \ldots, c_n$ such that

$$r_1 + \ldots + r_m = c_1 + \ldots + c_n = N.$$

The polytope P of non-negative $m \times n$ matrices (x_{ij}) with row sums $r_1, \ldots, r_m$ and column sums $c_1, \ldots, c_n$ is called a (two-index) *transportation polytope*. We have

$$\dim P = (m-1)(n-1).$$

Suppose r_i and c_j are integer. Integer points in P are called (two-way) *contingency tables* while 0-1 points in P are called (two-way) *binary contingency tables* with *margins* $r_1, \ldots, r_m$ and $c_1, \ldots, c_n$.

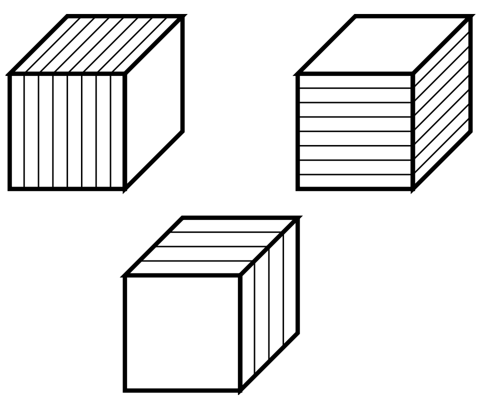
ν-index transportation polytopes. Let us fix an integer $\nu \geq 2$. The polytope P of ν-dimensional

$$k_1 \times \ldots \times k_\nu$$

arrays $(x_{j_1 \ldots j_\nu})$ with prescribed sectional sums

$$\sum_{\substack{1 \leq j_1 \leq k_1 \\ \cdots\cdots\cdots \\ 1 \leq j_{i-1} \leq k_{i-1} \\ 1 \leq j_{i+1} \leq k_{i+1} \\ \cdots\cdots\cdots \\ 1 \leq j_\nu \leq k_\nu}} x_{j_1 \ldots j_{i-1}, j, j_{i+1} \ldots j_\nu}$$

are called ν-way transportation polytopes.



As long as the natural balance conditions are met, P is a polytope with

$$\dim P = k_1 \cdots k_\nu - (k_1 + \ldots + k_\nu) + \nu - 1.$$

Integer points in P are called ν-way contingency tables and 0-1 points in P are called ν-way binary contingency tables.

Some results

Fix ν and let $k_1, \dots, k_\nu$ grow roughly proportionately.
We can prove:

- The volume of P is asymptotically Gaussian provided $\nu \geq 5$. We suspect it is Gaussian already for $\nu \geq 3$;

- The number of integer points and the number of 0-1 points in P are asymptotically Gaussian provided $\nu \geq 6$. We suspect it is Gaussian already for $\nu \geq 3$.

- In particular, for the volume of the (dilated) polytope P_k of *polystochastic tensors*, that is, the polytope of $k \times \cdots \times k$ arrays with all sectional sums equal to $k^{\nu-1}$ is

$$\operatorname{vol} P_k = \Big(1 + o(1)\Big) \frac{e^{k^\nu}}{(2\pi)^{(\nu k - \nu + 1)/2}} \quad \text{as} \quad k \longrightarrow +\infty,$$

provided $\nu \geq 5$.

- In particular, the number of non-negative integer $k \times \cdots \times k$ *magic cubes*, that is, contingency tables with all sectional sums equal to $r = \alpha k^{\nu-1}$ is

$$\Big(1 + o(1)\Big) \left((\alpha+1)^{\alpha+1} \alpha^{-\alpha}\right)^{k^\nu} \left(2\pi\alpha^2 + 2\pi\alpha\right)^{-(k\nu - \nu + 1)/2} k^{(\nu - \nu^2)(k-1)/2}$$
$$\text{as} \quad k \longrightarrow +\infty,$$

provided $\nu \geq 6$, r is integer and α is separated away from 0;

- In particular, the number of $k \times \cdots \times k$ *regular ν-partite hypergraphs*, that is binary contingency tables with all sectional sums equal to $r = \alpha k^{\nu-1}$ is

$$\Big(1 + o(1)\Big) \left(1 - \alpha)^{1-\alpha} \alpha^{\alpha}\right)^{-k^\nu} \left(2\pi\alpha - 2\pi\alpha^2\right)^{-(k\nu - \nu + 1)/2} k^{(\nu - \nu^2)(k-1)/2}$$
$$\text{as} \quad k \longrightarrow +\infty,$$

provided $\nu \geq 6$, r is integer and α is separated away from 0 and 1.

An interesting case of $\nu = 2$

The asymptotic of the volume of the polytope of doubly stochastic matrices (the *Birkhoff polytope*) was computed recently by E.R. Canfield and B. McKay. It is *not* asymptotically Gaussian, it differs from the Gaussian estimate by an additional factor of $e^{1/3}$. This corresponds to the 4-th order *kurtosis* correction to the Gaussian distribution.

The asymptotic of the number of integer contingency tables with equal row sums and equal column sums was recently computed by E.R. Canfield and B. McKay. It is differs from the Gaussian estimate by a constant factor (generally, greater than 1).

The asymptotic of the number of binary contingency tables with equal row sums and equal column sums was recently computed by E.R. Canfield, C. Greenhill and B. McKay. It differs from the Gaussian estimate by a constant factor (generally, smaller than 1).

If not clear whether the Gaussian approximation falls within a constant factor from the true asymptotic if the margins are allowed to be different (within reason). For general margins, the main terms in logarithmic order are $e^{g(z)}$ (integer tables) and $e^{h(z)}$ (binary tables) respectively.