The asymptotic volume of the Birkhoff polytope

E. Rodney Canfield*  Brendan D. McKay†
Department of Computer Science  Department of Computer Science
University of Georgia  Australian National University
Athens, GA 30602, USA  Canberra ACT 0200, Australia
ercanfie@uga.edu  bdm@cs.anu.edu.au

Abstract

Let \(m, n \geq 1\) be integers. Define \(T_{m,n}\) to be the transportation polytope consisting of the \(m \times n\) non-negative real matrices whose rows each sum to 1 and whose columns each sum to \(m/n\). The special case \(B_n = T_{n,n}\) is the much-studied Birkhoff-von Neumann polytope of doubly-stochastic matrices. Using a recent asymptotic enumeration of non-negative integer matrices (Canfield and McKay, 2007), we determine the asymptotic volume of \(T_{m,n}\) as \(n \to \infty\) with \(m = m(n)\) such that \(m/n\) neither decreases nor increases too quickly. In particular, we give an asymptotic formula for the volume of \(B_n\).

1 Introduction

Let \(m, n \geq 1\) be integers. Define \(T_{m,n}\) to be the transportation polytope consisting of the \(m \times n\) non-negative real matrices whose rows each sum to 1 and whose columns each sum to \(m/n\). The special case \(B_n = T_{n,n}\) is the famous Birkhoff-von Neumann polytope of doubly-stochastic matrices.

It is well known (see Stanley [7, Chap. 4] for basic theory and references) that \(T_{m,n}\) spans an \((m-1)(n-1)\)-dimensional affine subspace of \(\mathbb{R}^{m \times n}\). The vertices of \(T_{m,n}\) were described by Klee and Witzgall [6] and are moderately complicated. The special case of \(B_n\) is however very simple: the vertices are precisely the \(n \times n\) permutation matrices.

Two types of volume are customarily defined for such polytopes. We can illustrate the difference using the example

\[B_2 = \left\{ \begin{pmatrix} z & 1-z \\ 1-z & z \end{pmatrix} \mid 0 \leq z \leq 1 \right\} = \left[ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right].\]

*Research supported by the NSA Mathematical Sciences Program.
†Research supported by the Australian Research Council.
where the last notation indicates a closed line-segment in \( \mathbb{R}^{2 \times 2} \). The length of this line-segment is the \textit{volume} \( \text{vol}(B_2) = 2 \). We can also consider the lattice induced by \( \mathbb{Z}^{2 \times 2} \) on the affine span of \( B_2 \): this consists of the points \( \left( \begin{array}{cc} 1 & -z \\ z & 1 \end{array} \right) \) for integer \( z \). The polytope \( B_2 \) consists of a single basic cell of this lattice, so it has \textit{relative volume} \( \nu(B_2) = 1 \). In general, \( \text{vol}(T_{m,n}) \) is the volume in units of the ordinary \((m-1)(n-1)\)-dimensional Lebesgue measure, while \( \nu(T_{m,n}) \) is the volume in units of basic cells of the lattice induced by \( \mathbb{Z}^{m \times n} \) on the affine span of \( T_{m,n} \).

\begin{lemma}
For \( m, n \geq 2 \), \( \text{vol}(T_{m,n}) = m^{(n-1)/2} n^{(m-1)/2} \nu(T_{m,n}) \).
\end{lemma}

\textbf{Proof.} This is established in [5, Theorem 3]. Also see the Appendix of [2]. \hfill \Box

Next, define the function \( H_{m,n} : \mathbb{Z} \to \mathbb{Z} \) by
\[
H_{m,n}(z) = |zT_{m,n} \cap \mathbb{Z}^{m \times n}|.
\]

Clearly \( zT_{m,n} \cap \mathbb{Z}^{m \times n} \) is the set of \( m \times n \) non-negative integer matrices with row sums equal to \( z \) and column sums equal to \( zm/n \). This set is non-empty when \( zm/n \in \mathbb{Z} \); that is, when \( z \) is a multiple of \( z_0 = n/\gcd(m,n) \). The base case \( z = z_0 \) corresponds to an expanded polytope \( z_0 T_{m,n} \) whose vertices are integral [6, Cor. 1]. Therefore, by the celebrated theorem of Ehrhart (see [7]), there are constants \( c_i(m,n) \) for \( i = 0, 1, \ldots, (m-1)(n-1) \) such that
\[
H_{m,n}(z) = \left\{ \begin{array}{ll}
\sum_{i=0}^{(m-1)(n-1)} c_i(m,n) z^{(m-1)(n-1)-i}, & \text{if } z_0 \text{ divides } z; \\
0, & \text{otherwise.}
\end{array} \right.
\] (1)

This is the \textit{Ehrhart pseudo-polynomial} of \( T_{m,n} \). Applying [7, Prop. 4.6.30] to \( z_0 T_{m,n} \), we find that
\[
\nu(T_{m,n}) = c_0(m,n).
\] (2)

We turn now to asymptotics. Our main tool will be the following theorem of the present authors [3].

\begin{theorem}
Suppose \( m = m(n) \), \( s = s(n) \) and \( t = t(n) \) are positive integer functions such that \( ms = nt \). Let \( M(m,s;n,t) \) be the number of \( m \times n \) non-negative integer matrices with row sums equal to \( s \) and column sums equal to \( t \). Define \( \lambda = \lambda(n) \) by \( ms = nt = \lambda mn \). Let \( a, b > 0 \) be constants such that \( a + b < \frac{1}{2} \). Suppose that \( n \to \infty \) and that, for large \( n \),
\[
\frac{(1 + 2\lambda)^2}{4\lambda(1 + \lambda)} \left( 1 + \frac{5m}{6n} + \frac{5n}{6m} \right) \leq a \log n.
\] (3)
\end{theorem}
Then

\[ M(m, s; n, t) = \frac{(n+s-1)^m}{n-1} \frac{(m+t-1)^n}{m-1} \frac{\exp\left(\frac{1}{2} + O(n^{-b})\right)}{(mn + \lambda mn - 1)_{mn - 1}}. \]

Using this result, we can prove the following theorem concerning the volumes of \( T_{m,n} \) and \( B_n \).

**Theorem 2.** Let \( a, b > 0 \) be constants such that \( a + b < \frac{1}{2} \). Then

\[ \text{vol}(T_{m,n}) = \frac{1}{(2\pi)^{(m+n-1)/2} n^{(m-1)(n-1)}} \exp\left(\frac{1}{3} + mn - \frac{(m-n)^2}{12mn} + O(n^{-b})\right) \]

when \( m, n \to \infty \) in such a way that \( \max\left(\frac{m}{n}, \frac{n}{m}\right) \leq \frac{6}{5} a \log n \). In particular, for any \( \epsilon > 0 \)

\[ \text{vol}(B_n) = \frac{1}{(2\pi)^{n-1/2} n^{(n-1)^2}} \exp\left(\frac{1}{3} + n^2 + O(n^{-1/2+\epsilon})\right) \]

as \( n \to \infty \).

**Proof.** From (1) and (2), we have

\[ \nu(T_{m,n}) = \lim_{z \to \infty} \frac{H_{m,n}(z)}{z^{(m-1)(n-1)}} = \lim_{\lambda \to \infty} \frac{M(m, \lambda n; n, \lambda m)}{(\lambda n)^{(m-1)(n-1)}}, \tag{4} \]

where we restrict \( z \) to multiples of \( z_0 \) and \( \lambda \) to multiples of \( z_0 / n \). If \( a' > a \) and \( a' + b < \frac{1}{2} \), then the left side of (3) is less than \( a' \log n \) for sufficiently large \( \lambda \). Thus the conditions for Theorem \([1]\) hold. It remains to apply that theorem to (4) using Stirling’s formula, and to infer the value of \( \text{vol}(T_{m,n}) \) using Lemma \([1]\) \( \square \)

It is of interest to note that the same asymptotic formula for the volume (except for the error term) follows from the estimate of \( M(m, s; n, t) \) that Diaconis and Efron proposed without proof in 1985 \([3]\).

Exact values of \( \text{vol}(B_n) \) are known up to \( n = 10 \) \([1]\). In Table \([1]\) we compare the exact values to the approximation given in Theorem \([2]\). It appears that the true magnitude of the error term might be \( O(n^{-1}) \). This would indeed be the case if the well-tested conjecture made in \([3]\) about the value of \( M(n, s; n, t) \) was true. The same conjecture implies a value of \( \text{vol}(T_{m,n}) \) with relative error \( O((m + n)^{-1}) \) for all \( m, n \).

Recently, a summation with \( O(n^n n!) \) terms was found for \( \text{vol}(B_n) \) \([4]\). Whether it is useful for asymptotics remains to be seen.
| $n$ | estimate/actual |
|-----|-----------------|
| 1   | 1.51345         |
| 2   | 1.20951         |
| 3   | 1.25408         |
| 4   | 1.22556         |
| 5   | 1.19608         |
| 6   | 1.17258         |
| 7   | 1.15403         |
| 8   | 1.13910         |
| 9   | 1.12684         |
| 10  | 1.11627         |

Table 1: Accuracy of Theorem 2 for $\text{vol}(\mathcal{B}_n)$.

References

[1] M. Beck and D. Pixton, The Ehrhart polynomial of the Birkhoff polytope, *Discrete Comput. Geom.*, 30 (2003) 623–637.

[2] C.S. Chan and D.P. Robbins, On the volume of the polytope of doubly stochastic matrices, *Experiment. Math.*, 8 (1999), 291–300.

[3] E.R. Canfield and B.D. McKay, Asymptotic enumeration of contingency tables with constant margins, submitted (2007). Preprint available at http://www.arxiv.org/abs/math.CO/0703600.

[4] J.A. DeLoera, F. Liu and R. Yoshida, Formulas for the volumes of the polytope of doubly-stochastic matrices and its faces (2007), preprint available at http://www.arxiv.org/abs/math.CO/0701866.

[5] P. Diaconis and B. Efron, Testing for independence in a two-way table: New interpretations of the chi-square statistic, *Ann. Stat.*, 13 (1985) 845–874.

[6] V. Klee and C. Witzgall, Facets and vertices of transportation polytopes, Lectures in Applied Mathematics, 11, Amer. Math. Soc., Providence, RI, 1968, 257–282.

[7] R.P. Stanley, *Enumerative Combinatorics I*, Wadsworth, Belmont, 1986.