An algorithmic proof of Brégman-Minc theorem

Li Xu *, Heng Liang †, Fengshan Bai ‡

* Department of Mathematical Sciences, Tsinghua University, Beijing, P.R.China

First published on: 18 March 2009

To cite this Article Xu, Li, Liang, Heng and Bai, Fengshan(2009) 'An algorithmic proof of Brégman-Minc theorem', International Journal of Computer Mathematics, 86: 12, 2181 — 2185, First published on: 18 March 2009 (iFirst)

To link to this Article: DOI: 10.1080/00207160802441264

URL: http://dx.doi.org/10.1080/00207160802441264
An algorithmic proof of Brégman–Minc theorem

Li Xu, Heng Liang* and Fengshan Bai

Department of Mathematical Sciences, Tsinghua University, Beijing, P.R.China

(Received 16 August 2008; revised version received 27 August 2008; accepted 27 August 2008)

Brégman-Minc theorem gives the best known upper bound of the permanent of (0, 1)-matrices. A new proof of the theorem is presented in this paper, using an unbiased estimator of permanent [L.E. Rasmussen, Approximating the permanent: a simple approach, Random Structures Algorithms 5 (1994), p. 349]. This proof establishes a connection between the randomized approximate algorithm and the bound estimation for permanents.

Keywords: permanent; (0, 1)-matrix; upper bound; algorithm

2000 AMS Subject Classification: 15A15; 15A45

1. Introduction

Let \( A = [a_{ij}] \) be an \( n \times n \) matrix with all entries either 0 or 1, which is called a (0, 1)-matrix for brevity. The permanent of matrix \( A \) is defined as

\[
\text{per}(A) = \sum_{\sigma} \prod_{i=1}^{n} a_{i,\sigma(i)},
\]

where the sum goes over every permutation \( \sigma \) of the set \{1, 2, \ldots, n\}. The permanent looks similar to the determinant of matrices. However, it is much harder to compute. Valiant [17] proved that evaluating the permanent of a (0, 1)-matrix is a \#P-complete problem. Hence randomized approximate algorithms, which can give a reasonable estimation for \( \text{per}(A) \) within acceptable computer time, and bound estimations now attract a great deal of attention [4,8,9].

The Brégman–Minc theorem gives the best known upper bound for the permanent of (0, 1)-matrices, as follows.

**Theorem 1.1** Let \( A = [a_{ij}] \) be a (0, 1)-matrix with row sums \( r_1, r_2, \ldots, r_n \). Then

\[
\text{per}(A) \leq \prod_{i=1}^{n} (r_i!)^{1/r_i}.
\]
This result was conjectured by Minc [12], and first proved by Brégman [3]. A shorter proof was then given by Schrijver [16]. Later, a similar proof was obtained by Alon and Spencer [1] by analysing a randomized procedure for estimating the permanent; and a proof based on entropy was presented by Radhakrishnan [14].

The Rasmussen’s estimator (RAS) [15] is an efficient and very simple unbiased estimator for the permanent of $(0, 1)$-matrices. It gives a randomized algorithm for the permanents of matrices. Based on this algorithm, a new proof of the Brégman–Minc theorem is obtained in this paper. This proof actually shows that the tightness of Brégman–Minc theorem is related to the efficiency of the RAS. Our result establishes the connection between the randomized approximate algorithm and the bound estimate for the permanent of a $(0, 1)$-matrix. Moreover, it is shown that the Brégman–Minc upper bound is a good estimate of permanents for almost all $(0, 1)$-matrices as the matrix dimension $n$ becomes large.

In the next section, RAS is briefly introduced. A new proof of the Brégman–Minc theorem is presented based on the algorithm in Section 2, which is the main result of this article. Some discussions are finally made.

2. The Rasmussen’s estimator

Let $A[i,j]$ denote the sub-matrix obtained by removing the $i$th row and $j$th column from the matrix $A$; and $A(i,:)$ be the $i$th row of matrix $A$. For any set $S$, let $|S|$ be the number of its elements. Algorithm 2.1 gives RAS for permanent [15].

**Algorithm 2.1 Rasmussen’s estimator (RAS)**

**Input:** $A$: an $n \times n$ $(0, 1)$-matrix, $X = 1$.

**Output:** $X_A$: the estimate for $\text{perm}(A)$.

**Step 1:** for $i = 1$ to $n$

- $W = \{s \mid a_{1s} = 1\}$;
- if $W = \emptyset$ then $X = 0$;
- elseif $W \neq \emptyset$ then choose $k$ from $W$ with probability $p_i = \frac{1}{|W|}$;
- $X = X \cdot |W|$, $A = A[1][k]$;

**Step 2:** $X_A = X$.

Running Algorithm 2.1 once gives one sample, which would end up with either a permutation $\sigma$ of $\{1, 2, \ldots, n\}$ such that $a_{i,\sigma(i)} = 1$ for all $i = 1, 2, \ldots, n$, or a permutation $\sigma'$ of a subset of $\{1, 2, \ldots, n\}$ such that $[\sigma'(1), \ldots, \sigma'(j)] \ (j < n)$. We call the permutation obtained in this way a ‘random path’. $X_A$ obtained from Algorithm 2.1 is called the value of the random path, which defines a random variable.

**Definition** A random path $\sigma$ is said to be feasible if $X_A(\sigma) \neq 0$.

Note that permutations $\sigma$ that satisfy $\prod_{i=1}^{n} a_{i,\sigma(i)} = 1$ are one-to-one correspondents to feasible paths. For a feasible path $\sigma$ with path value $X_A(\sigma)$, the probability of the path being sampled in Algorithm 2.1 is $1/X_A(\sigma)$. The following theorem shows that the RAS estimator is unbiased.

**Theorem 2.2 (see [10,15])** Let $X_A$ be the random variable given by Algorithm 2.1. Then $E[X_A] = \text{per}(A)$. 

Downloaded By: [Tsinghua University] At: 07:36 23 December 2009
The efficiency of the algorithm relies on the critical ratio $E[X_A^2]/E[X_A]^2$. This ratio is determined by the distribution of the path values.

**Theorem 2.3** Let $S$ denote the set of all feasible paths of matrix $A$. Then for the RAS estimator $X_A$, one has

$$E[X_A]^2 = \text{per}^2(A), \quad E[X_A^2] = \sum_{\sigma \in S} X_A(\sigma).$$

3. The main results

Let $A = [a_{ij}]$ be an $n \times n$ $(0, 1)$-matrix with row sums $r_1, r_2, \ldots, r_n$ and $A_k$ ($k = 1, 2, \ldots, n!$) be matrices induced by reordering the rows of $A$. Each possible permutation of $\{1, 2, \ldots, n\}$ gives a matrix $A_k$. Hence $n!$ is the number of such matrices. Let $S$ be the set of all feasible paths of matrix $A$, which are the permutations $\sigma$ of $\{1, 2, \ldots, n\}$ such that $a_i, \sigma(i) = 1$ for all $1 \leq i \leq n$. Denote $S = \{\sigma_1, \sigma_2, \ldots, \sigma_N\}$. Thus

$$\text{per}(A) = |S| = N.$$ 

Assume that a feasible path $\sigma_j \in S$ of $A$ corresponds to the path $\sigma_j^{(k)}$ with matrix $A_k$ for $1 \leq k \leq n!$. Let $P_k (1 \leq k \leq n!)$ be the probability of $X_{A_k}$ nonzero. Then

$$P_k = \sum_{j=1}^N \frac{1}{X_{A_k}(\sigma_j^{(k)})} \leq 1, \quad k = 1, 2, \ldots, n!.$$ 

**Theorem 3.1** Let $C_k$ denote the critical ratio of $X_{A_k}$, i.e., $C_k = E[X_{A_k}^2]/E[X_{A_k}]^2$. Then

$$\frac{1}{\prod_{k=1}^{n!} C_k^{1/n!}} \prod_{i=1}^n (r_i!)^{1/r_i} \leq \text{per}(A) \leq \left[ \prod_{k=1}^N P_k \right]^{1/n!} \prod_{i=1}^n (r_i!)^{1/r_i}.$$ 

**Proof** For any feasible path $\sigma_j$ of $A$, it corresponds to the path $\sigma_j^{(k)}$ with matrix $A_k$ for $1 \leq k \leq n!$. One has

$$\prod_{k=1}^{n!} X_{A_k}(\sigma_j^{(k)}) = \prod_{i=1}^n \prod_{t=1}^{r_i} t^{n!/r_i} = \left[ \prod_{i=1}^n (r_i!)^{1/r_i} \right]^{n!}.$$ 

(1)

Consider all $A$’s corresponding paths with matrix $A_k$ ($1 \leq k \leq n!$),

$$\prod_{j=1}^N \frac{1}{X_{A_k}(\sigma_j^{(k)})} \leq \left[ \sum_{j=1}^N 1/ X_{A_k}(\sigma_j^{(k)}) \right]^N = \left( \frac{P_k}{N} \right)^N.$$ 

(2)

Hence it is obvious that

$$\left( \frac{N}{P_k} \right)^N \leq \prod_{j=1}^N X_{A_k}(\sigma_j^{(k)}).$$ 

(3)

From Equations (1) and (3)

$$\prod_{k=1}^{n!} \left( \frac{N}{P_k} \right)^N \leq \prod_{k=1}^{n!} \prod_{j=1}^N X_{A_k}(\sigma_j^{(k)}) = \prod_{j=1}^N \left[ \prod_{i=1}^n (r_i!)^{1/r_i} \right]^{n!}.$$ 

(4)
Using Equation (1) and Theorem 2.2

\[
\prod_{j=1}^{N} \prod_{i=1}^{n} (r_i!)^{1/r_i} = \prod_{k=1}^{n} X_{A_k}(\sigma_j^{(k)}) \leq \prod_{k=1}^{n!} \left[ \frac{\sum_{j=1}^{N} X_{A_k}(\sigma_j^{(k)})}{N} \right]^{N} = \left[ \prod_{k=1}^{n!} \left( \frac{E[X_{A_k}^2]}{E[X_{A_k}]} \right) \right]^{N} N^{N} n!. \tag{5}
\]

By Combining Equations (4) and (5), we have that

\[
\frac{1}{\prod_{k=1}^{n!} C_k^{1/n!}} \prod_{i=1}^{n} (r_i!)^{1/r_i} \leq \text{per}(A) \leq \prod_{k=1}^{n!} P_k^{1/n!} \prod_{i=1}^{n} (r_i!)^{1/r_i}. \tag{6}
\]

Thus the result follows.

It is clear that the results of Theorem 3.1 certainly imply Theorem 1.1, Brégman–Minc Theorem, since all \(P_k\)'s in Equation (6) are no more than one. A lower bound is also given in Theorem 3.1. The theorem shows that the tightness of Brégman–Minc bound is related to the efficiency of the RAS, which establishes the connection between the randomized approximate algorithm and the bound estimation for the permanent of a \((0, 1)\)-matrix.

4. Discussions

The asymptotic property of RAS is known as follows:

**Theorem 4.1 (Rasmussen [15])** Let \(\omega = \omega(n)\) be any function satisfying \(\omega \to \infty\) as \(n \to \infty\). Then

\[
E[X_A^2]/E[X_A]^2 \leq O(n\omega)
\]

holds for almost all \((0, 1)\)-matrices \(A\), where \(X_A\) is the output of the estimator RAS of Algorithm 2.1.

This result implies that the permanent of a random \((0, 1)\)-matrix is tightly concentrated for large \(n\). Now by Theorem 3.1, the point at which the permanents of almost all \((0, 1)\)-matrices \(A\) are concentrated is the Brégman–Minc upper bound in the sense that

\[
n^{-1} \log \frac{\prod_{i=1}^{n} (r_i!)^{1/r_i}}{\text{per} A} \to 0 \quad \text{as} \quad n \to \infty.
\]

Hence the Brégman–Minc upper bound is a good estimation of permanents for almost all \((0, 1)\)-matrices as \(n\) becomes large.

Friedland et al. [5] have recently given an interesting and permanent result on the concentration property. Let \(A_n = [a_{ij}]\) be an \(n \times n\) matrix with entries in the closed interval \([a, b]\), \(0 < a \leq b\), and let \(X_n = (\sqrt{a_{ij}}x_{ij})\) be an \(n \times n\) random matrix, where \(\{x_{ij}\}\) are independent identically
distributed \( N(0, 1) \) random variables. It is shown that for large \( n \), the determinant of the matrix \( X_n^T X_n \) concentrates sharply at the permanent of \( A_n \), in the sense that

\[
\lim_{n \to \infty} \frac{1}{n} \log \frac{\det(X_n^T X_n)}{\per(A_n)} = 0
\]

in probability.

There are wide applications of the permanent of \((0, 1)\)-matrices. The permanent of the adjacency matrix of a chemical graph is expected to be computed in computational molecular chemistry [6]. In statistical physics, the permanent is used to solve the dimer covering problem [2] and monomer–dimer model [7]. The permanent is also an important issue in combinatorial counting and graph theory [4,13].

The bound estimates of permanents are useful tool in computations. For example, based on van der Waerden lower bound and Brégman–Minc upper bound, Linial et al. [11] develop a deterministic strongly polynomial algorithm to approximate the permanent. The connection between the randomized approximate algorithm and the bound estimation, which is established in Theorem 3.1, should be useful to construct more efficient algorithms for the permanent of \((0, 1)\)-matrices arising from application problems.

**Acknowledgements**

This research is supported by National Science Foundation of China 10501030.

**References**