# 应用统计

## 第7讲　重要度抽样　importance sampling

# 计算积分的Monte Carlo方法

$$\iint_D f(x)\,dx$$

生成$D$内的均匀随机数$x_1, x_2, \cdots x_n$,

$$\iint_D f(x)\,dx \approx S_D \cdot \frac{f(x_1)+f(x_2)+\cdots+f(x_n)}{n}$$

$$X \sim U(D), \quad E(f(X)) = \iint_D f(x)\cdot\frac{1}{S_D}dx, \text{ 其中 } S_D = \iint_D dx$$

$$E(f(X)) \approx \frac{1}{n}\sum_{k=1}^{n} f(x_k) \quad \Rightarrow \quad \iint_D f(x)\,dx \approx S_D \cdot \frac{1}{n}\sum_{k=1}^{n} f(x_k)$$

理论上g取任意概率密度函数都可以，但是选取的时候　想要让方差尽量小，还要便于计算

# 重要度抽样 importance sampling

$$\iint\limits_D f(x)\,dx$$

更一般地，生成$D$内的随机数$x_1, x_2, \cdots x_n$，

满足概率密度$g(x)$

$$\iint\limits_D f(x)\,dx \approx \frac{1}{n}\cdot\left(\frac{f(x_1)}{g(x_1)}+\frac{f(x_2)}{g(x_2)}+\cdots+\frac{f(x_n)}{g(x_n)}\right)$$

$$X \sim g(x),\quad E\left(\frac{f(X)}{g(X)}\right)=\iint\limits_D \frac{f(x)}{g(x)}\cdot g(x)\,dx=\iint\limits_D f(x)\,dx,$$

$$E\left(\frac{f(X)}{g(X)}\right)\approx\frac{1}{n}\sum_{k=1}^{n}\frac{f(x_k)}{g(x_k)},$$

只要$Var_g\left(\frac{f(X)}{g(X)}\right)<Var_u\left(\frac{f(X)}{1/S_D}\right)$

无偏的估计

$$Var\left(\frac{f(X)}{g(X)}\right)=E\left(\left(\frac{f(X)}{g(X)}\right)^2\right)-E\left(\frac{f(X)}{g(X)}\right)^2=\iint_D \frac{f^2(x)}{g^2(x)}\cdot g(x)\,dx-\left(\iint_D f(x)\,dx\right)^2$$

$$\iint_D \frac{f^2(x)}{g(x)}dx=\iint_D g(x)\,dx\cdot\iint_D \frac{f^2(x)}{g(x)}dx \ge \left(\iint_D \sqrt{g(x)}\cdot\frac{f(x)}{\sqrt{g(x)}}dx\right)^2=\left(\iint_D f(x)\,dx\right)^2$$

当$\sqrt{g(x)}=c\cdot\frac{f(x)}{\sqrt{g(x)}}$，上面不等式取等号，即$g(x)=c\cdot f(x)\Rightarrow g(x)=\frac{f(x)}{\iint_D f(x)dx}$

理想的最好的估计，方差为0了

$$E\left(\frac{f(X)}{g(X)}\right)\approx\frac{1}{n}\sum_{k=1}^{n}\frac{f(x_k)}{g(x_k)}\quad (X\sim g(x))$$，使得 $g(x)\approx\frac{f(x)}{\iint_D f(x)dx}$

显然g的最好理想估计没法达到，但是可以让其分布规律类似，f大的地方让g也比较大，这样方差就比较小了

# 重要度抽样的例子

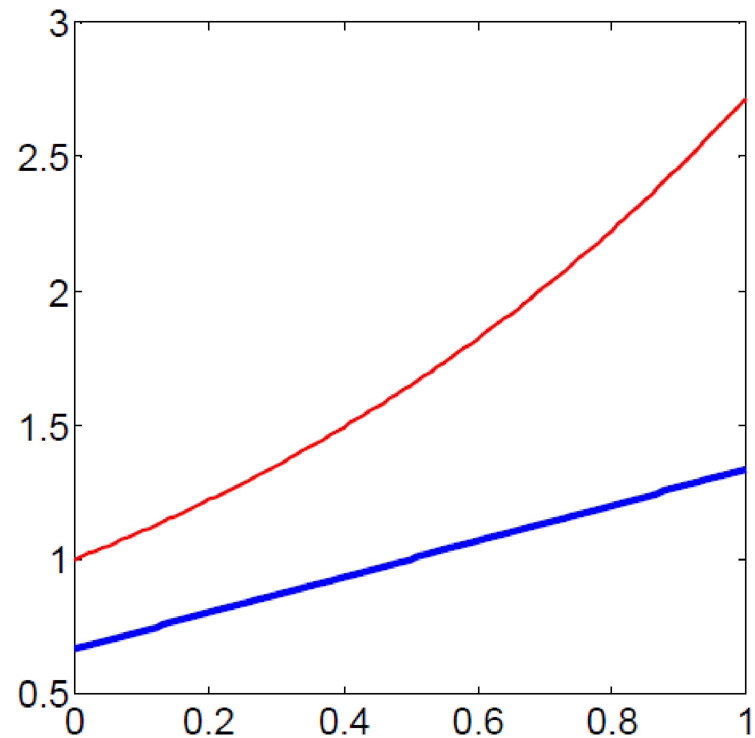
$$\theta = E\left(e^Y\right) = \int_0^1 e^x dx \quad Y \sim U[0,1]$$

$$Var\left(e^Y\right) = E\left(e^{2Y}\right) - E\left(e^Y\right)^2 = 0.242$$

$$X \sim \frac{2}{3}(1+x) \quad x \in [0,1]$$

$$\theta = E\left(\frac{e^X}{\frac{2}{3}(1+x)}\right) = \int_0^1 e^x dx$$

$$Var\left(\frac{e^X}{\frac{2}{3}(1+X)}\right) = \frac{3}{2}\int_0^1 \frac{e^{2x}}{(1+x)} dx - (e-1)^2 = 0.0269$$



# 重要度抽样的例子

若 $X=(X_1,X_2,\cdots,X_{100})$ 是 $(1,2,\cdots,100)$ 的一个随机排列，也就是说 $X$ 等可能地等于 $100!$ 个排列中的任意一个。估计概率

$$\theta=P\left\{\sum_{j=1}^{100} j\cdot X_j>290000\right\}。$$

$$i_1<i_2, j_1<j_2, i_1j_1+i_2j_2-(i_1j_2+i_2j_1)=(i_2-i_1)(j_2-j_1)>0\Rightarrow i_1j_1+i_2j_2>i_1j_2+i_2j_1$$

[171700, 338350]

首先对$\theta$值做一大致的估计

$$E\left(\sum_{j=1}^{100} j\cdot X_j\right)=\sum_{j=1}^{100} j\cdot E\left(X_j\right)=\frac{100\times 101}{2}\times\frac{100\times 101}{2\times 100}=255025$$

$$Std\left(\sum_{j=1}^{100} j\cdot X_j\right)=\left(\frac{99\times 100^2\times 101^2}{144}\right)^{1/2}=8374.5$$

$$\sum_{j=1}^{100} j\cdot X_j \dot\sim N\left(255025, 8374.5^2\right)$$

概率太小了，说明如果利用Monte Carlo方法的话要特别大的样本量才能保证一定的可靠性

$$P\left\{\sum_{j=1}^{100} j\cdot X_j>290000\right\}=P\left\{\frac{\sum_{j=1}^{100} j\cdot X_j-255025}{8374.5}>\frac{290000-255025}{8374.5}\right\}$$

$$\approx 1-\Phi\left(4.176\right)=0.0000148$$

本质上也就是说利用Monte Carlo方法均匀撒点的时候，在空白的地方浪费比较多，因此用重要度在可能性大的地方撒较多点

$$E\left(\sum_{k=1}^{n} kX_k\right)=\sum_{k=1}^{n} k\cdot E\left(X_k\right)=\frac{n(n+1)}{2}\cdot\frac{(n+1)}{2}=\frac{n(n+1)^2}{4}$$

$$E\left(\left(\sum_{k=1}^{n} kX_k\right)^2\right)=E\left(\sum_{k=1}^{n} k^2 X_k{}^2+\sum_{\substack{i\neq j\\ 1\le i,j\le n}} ij\cdot X_i X_j\right)=\sum_{k=1}^{n} k^2 E\left(X_k{}^2\right)+\sum_{\substack{i\neq j\\ 1\le i,j\le n}} ij\cdot E\left(X_i X_j\right)$$

$$=\sum_{k=1}^{n} k^2\cdot\frac{\sum_{k=1}^{n} k^2}{n}+\sum_{\substack{i\neq j\\ 1\le i,j\le n}} i\cdot j\cdot\frac{\sum_{\substack{i\neq j\\ 1\le i,j\le n}} i\cdot j}{n(n-1)}=\frac{\left(\sum_{k=1}^{n} k^2\right)^2}{n}+\frac{\left(\left(\frac{n(n+1)}{2}\right)^2-\sum_{k=1}^{n} k^2\right)^2}{n(n-1)}$$

$$=\frac{n^2(n+1)^2(2n+1)^2}{36n}+\frac{\left(\frac{n^2(n+1)^2}{4}-\frac{n(n+1)(2n+1)}{6}\right)^2}{n(n-1)}=\frac{n(n+1)^2(2n+1)^2}{36}+\frac{n(n+1)^2(3n+2)^2(n-1)}{144}$$

$$Var\left(\sum_{k=1}^{n} kX_k\right) = E\left(\left(\sum_{k=1}^{n} kX_k\right)^2\right) - E\left(\sum_{k=1}^{n} kX_k\right)^2$$

$$= \frac{n(n+1)^2(2n+1)^2}{36} + \frac{n(n+1)^2(3n+2)^2(n-1)}{144} - \frac{n^2(n+1)^4}{16}$$

$$= \frac{n(n+1)^2}{144}\left(4\cdot(2n+1)^2 + (3n+2)^2(n-1) - 9n(n+1)^2\right)$$

$$= \frac{(n-1)n^2(n+1)^2}{144}$$

但是这个时候方差较大，这样做是为了避免概率太小的时候的误差

重要度抽样，我们想要产生能够使得 $P\left\{\sum_{j=1}^{100} j\cdot X_j > 290000\right\}$ 较大的 $X$ 的排列。实际上我们希望得到0.5附近的概率。当 $X_j = j, j = 1,2,\cdots,100$ 时，$\sum_{j=1}^{100} j\cdot X_j$ 将达到最大值。

实际上，当 $j$ 较大时，若 $X_j$ 也较大，则求和会倾向于较大的值。我希望生成倾向于这种类型的 $X$ 排列的方法。

**以此来增加所求的概率**

$$X \sim b(1,p), Y = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

$$E\left(\left|Y - E(Y)\right| \le \varepsilon\right) \ge 1 - \frac{Var(Y)}{\varepsilon^2} = 1 - \frac{Var(X)}{n\varepsilon^2} \ge 1 - \delta$$

$$\Rightarrow \quad n \ge \frac{Var(X)}{\varepsilon^2 \delta}$$

$$\Pr[(1-\varepsilon)\text{per } A \le Y \le (1+\varepsilon)\text{per } A] \ge 1-\delta$$

$$E\left(\left|\frac{Y-p}{p}\right| \le \varepsilon\right) = E\left(\left|Y - p\right| \le p\varepsilon\right) \ge 1 - \frac{Var(Y)}{p^2\varepsilon^2} = 1 - \frac{p(1-p)}{np^2\varepsilon^2} \ge 1 - \delta$$

$$\Rightarrow \quad n \ge \frac{1}{\varepsilon^2\delta}\frac{1-p}{p}$$

生成参数为$\lambda_j$的独立的指数分布随机变量$Y_j, j=1,2,\cdots,100$，其中$\lambda_j$，$j=1,2,\cdots,100$是单调递增的序列，具体值将在下面给出。对于$j=1,2,\cdots,100$，用$X_j$表示生成的这些值中第$j$大的数的角标，也就是说$Y_{X_1} > Y_{X_2} > \cdots > Y_{X_{100}}$。

---

因为$\lambda_j$递增，所以较大的$j$对应较小的$Y_j$，则$Y_j$排序较后，$X_j$则较大。相对于从所有可能排列中均匀取出的排列，现在得到的$(X_1, X_2, \cdots, X_{100})$，$\sum_{j=1}^{100} j \cdot X_j$倾向于更大的值。

现在计算 $E\left(\sum_{j=1}^{100} j\cdot X_j\right)$，用 $R(j)$ 表示 $Y_j$ 的位次，

$Y_{X_1}$ 的位次为 1，$Y_{X_2}$ 的位次为 2，则 $R(X_j)=j$，$j=1,2,\cdots,100$

$$E\left(\sum_{j=1}^{100} j\cdot X_j\right)=E\left(\sum_{j=1}^{100} R(X_j)\cdot X_j\right)=E\left(\sum_{j=1}^{100} j\cdot R(j)\right)=\sum_{j=1}^{100} j\cdot E\left[R(j)\right]$$

为了计算 $E(R(j))$，

当 $Y_j<Y_i$ 时，令 $I(i,j)=1$，否则令 $I(i,j)=0$，

$$R(j)=1+\sum_{i=1,i\neq j}^{n} I(i,j),$$

此式的意思即为 $Y_j$ 的位次是1加上比 $Y_j$ 大的 $Y_i$ 的个数

利用指数分布性质 $P\left(Y_j < Y_i\right) = \dfrac{\lambda_j}{\lambda_i + \lambda_j}$，得到期望

$$E\left[R(j)\right] = 1 + \sum_{i:i\neq j} \frac{\lambda_j}{\lambda_i + \lambda_j}。$$

$$E\left(\sum_{j=1}^{100} j \cdot X_j\right) = \sum_{j=1}^{100} j \cdot E\left[R(j)\right] = \sum_{j=1}^{100} j \cdot \left(1 + \sum_{i:i\neq j} \frac{\lambda_j}{\lambda_i + \lambda_j}\right)。$$

如果我们令 $\lambda_j = j^{0.7}$，，可得 $E\left(\sum_{j=1}^{100} j \cdot X_j\right) = 290293.6$，此时，

将会有 $P\left\{\sum_{j=1}^{100} j \cdot X_j > 290000\right\} \approx 0.5$。

具体的模拟估计步骤为：

先以各自的参数 $\lambda_j = j^{0.7}$ 生成独立的指数分布随机变量 $Y_j$，

$j = 1, 2, \cdots, 100$；

再将 $Y_j$ 排序，令 $X_k$ 等于位次第 $k$ 位的 $Y_j$ 的下标；

若 $\sum_{j=1}^{100} j \cdot X_j > 290000$，则令 $I = 1$，否则令其为 0。

因此由 $Y_{X_1} > Y_{X_2} > \cdots > Y_{X_{100}}$ 得到排列 $X = \left(X_1, X_2, \cdots, X_{100}\right)$ 的

概率为 $\dfrac{\left(X_{100}\right)^{0.7}}{\sum_{j=1}^{100}\left(X_j\right)^{0.7}} \cdot \dfrac{\left(X_{99}\right)^{0.7}}{\sum_{j=1}^{99}\left(X_j\right)^{0.7}} \cdots \dfrac{\left(X_2\right)^{0.7}}{\sum_{j=1}^{2}\left(X_j\right)^{0.7}} \dfrac{\left(X_1\right)^{0.7}}{\left(X_1\right)^{0.7}}$，

因此单个抽样的重要度估计为

$$\hat{\theta} = \frac{I}{100!} \frac{\prod_{n=1}^{100}\left(\sum_{j=1}^{n}\left(X_j\right)^{0.7}\right)}{\left(\prod_{n=1}^{100} n\right)^{0.7}} = \frac{I \cdot \prod_{n=1}^{100}\left(\sum_{j=1}^{n}\left(X_j\right)^{0.7}\right)}{\left(\prod_{n=1}^{100} n\right)^{1.7}}$$

$$E\left(\frac{I \cdot \frac{1}{100!}}{\frac{(100!)^{0.7}}{\prod_{n=1}^{100}\sum_{k=1}^{n}(X_k)^{0.7}}}\right) = \theta$$

Available online at www.sciencedirect.com

ScienceDirect

Applied Mathematics and Computation 185 (2007) 59–71

APPLIED MATHEMATICS AND COMPUTATION

www.elsevier.com/locate/amc

# Random path method with pivoting for computing permanents of matrices ☆

Heng Liang [a,*], Linsong Shi [a], Fengshan Bai [a], Xiaoyan Liu [b]

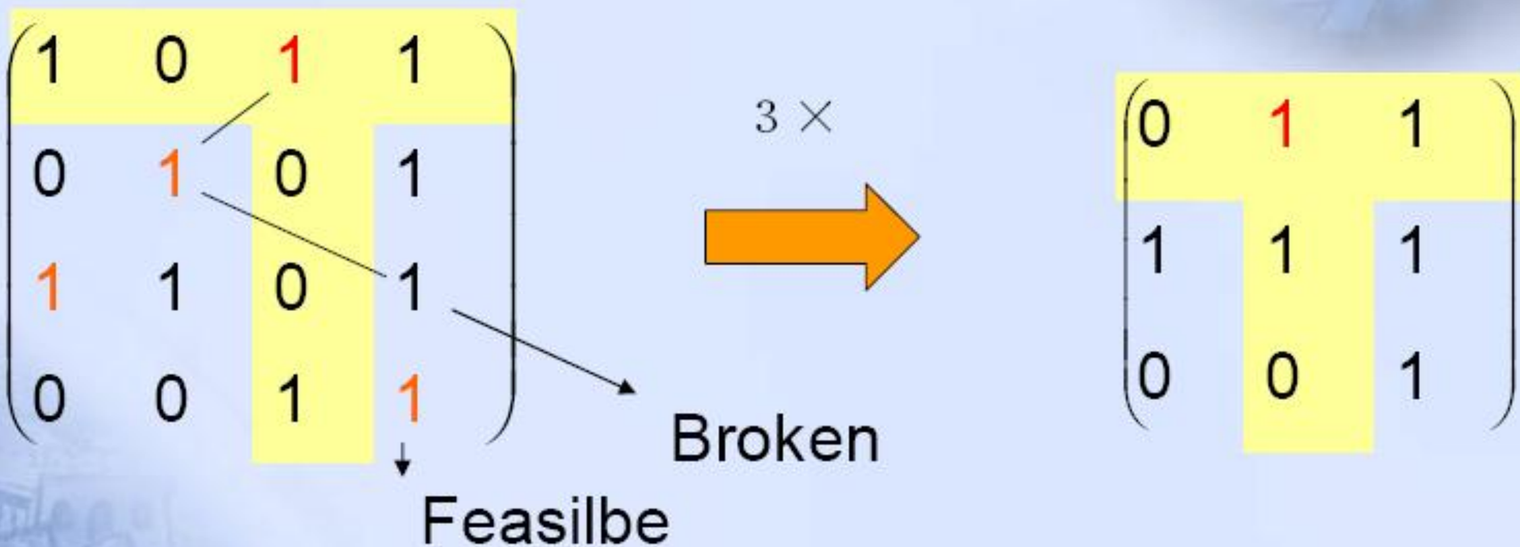
[a] *Department of Mathematical Sciences, Tsinghua University, Beijing 100084, PR China*
[b] *Department of Information Systems, City University of Hong Kong, Hong Kong*

**Abstract**

The permanent of matrix is important in mathematics and applications. Its computation, however, is #*P*-complete. Randomized algorithms are natural consideration to deal with such kind of problems. A Monte Carlo algorithm for approximating permanents of matrices is proposed in this paper, which improves a method by Rasmussen. Mathematical analysis and numerical computations show the efficiency of the method.


*Keywords:* Permanent; Monte Carlo method; Unbiased estimator; Random path

# 积和式的随机Laplace展开




$$X_A = p_1 \times p_2 \times \cdots \times p_n$$

$$E[X_A] = Per(A)$$

# 随机Laplace展开

**Example 3.1.** Consider a matrix given by

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

All its random paths are listed as follows.

$$\begin{array}{ccccccccc} i_1 & & i_2 & & i_3 & & i_4 & & \\ 1 & \Rightarrow & 2 & \Rightarrow & 3 & \Rightarrow & 4 & \sigma_1 = \{1,2,3,4\} & K_{\sigma_1} = 4 \\ 1 & \Rightarrow & 3 & \Rightarrow & \times & & & \sigma_2 = \{1,3,\sim\} & K_{\sigma_2} = 0 \\ 4 & \Rightarrow & 2 & \Rightarrow & 1 & \Rightarrow & \times & \sigma_3 = \{4,2,1,\sim\} & K_{\sigma_3} = 0 \\ 4 & \Rightarrow & 2 & \Rightarrow & 3 & \Rightarrow & 1 & \sigma_4 = \{4,2,3,1\} & K_{\sigma_4} = 8 \\ 4 & \Rightarrow & 3 & \Rightarrow & 1 & \Rightarrow & 2 & \sigma_5 = \{4,3,1,2\} & K_{\sigma_5} = 4 \end{array}$$

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

**Example 3.2.** Assume a matrix

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

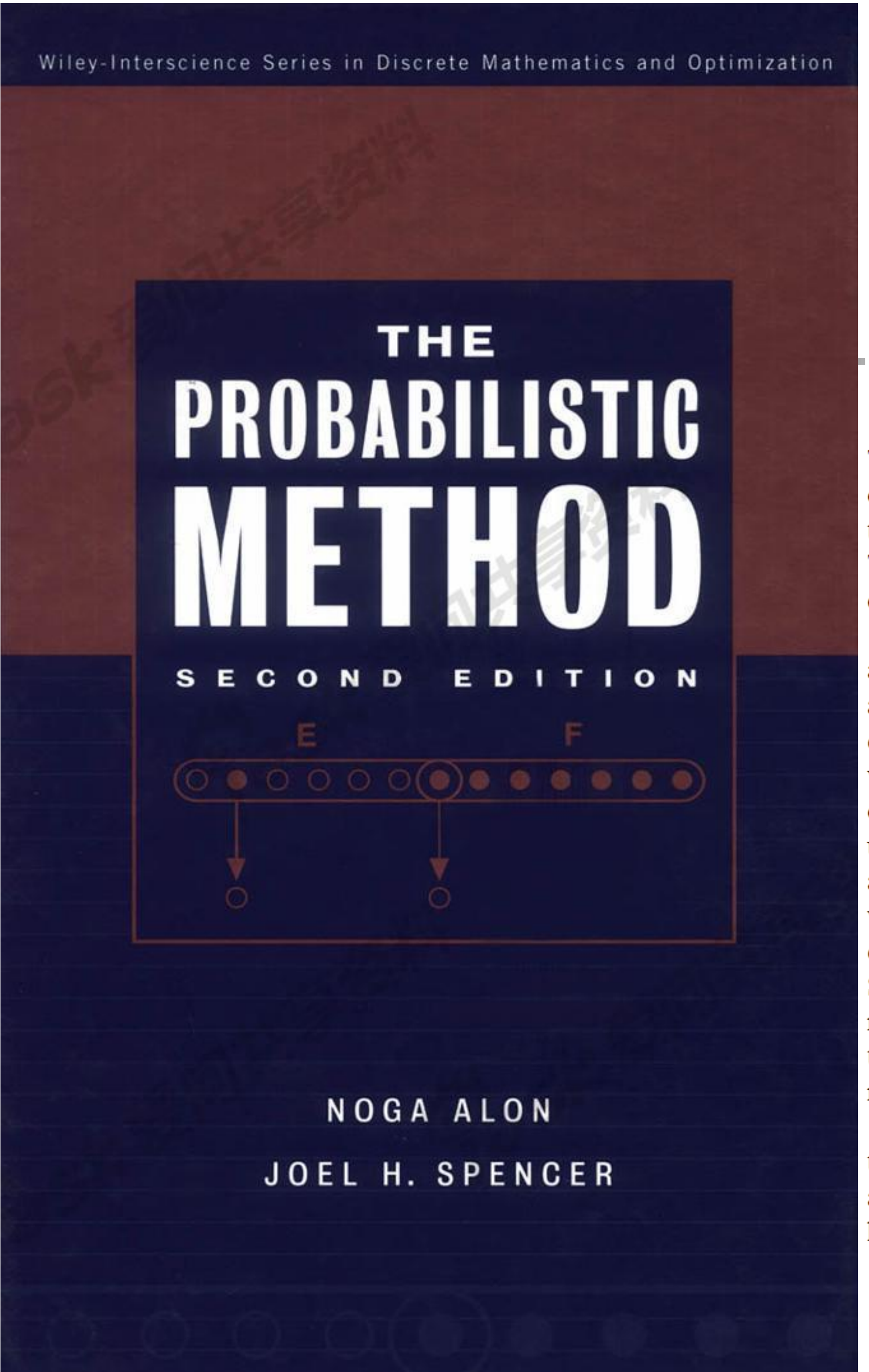
Hence all its random paths, together with their path values, are given as follows:

$$\begin{array}{ccccccc}
i_1 & & i_2 & & i_3 & & \\
1 & \Rightarrow & 3 & \Rightarrow & 2, & \sigma_1 = \{1, 3, 2\}, & K_{\sigma_1} = 3 \\
2 & \Rightarrow & 1 & \Rightarrow & \times, & \sigma_2 = \{2, 1, \sim\}, & K_{\sigma_2} = 0 \\
2 & \Rightarrow & 3 & \Rightarrow & \times, & \sigma_3 = \{2, 3, \sim\}, & K_{\sigma_3} = 0 \\
3 & \Rightarrow & 1 & \Rightarrow & 2, & \sigma_4 = \{3, 1, 2\}, & K_{\sigma_4} = 3
\end{array}$$

$$\overline{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

All its random paths of $\overline{A}$ are

$$\begin{array}{ccccccc}
i_1 & & i_2 & & i_3 & & \\
2 & \Rightarrow & 1 & \Rightarrow & 3, & \sigma_1 = \{2, 1, 3\}, & K_{\sigma_1} = 2 \\
2 & \Rightarrow & 3 & \Rightarrow & 1, & \sigma_2 = \{2, 3, 1\}, & K_{\sigma_2} = 2
\end{array}$$




# Preface

The Probabilistic Method has recently been developed intensively and became one of the most powerful and widely used tools applied in Combinatorics. One of the major reasons for this rapid development is the important role of randomness in Theoretical Computer Science, a field which is recently the source of many intriguing combinatorial problems.

The interplay between Discrete Mathematics and Computer Science suggests an algorithmic point of view in the study of the Probabilistic Method in Combinatorics and this is the approach we tried to adopt in this book. The manuscript thus includes a discussion of algorithmic techniques together with a study of the classical method as well as the modern tools applied in it. The first part of the book contains a description of the tools applied in probabilistic arguments, including the basic techniques that use expectation and variance, as well as the more recent applications of Martingales and Correlation Inequalities. The second part includes a study of various topics in which probabilistic techniques have been successful. This part contains chapters on discrepancy and random graphs, as well as on several areas in Theoretical Computer Science; Circuit Complexity , Computational Geometry, and Derandomization of randomized algorithms. Scattered between the chapters are gems described under the heading "The Probabilistic Lens". These are elegant proofs that are not necessarily related to the chapters after which they appear and can be usually read separately.

The basic Probabilistic Method can be described as follows: in order to prove the existence of a combinatorial structure with certain properties, we construct an appropriate probability space and show that a randomly chosen element in this space has the desired properties with positive probability. This method has been initiated

# *THE PROBABILISTIC LENS:*

# *Brégman's Theorem*

Let $A = [a_{ij}]$ be an $n \times n$ matrix with all $a_{ij} \in \{0, 1\}$. Let $r_i = \sum_{1 \le j \le n} a_{ij}$ be the number of ones in the $i$-th row. Let $S$ be the set of permutations $\sigma \in S_n$ with $a_{i,\sigma i} = 1$ for $1 \le i \le n$. Then the permanent $per(A)$ is simply $|S|$. The following result was conjectured by Minc and proved by Brégman (1973). The proof presented here is similar to that of Schrijver (1978).

**Theorem 1 [Brégman's Theorem]**

$$per(A) \le \prod_{1 \le i \le n} (r_i!)^{1/r_i}.$$

*International Journal of Computer Mathematics*
Vol. 86, No. 12, December 2009, 2181–2185

Taylor & Francis
Taylor & Francis Group

# 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

## 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 $perm(A)$.
*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 $p_i = \dfrac{1}{|W|}$;

$X = X \cdot |W|; \quad A = A[1|k]$;

*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.

THEOREM 2.2 (see [10,15]) *Let $X_A$ be the random variable given by Algorithm 2.1. Then*

$$E[X_A] = \mathrm{per}(A).$$

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*

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

## 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 $1 \le i \le n$. Denote $S = \{\sigma_1, \sigma_2, \ldots, \sigma_N\}$. Thus

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

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

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

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*

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

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

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

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

$$\prod_{j=1}^{N} \frac{1}{X_{A_k}(\sigma_j^{(k)})} \le \left[\frac{\sum_{j=1}^{N} 1/X_{A_k}(\sigma_j^{(k)})}{N}\right]^{N} = \left(\frac{P_k}{N}\right)^{N}. \tag{2}$$

Hence it is obvious that

$$\left(\frac{N}{P_k}\right)^{N} \le \prod_{j=1}^{N} X_{A_k}(\sigma_j^{(k)}). \tag{3}$$

From Equations (1) and (3)

$$\prod_{k=1}^{n!} \left(\frac{N}{P_k}\right)^{N} \le \prod_{k=1}^{n!} \prod_{j=1}^{N} X_{A_k}(\sigma_j^{(k)}) = \prod_{j=1}^{N} \left[\prod_{i=1}^{n} (r_i!)^{1/r_i}\right]^{n!}. \tag{4}$$

Journal of Computational Physics **149**, 128–147 (1999)
Article ID jcph.1998.6149, available online at http://www.idealibrary.com on IDEAL®

# Approximating the Permanent via Importance Sampling with Application to the Dimer Covering Problem

Isabel Beichl* and Francis Sullivan†

**NIST, Gaithersburg, Maryland 20899; †IDA/Center for Computing Sciences, Bowie, Maryland 20715*
E-mail: isabel@cam.nist.gov, fran@super.org

Received April 16, 1998; revised November 3, 1998

# 重要度抽样方法

$$m-\mathrm{bal}(A)=\begin{bmatrix} \frac{a_{11}|A_{11}|}{|A|} & \frac{a_{12}|A_{12}|}{|A|} & \cdots & \frac{a_{1n}|A_{1n}|}{|A|} \\ \frac{a_{21}|A_{21}|}{|A|} & \frac{a_{22}|A_{22}|}{|A|} & \cdots & \frac{a_{2n}|A_{2n}|}{|A|} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{a_{n1}|A_{n1}|}{|A|} & \frac{a_{n2}|A_{n2}|}{|A|} & \cdots & \frac{a_{nn}|A_{nn}|}{|A|} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$|A| = \frac{|A|}{\left|A_{(1,j_1)}\right|}\frac{\left|A_{(1,j_1)}\right|}{\left|A_{(1,j_1)(2,j_2)}\right|}\frac{\left|A_{(1,j_1)(2,j_2)}\right|}{\left|A_{(1,j_1)(2,j_2)(3,j_3)}\right|}\cdots\left|A_{(1,j_1)\cdots(n-1,j_{n-1})}\right|.$$

## 重要度抽样，理想的重要度函数

$$\iint_D f(x)\,dx$$

更一般地，生成$D$内的随机数$x_1, x_2, \cdots x_n$，满足概率密度$g(x)$

$$\iint_D f(x)\,dx \approx \frac{1}{n}\cdot\left(\frac{f(x_1)}{g(x_1)}+\frac{f(x_2)}{g(x_2)}+\cdots+\frac{f(x_n)}{g(x_n)}\right)$$

$$E\left(\frac{f(X)}{g(X)}\right) \approx \frac{1}{n}\sum_{k=1}^{n}\frac{f(x_k)}{g(x_k)} \quad \left(X \sim g(x)\right)，使得\ g(x) \approx \frac{f(x)}{\iint_D f(x)\,dx}$$

# Von Neumann的取舍原则

- 取常数M，满足 $f(x) < Mg(x)$

- Sample $x$ from $g(x)$ and $u$ from $U(0,1)$
- Check whether or not $u < \frac{f(x)}{Mg(x)}$ .
  - If this holds, accept $x$ as a realization of $f(x)$;
  - if not, reject the value of $x$ and repeat the sampling step.