Approximating the permanent of a matrix with nonnegative entries

Mark Huber[1] and Jenny Law[2]

[1]Department of Mathematics and Computer Science, Claremont McKenna College

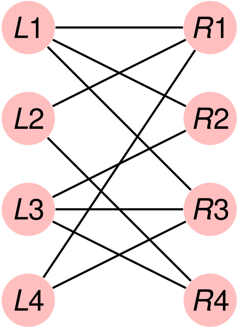
[2]Department of Mathematics, Georgia Perimeter College

31 Mar, 2010

A Permanent Problem

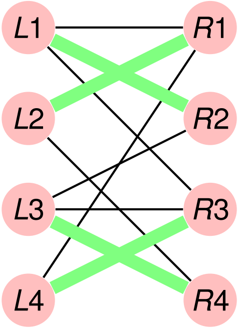
Perfect Matchings

A perfect matching in a graph is a collection of edges such that every node is adjacent to exactly one edge



Perfect Matchings

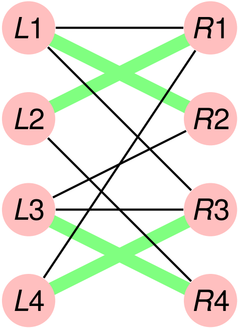
A perfect matching in a graph is a collection of edges such that every node is adjacent to exactly one edge



Encodings

More than one way to encode matching...

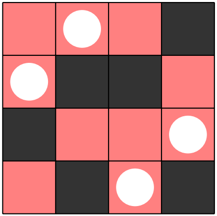
Perfect Matchings



Permutation

i	$\sigma(i)$
1	2
2	1
3	4
4	3

Rook Placement



Applications

- Nonparametric statistics (permutation statistical models)
- Contingency tables
- Graphs with given degree sequence
- One of first #P-complete problems

Goal

Basic Problem

- Count number of perfect matchings/restricted permutations/rook placements
- #P-complete problem, best can do is approximation
- Jerrum-Sinclair-Vigoda used Markov chain approach
- Even with improvements, still an $O(n^{10}(\log n)^4)$ algorithm
- Today: how to get $O(n^4 \log n)$ but only for dense problems

Make the problem harder

Weighted problem

- Give each edge a nonnegative weight
- (Chessboard now matrix with nonnegative entries)
- Weight of perfect matching product of weights of edges:

$$\text{wt}(\text{matching}) = \prod_{e \in \text{matching}} \text{wt}(e) = \prod_{i=1}^{n} A(i, \sigma(i))$$

New goal

- New goal, find

$$Z = \sum_{\text{matchings}} \text{wt}(\text{matching})$$

- Call it the *permanent* of the matrix A:

$$\text{per}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} A(i, \sigma(i))$$

D v P

- Recall determinant of a matrix is:

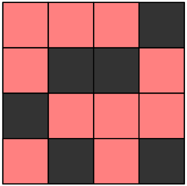
$$\det(A) = \sum_{\sigma \in S_n} (-1)^{\text{sgn}(\sigma)} \prod_{i=1}^{n} A(i, \sigma(i))$$

- Remove the sign of the permuation to get the permanent:

$$\text{per}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} A(i, \sigma(i))$$

Original problem

Allowable spaces get value "1", forbidden "0":



$$\Rightarrow \quad \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Comparison

Differences

- Determinant geometric/algebraic, simple invariants
- Can be found in $O(n^3)$ time
- Permanent combinatorial, #P-complete
- Best known algorithm Ryser: $O(n2^n)$ time

Similarities

- Expansion by minors works with both
- Multiply row by α, multiply per/det by α

Acceptance/Rejection

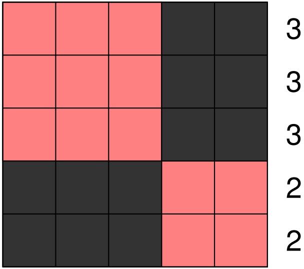
How to roll a seven sided die

1. Repeat
2.   Roll an 8-sided die
3. Until result at most 7

On average, takes $8/7$ rolls of 8-sided die
To get roll of 3-sided die, takes $8/3$ rolls on average
This works because $8 > 7$ and $8 > 3$

Upper bounds

Consider the following example:



$\text{per}(A) = 3!2!$

Minc (1963) conjectured that for given row sums, block structure maximizes permanent

Minc's conjecture

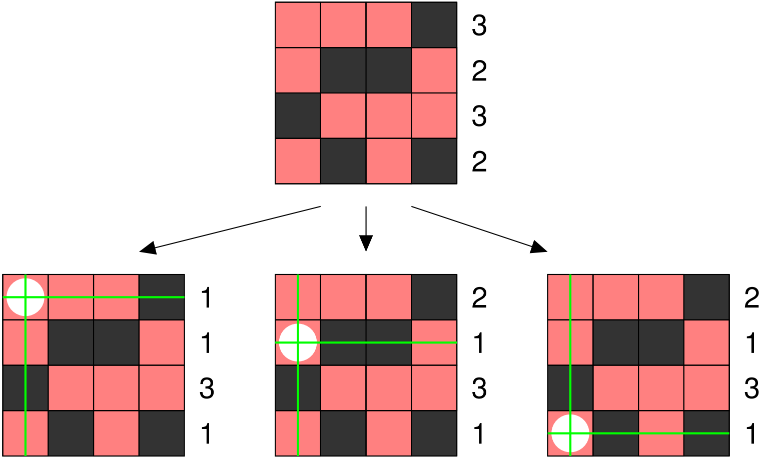
[Finally proved by Bregman in 1973]

For matrices whose entries are 0 or 1:

$$\text{per}(A) \leq \prod_{i=1}^{n} [r(i)!]^{1/r(i)},$$

where $r(i)$ is the sum of row i.

Problem: cannot prove via induction on minors



To work for upper bound...

Need

$$\text{bound}\begin{pmatrix}3\\2\\3\\2\end{pmatrix} \geq \text{bound}\begin{pmatrix}1\\1\\3\\1\end{pmatrix} + \text{bound}\begin{pmatrix}2\\1\\3\\1\end{pmatrix} + \text{bound}\begin{pmatrix}2\\1\\3\\1\end{pmatrix}$$

but for Minc/Bregman:

$$6.6038... \leq 1.81... + 2.57... + 2.57...$$

So can't use Minc's inequality exactly

By Stirling's approximation:

$$(r!)^{1/r} \approx \sqrt{2\pi r}\left(\frac{r}{e}\right)^{1/r} = (2\pi r)^{1/2r} \cdot \frac{r}{e}.$$

Get a handle on $r^{1/2r}$ using $e^x \approx 1 + x$ for small x:

$$\begin{aligned}
(r!)^{1/r} &= e^{(1/2r)(\log r + \log 2\pi)}(r/e) \\
&\approx \left[1 + \frac{1}{2r}(\log r + \log 2\pi)\right]\frac{r}{e} \\
&= \frac{r + (1/2)\ln r + (1/2)\ln(2\pi)}{e}
\end{aligned}$$

The Theorem/Method

Theorem

For a matrix A with entries either 0 or 1 and row sums $r(i)$:

$$\mathrm{per}(A) \leq \prod_{i=1}^{n} \frac{h(r(i))}{e},$$

where

$$h(r) = r + \frac{1}{2} \ln r + e - 1.$$

Moreover, this bound can be proved by induction on minors.

Example from earlier

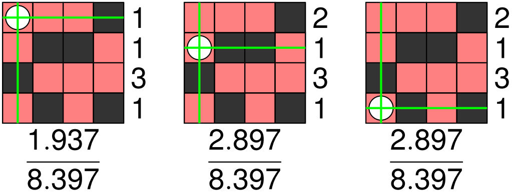
Need

$$\text{bound}\begin{pmatrix}3\\2\\3\\2\end{pmatrix} \geq \text{bound}\begin{pmatrix}1\\1\\3\\1\end{pmatrix} + \text{bound}\begin{pmatrix}2\\1\\3\\1\end{pmatrix} + \text{bound}\begin{pmatrix}2\\1\\3\\1\end{pmatrix}$$

and for new bound:

$$8.397... \geq 1.937... + 2.897... + 2.897 = 7.733...$$

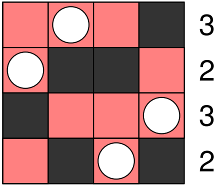
Sampling algorithm



Procedure

- Choose rook placement for first column...
- ...with probability bound of minor over original bound
- By Theorem, these ratios add to less than one
- If no placement chosen, reject and start over
- Otherwise continue until all rooks placed

Why does this work?



Probability of making this choice:

$$\frac{h(2,1,3,1)}{h(3,2,3,2)} \cdot \frac{h(1,1,2,1)}{h(2,1,3,1)} \cdot \frac{h(1,1,1,1)}{h(1,1,2,1)} \cdot \frac{h(1,1,1,1)}{h(1,1,1,1)} = \frac{1}{h(3,2,3,2)}$$

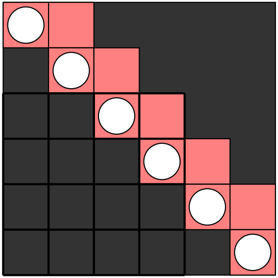
where $(3,2,3,2)$ are original row sums

How long does in take to run?

Average number of steps:

$$\frac{\text{bound}(\text{per}(A))}{\text{per}(A)}$$

Can be exponential in n
Example: A diagonal plus superdiagonal:



$\text{per}(A) = 1$

$\text{bound}(\text{per}(A)) = (1.495\ldots)^{n-1}$

Speeding things up

Upper and lower bounds

Regular graphs

- Minc: permanent large when 1's are clumped
- Van der Waerden: permanent small when entries spread out as much as possible

Lower bounding the permanent

Regular graphs

- Regular graphs have all degrees of all nodes are same
- Matrices all row and column sums are same, say r
- Van der Waerden lower bound for regular matrices

$$\text{per}(A) \geq \text{per}\begin{pmatrix} r/n & r/n & \cdots & r/n \\ r/n & r/n & \cdots & r/n \\ \vdots & \vdots & \ldots & \vdots \\ r/n & r/n & \cdots & r/n \end{pmatrix} = r^n \frac{n!}{n^n}$$

- Proved indep. by Egorychev (1981), Falikman (1981)

Use Stirling as before

For regular problems:

$$\sqrt{2\pi n}\left(\frac{r}{e}\right)^n \leq \text{per}(A) \leq \left(\frac{r + (1/2)\ln r + e - 1}{e}\right)^n$$

Looking at ratio of upper to lower bound:

$$\frac{\text{upper}}{\text{lower}} = \frac{1}{\sqrt{2\pi n}}\left(1 + \frac{1}{2r}\ln r + e - 1\right)^n \approx \frac{1}{\sqrt{2\pi n}} r^{n/2r}(e-1)^{n/r}$$

So if problem is dense:

$$r \geq \gamma n, \quad \gamma \in [0, 1]$$

then runtime polynomial

Achieving regularity

Sinkhorn step

1. Divide each row by its row sum
2. Divide each column by its column sum

Repeat until nearly regular

- Row and col sums in $[1 - n^{-2}, 1 + n^{-2}]$ in $O(n^{4.5})$ steps
- Can be done in $O(n^4)$ using ellipsoid method

Sinkhorn in action

$$\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 3 \\ 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 & 3 \\ 1 & 0 & 1 & 0 & 2 \\ \hline 3 & 2 & 3 & 2 & \end{array} \to \begin{array}{cccc|c} 1/3 & 1/3 & 1/3 & 0 & 1 \\ 1/2 & 0 & 0 & 1/2 & 1 \\ 0 & 1/3 & 1/3 & 1/3 & 1 \\ 1/2 & 0 & 1/2 & 0 & 1 \\ \hline 4/3 & 2/3 & 7/6 & 3/6 & \end{array}$$

$$\to \begin{array}{cccc|c} 1/4 & 1/2 & 2/7 & 0 & 29/28 \\ 3/8 & 0 & 0 & 3/5 & 39/40 \\ 0 & 1/2 & 2/7 & 2/5 & 73/70 \\ 3/8 & 0 & 3/4 & 0 & 45/56 \\ \hline 1 & 1 & 1 & 1 & \end{array}$$

After many Sinkhorn steps

$$\begin{array}{cccc|c}
.2069 & .5450 & .2479 & 0 & \approx 1 \\
.3381 & 0 & 0 & .6618 & \approx 1 \\
0 & .4549 & .2069 & .3381 & \approx 1 \\
.4549 & 0 & .5450 & 0 & \approx 1 \\
\hline
1 & 1 & 1 & 1 &
\end{array}$$

This has permanent about .1359

Making rows as large as possible

Last scaling of rows

- Want row sums as large as possible
- Divide each row through by its maximum entry
- All entries still in $[0, 1]$

$$\begin{array}{cccc|c} .3797 & 1 & .4549 & 0 & 1.8346 \\ .5108 & 0 & 0 & 1 & 1.5108 \\ 0 & 1 & .4549 & .7432 & 2.1982 \\ .8346 & 0 & 1 & 0 & 1.8346 \end{array}$$

The Theorem/Method

Theorem

For a matrix A with entries in $[0,1]$ and row sums $r(i)$:

$$\operatorname{per}(A) \leq \prod_{i=1}^{n} \frac{h(r(i))}{e},$$

where

$$h(r) = \begin{cases} r + \frac{1}{2}\ln r + e - 1 & r > 1 \\ 1 + (e-1)r & r \leq 1 \end{cases}$$

Moreover, this bound can be proved by induction on minors.

Summary: Speeding things up

What to do

- Regularize matrix using Sinkhorn steps or ellipsoid method
- Van der Waerden's inequality lower bounds permanent

Theorem (Huber,Law 2008)

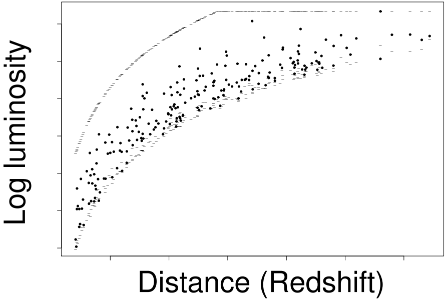
If original row sums at least γn *for* $\gamma \in (.5, 1]$*, then*

$$\frac{\text{bound}(\text{per}(A))}{\text{per}(A)} \leq n^{-.5+.5/(2\gamma-1)}.$$

An application

Nonparametric regression

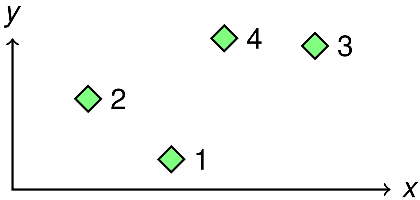
Are quasars that are far away brighter? Some data:



Is sample correlation high only because of truncation?

Nonparametric approach

Data set encodes permuation

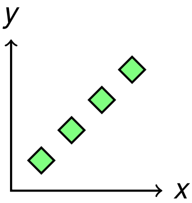


Measure relationship by counting inversions

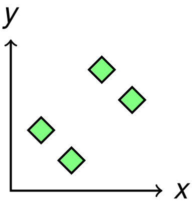
- Inversion when items “out of order” in permutation
- Ex: 2143 has four inversions
- 2 before 1, 1 before 4, 1 before 3, 4 before 3

Possbilities

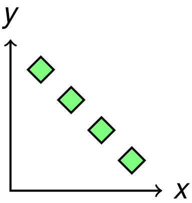
Three possibilities:



No inversions

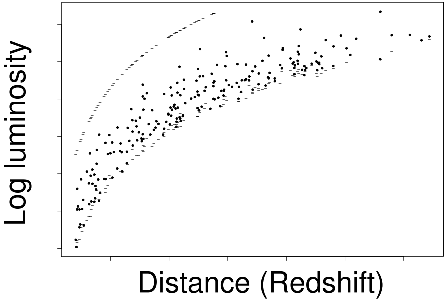


2 inversions



6 inversions

Understanding distribution of number of inversions



Take random permutations that don't violate truncation
So: perfect matching/restricted permutation/rook placement
Count number of inversions in random permuation

Conclusions

First perfect simulation for permanent

- Sadly, only provably poly time on dense problems
- But on these dense problems, much faster than Markov chain method
- First Markov chain approach only poly time on dense problems

Open question

- Can general problem be reduced to dense problem?
- Current Markov chain approach doesn't do that

References

M. Jerrum, L. Valiant, and V. Vazirani.
Random generation of combinatorial structures from a uniform distribution
Theoret. Comput. Sci., **43**, 169–188, 1986

M. Huber.
Exact Sampling from perfect matchings from dense regular bipartite graphs
Algorithmica, **44**, 183–193, 2006

M. Huber and J. Law.
Fast approximation of the permanent for very dense problems
Proc. of Syposium on Discrete Algorithms, 681–689, 2008