Counting and Sampling **Fall 2017**

Lecture 14: Barvinok's Method: A Deterministic Algorithm for Permanent

Lecturer: Shayan Oveis Gharan *November 15th*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

In this lecture we will prove the following theorem:

Theorem 14.1. *For any $\delta < 0.5$, $\epsilon > 0$, and any matrix $A \in \mathbb{C}^{n\times n}$ such that*

$$|1 - A_{i,j}| \le \delta, \forall i, j$$

there exists a polynomial $p_{n,\delta,\epsilon}$ of degree $O(\ln n - \ln \epsilon)$ such that

$$|\ln \operatorname{per} A - p(A)| \le \epsilon.$$

Furthermore, the polynomial $p(A)$ can be computed in quasi-polynomial time in n.

Recall that the theorem of Jerrum-Sinclair-Vigoda [JSV04] shows that as long as $A \ge 0$ we can use MCMC technique to give a $1+\epsilon$ approximation to $\operatorname{per}(A)$. But, if the entries of A can be negative (or even a complex number) we have no other tool besides this theorem to estimate $\operatorname{per}(A)$.

To prove this theorem, we use an elegant machinery of Barvinok. A weaker version of this theorem first appeared in [Bar16]. Parts of the proof that we are going to present here is from a more recent proof in [Bar17]. In the future lectures we will see many more applications of this machinery in other counting problems.

14.1 Estimating a Polynomial in the Zero Free Region

Essentially Lemma 14.2 shows that because the polynomial $g(z)$ is zero-free around zero the first few coefficients have enough information to estimate the polynomial in this regions.

Lemma 14.2. *Let $g(z)$ be a (complex) polynomial of degree d and suppose $g(z) \neq 0$ for all $|z| \le \beta$ where $\beta > 1$. Consider degree m taylor approximation of $f(z) = \ln g(z)$,*

$$p_m(z) = f(0) + \sum_{k=1}^{m} \frac{d^k}{dz^k} f(z)|_{z=0} \frac{z^k}{k!}.$$

Then, for all $|z| \le 1$,

$$|f(z) - p_m(z)| \le \frac{d}{(m+1)\beta^m(\beta - 1)}.$$

In other words, for $m = O_\beta(\ln d/\epsilon)$, $p_m(z)$ approximates $f(z)$ up to an additive ϵ error. Note that if β is very close to 1, we need to choose $m = O(\frac{1}{1-\beta} \ln(d\beta/\epsilon))$.

Proof. Let $r_1, \dots, r_d$ be the roots of $g(z)$. So,

$$g(z) = \prod_{i=1}^{d}(r_i - z) = g(0)\prod_{i=1}^{d}(1 - z/r_i).$$

So,

$$f(z) = \ln g(z) = f(0) + \sum_{i=1}^{d} \ln(1 - z/r_i)$$

Expanding the taylor series of the logarithm up to degree n,

$$\ln(1 - z/r_i) = -\sum_{k=1}^{m} \frac{z^k}{k r_i^k} + \zeta_{i,m},$$

and we can upper bound $\zeta_{i,m}$ by

$$\zeta_{i,m} = \left| \sum_{k=m+1}^{\infty} \frac{z^k}{k r_i^k} \right| \leq \frac{1}{(m+1)\beta^m(\beta - 1)}$$

where we used that $|z| \leq 1$ and that $|r_i| \geq \beta$. It follows that

$$f(z) = f(0) - \sum_{i=1}^{d} \sum_{k=1}^{n} \frac{z^k}{k r_i^k} + \zeta_m,$$

where $\zeta_m \leq \frac{d}{(m+1)\beta^m(\beta-1)}$. Finally, observe that the above equation gives the taylor series of $f(z)$ up to degree m, so

$$\frac{1}{k!} \frac{d^k}{dz^k} f(z)|_{z=0} = \sum_{i=1}^{d} \frac{1}{k r_i^k}.$$

This completes the proof. □

To compute the polynomial p_m we need to know the first m derivates of f at $z = 0$. Here, we show that if we have access to the first m derivatives of g at $z = 0$ we can use them to efficiently compute the first m derivatives of $f = \ln g$ at $z = 0$. The idea is to just use a system of linear equations. First observe that

$$f'(z) = \frac{g'(z)}{g(z)} \Rightarrow g'(z) = f'(z)g(z).$$

In general one can observe that

$$\frac{d^k}{dz^k} g(z)|_{z=0} = \sum_{j=0}^{k-1} \binom{k-1}{j} \left(\frac{d^{k-j}}{dz^{k-j}} f(z)|_{z=0} \right) \left(\frac{d^j}{dz^j} g(z)|_{z=0} \right)$$

So, in particular,

$$\begin{aligned} g''(0) &= g'(0)f'(0) + f''(0)g(0), \\ g'''(0) &= g''(0)f'(0) + 2g'(0)f''(0) + g(0)f'''(0), \ldots \end{aligned}$$

So, from the above we can compute all n derivatives of f at $z = 0$ in $O(m^2)$ time. In the next lecture we will discuss an extension of the polynomial approximation method under a somewhat weaker assumption on the no-root region of the polynomial $g(z)$.

14.2 Approximating Permanent with a Low Degree Polynomial

Armed with Lemma 14.2 all we need to do to prove Theorem 14.1 is to construct a polynomial $g(z)$ such that at for some $|z| \leq 1$, $g(z) = \text{per}(A)$, and that $g(z) \neq 0$ for all $|z| \leq \beta$ for some $\beta > 1$.

Consider the following polynomial:

$$g(z) = \text{per}(J + z(A - J)),$$

where $J \in \mathbb{R}^{n \times n}$ is the all-ones matrix. The following facts are immediate:

$$\begin{aligned} g(0) &= \text{per}(J) = n!, \\ g(1) &= \text{per}(A). \end{aligned}$$

So, all we need to do is to estimate $g(1)$. Now, we need to show that g has no roots in the ball of a radius $\beta > 1$ around the origin. This is in fact the main technical part of the proof.

Theorem 14.3. *Let $A \in \mathbb{C}^{n \times n}$. There exists an absolute constant $\delta_0 > 0$ such that if for all i, j*

$$|1 - A_{i,j}| \leq \delta_0,$$

then $\text{per}(A) \neq 0$.

We will see later that in the proof of the above theorem we can let $\delta_0 \geq 0.5$.

Before proving the above theorem first we use it to prove Theorem 14.1. Let $\beta = \frac{\delta_0}{\delta}$. First observe that for any z such that $|z| \leq \beta$ and for any i, j we have

$$|(J + z(A - J))_{i,j} - 1| = |1 + z(A_{i,j} - 1) - 1| = |z(A_{i,j} - 1)| \leq |z| \cdot |A_{i,j} - 1| \leq \frac{\delta_0}{\delta} \cdot \delta = 1.$$

Therefore, by the above theorem for any $|z| \leq \beta$, $g(z) \neq 0$. Now, by Lemma 14.2 for $m = O(\frac{1}{1-\beta} \ln(n\beta/\epsilon))$ and

$$p_m(z) = n! + \sum_{k=1}^{m} \frac{d^k}{dz^k} \ln g(z)|_{z=0} \frac{z^k}{k!}$$

satisfies $|\ln g(z) - p_m(z)| \leq \epsilon$. So, all we need to do is to compute the k-th derivative of $\ln g(z)$ for $k \leq m$. As we discussed in the previous section, equivalently, it is enough to know the k-th derivative $g(z)$ for $k \leq m$.

We can write

$$\frac{d^k}{dz^k} g(z) = \frac{d^k}{dz^k} \sum_{\sigma} \prod_{i=1}^{n} (1 + z(A_{i,\sigma_i} - 1)) = k!(n-k)! \sum_{\substack{(i_1, i_2, \ldots, i_k) \\ (j_1, \ldots, j_k)}} (A_{i_1,j_1} - 1) \ldots (A_{i_k,j_k} - 1),$$

where the last sum is over all pairs of ordered k subsets $(i_1, \ldots, i_k)$ and $(j_1, \ldots, j_k)$ of indices between $1, \ldots, n$. The $(n-k)!$ constant is because each such pair appears in exactly $(n-k)!$ many permutations and the $k!$ is because of differentiating z^k, k times. Note that in other words, the RHS of the above is just proportional to the sum of the permanents of all $k \times k$ submatrices of the matrix $A - J$. Therefore we can compute $g^{(k)}(0)$ in time $n^O(k)$. For $k \leq m$ this can be done in quasi-polynomial time. This completes the proof of Theorem 14.1.

14.3 Zero Free Region of Permanent

In this section we prove Theorem 14.3. This part is in a sense the only part of the proof which heavily depends on the permanent as a function. As we will see in future in several applications of the Barvinok's method typically one can compute the first $\log(n/\epsilon)$ coefficients of the corresponding polynomial in quasi-polynomial time and some cases in polynomial time. Therefore, the main nontrivial part of the proof is to find zero-free region for the that polynomial.

The proof is by a clever induction. Let $\mathcal{U}_n$ be the set of all $\mathbb{C}^{n\times n}$ complex matrices such that for all $A \in \mathcal{U}_n$, and for all i, j

$$|1 - A_{i,j}| \leq \delta_0.$$

We want to induct on n. Ideally, we would just need that for all $A \in \mathcal{U}_n$, $\text{per}(A) \neq 0$. But that is not enough for the induction. We strengthen the hypothesis assuming that for all $A, B \in \mathcal{U}_n$ that differ in one row or one column only, the angle between $\text{per}(A), \text{per}(B)$ does not exceed α. The means that if we consider each complex number $\text{per}(A)$ as a vector in $\mathbb{R}^2$, the angle between any two vectors corresponding to two matrices that differ in exactly one row (or one column) is at most α. We leave α as a parameter now, but later we will see that we can take $\alpha = \pi/2$.

We leave the base case as an exercise. Here we prove the claim for $\mathcal{U}_n$ assuming it holds for $\mathcal{U}_{n-1}$. The main important property of the permanent that we use in the proof is that permanent is a linear function of any single row or a column of the matrix and that it is invariant under permuting rows/columns.

Fix a matrix $A \in \mathcal{U}_n$; we can write

$$\text{per}(A) = \sum_{j=1}^{n} A_{1,j} \,\text{per}(A_j),$$

where A_j is the matrices obtained from A by removing the first row and the j-th column of A. Since A_j is a submatrix of A, we have $A_j \in \mathcal{U}_{n-1}$. Now observe that any pair of matrices A_j, A_k differ in exactly one column (up to a permutation of columns). Therefore by induction hypothesis the angle between $\text{per}(A_j), \text{per}(A_k)$ is at most α. It follows form the following lemma that $\sum_i A_{1,i} \,\text{per}(A_i) \neq 0$.

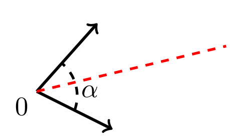
Lemma 14.4. *Let $u_1, \ldots, u_n \in \mathbb{C}$ be nonzero such that the angle between any two vectors u_i, u_j is at most α for some $0 < \alpha < 2\pi/3$. For $\delta_0 < \cos(\alpha/2)$ and any set of complex numbers $a_1, \ldots, a_n$ such that $|1 - a_i| \leq \delta_0$ for all i we have $\sum_i a_i u_i \neq 0$.*

Note that for $u_i = \text{per}(A_i)$ and $a_i = A_{1,i}$ we get from the lemma that $\text{per}(A) \neq 0$.

Proof. Let $u = u_1 + \cdots + u_n$.

Claim 14.5. $|u| \geq \cos\frac{\alpha}{2} \sum_{i=1}^{n} \|u_i\|$.

Proof. It turns out that all of these vectors lie in an angle at most α. This simply follows from the fact that $\alpha < 2\pi/3$. Note that if $\alpha = 2\pi/3$ then we could have three vectors with pairwise angle $2\pi/3$. Since $\alpha < 2\pi/3$ we can see that the origin is not in the convex all of these vectors, and therefore they lie in an angle of size α.



Now, project all vectors to the bisector of the angle; the projection of each u_i to the bisectors is at least $\cos\frac{\alpha}{2}\|u_i\|$. Since these projections do not cancel out each other, the projection of u onto the bisector is at least $\cos\frac{\alpha}{2}\sum_i\|u_i\|$ which proves the claim. □

Now, we are ready to finish the proof of the lemma. Let $v = \sum_i a_i u_i$. We use triangle inequality:

$$\begin{aligned} \|v\| &\geq \|u\| - \|v-u\| \\ &\geq \cos\frac{\alpha}{2}\sum_{i=1}^{n}\|u_i\| - \sum_{i=1}^{n}|1-a_i|\cdot\|u_i\|. \\ &\geq (\cos\frac{\alpha}{2} - \delta_0)\sum_{i=1}^{n}\|u_i\| \neq 0, \end{aligned}$$

where the second to last inequality uses that $|1-a_i| \leq \delta_0$ and the last inequality uses that $\delta_0 < \cos(\alpha/2)$. □

Now, we have proven part of the induction step; we know that for any matrix $A \in \mathcal{U}_n$, $\mathrm{per}(A) \neq 0$. But to finish the proof we also need to show that for any pair of matrices $A, B \in \mathcal{U}_n$ that differ in one row (or column), the angle between $\mathrm{per}(A), \mathrm{per}(B)$ is at most α. Fix two matrices $A, B \in \mathcal{U}_n$ and assume that they only differ in their first row. Therefore, we can write:

$$\begin{aligned} \mathrm{per}(A) &= \sum_{i=1}^{n} A_{1,i}\,\mathrm{per}(A_i), \\ \mathrm{per}(B) &= \sum_{i=1}^{n} B_{1,i}\,\mathrm{per}(A_i). \end{aligned}$$

We again use the proof strategy of Lemma 14.4; let u, v be the vectors defined in that lemma. It follows from the following simple fact that the angle between u, v is at most $\arcsin\frac{\delta_0}{\cos\frac{\alpha}{2}}$.

Fact 14.6. *For any two vectors x, y if $\|x\| < \|y\|$, then the angle between $y, x+y$ is at most* $\arcsin\frac{\|x\|}{\|y\|}$.

By symmetry we can show that the angle between $\sum_i B_{1,i}\,\mathrm{per}(A_i)$ and u is at most $\arcsin\frac{\delta_0}{\cos\frac{\alpha}{2}}$. Therefore, the angle between $\mathrm{per}(A), \mathrm{per}(B)$ is at most

$$2\arcsin\frac{\delta_0}{\cos\frac{\alpha}{2}}.$$

So, we only need that the above quantity is at most α. So, letting $\alpha = \pi/2$ and $\delta_0 = 0.5$ is enough for our purpose. This completes the proof of Theorem 14.3.

References

[Bar16] A. Barvinok. Computing the permanent of (some) complex matrices, 2016. 14-1

[Bar17] A. Barvinok. Approximating permanents of hafnians, 2017. 14-1

[JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, July 2004. 14-1