The Polynomial Paradigm in Algorithm Design	**Winter 2020**

Lecture 3: Log Concavity Property

Lecturer: Shayan Oveis Gharan — *Jan 17th*

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

3.1 Gurvits' Machinery

Given a matrix $M \in \mathbb{R}^{n\times n}_{\geq 0}$ the permanent of M is defined as follows:

$$\text{per}(M) = \sum_{\sigma\in\mathbb{S}_n}\prod_{i=1}^{n} M_{i,\sigma_i}.$$

This quantity is #P-hard to compute exactly so one can only appeal to approximation algorithms. The problem has been studied for years. After much effort, Jerrum, Sinclair and Vigoda [**JSV04**] managed to design a randomized $1\pm\epsilon$ approximation algorithm for this problem that runs in time polynomial in $n, 1/\epsilon$.

Designing a deterministic algorithm with the same approximation factor remains a *fundamental open problem*. This question relates to several fundamental problem in TCS under the umbrella of polynomial identity testing (PIT). It is believed that one can simulate any polynomial time randomized algorithm with a polynomial time deterministic algorithm. Permanent is perhaps the most fundamental problem which we have no deterministic approach for.

After a long line of works–see Linial, Samarodnitsky, Widgerson [**LSW**], Barvinok, Gurvits [**Gur02**], Gurvits, Samarodnitsky [**GS14**] and Anari, Rezaei [**AR18**]–the best deterministic approximation factor known is $(\sqrt{2})^n$. Note that if we can design a 2^{n^ϵ} approximation algorithm for any $\epsilon > 0$ then it is possible to improve the approximation factor to $1\pm\epsilon$.

In this section, we will prove the following result of Gurvits:

Theorem 3.1 (Gurvits [**Gur02**])**.** *Let $p \in \mathbb{R}[z_1,\dots,z_n]$ be a real stable polynomial with non-negative coefficients. Then,*

$$e^{-n}\inf_{z>0}\frac{p(z_1,\dots,z_n)}{z_1\dots z_n} \leq \partial_{z_1}\dots\partial_{z_n}p\big|_{z=0} \leq \inf_{z>0}\frac{p(z_1,\dots,z_n)}{z_1\dots z_n}.$$

In other words, the above theorem shows that we can approximate the coefficient of the monomial $z_1\dots z_n$ in the real stable polynomial p (with non-negative coefficient) with the mathematical program $\inf_{z>0}\frac{p(z_1,\dots,z_n)}{z_1\dots z_n}$.

First we show that $\inf_{z>0}\frac{p(z_1,\dots,z_n)}{z_1\dots z_n}$ can be formulated as a convex programming problem. The idea is to do a change of variables $e^{y_i} \leftrightarrow z_i$. Note that since $z_i > 0$ this is a valid change of variables. Furthermore to minimize the ratio we can equivalently minimize the log of the ratio. So, we claim,

$$\log\frac{p(e^y_1,\dots,e^{y_n})}{e^{y_1}\dots e^{y_n}} = \log p(e^y_1,\dots,e^{y_n}) - \sum_{i=1}^{n} y_i$$

is convex. To see this, it is enough to show $\log p(e^{y_1},\dots,e^{y_n})$ is convex, as the second term is linear. To see this, recall that log-sum-exp is a convex function i.e.,

$$\log \sum_{i=1}^{n} a_i e^{\langle b_i, y\rangle}$$

is convex as long as $a_1, \ldots, a_n \geq 0$ and $b_i \in \mathbb{R}^n$ for all i.

Now, we show how to use this theorem to get an e^n approximation for permanent. Consider the following polynomial

$$p(z_1, \ldots, z_n) = \prod_{i=1}^{n} \sum_{j=1}^{n} M_{i,j} z_j. \tag{3.1}$$

First, observe that p is real stable since all entries of M are nonnegative and the product of real stable polynomials is real stable. Second, observe the coefficient of the monomial $z_1 \ldots z_n$ is exactly per(M), i.e.,

$$\partial_{z_1} \ldots \partial_{z_n} p \mid_{z=0} = \text{per}(M).$$

Therefore, Theorem 3.1 gives an e^n approximation algorithm to the per(M).

Finally, we use this theorem to prove the van-der-Waerden conjecture:

Theorem 3.2. *Let $M \in \mathbb{R}^{n\times n}_{\geq 0}$ be a doubly stochastic matrix, i.e., the sum of the entries in every row and column is exactly 1. Then, per$(M) \geq e^{-n}$.*

Proof. By Theorem 3.1 we can write

$$\text{per}(M) \geq e^{-n} \inf_{z>0} \frac{p(z_1, \ldots, z_n)}{z_1 \ldots z_n}$$

where p is the same polynomial as in (3.1). So it is enough to show $\inf_{z>0} \frac{p(z_1,\ldots,z_n)}{\geq} z_1 \ldots z_n$. To see this we use the weighted AM-GM inequality.

Theorem 3.3 (Weighted AM-GM Inequality)**.** *Let $a_1, \ldots, a_n \geq 0$ and $\lambda_1, \ldots, \lambda_n \geq 0$ such that $\sum_i \lambda_i = 1$. Then,*

$$\sum_i \lambda_i a_i \geq \prod_i a_i^{\lambda_i}.$$

By the above theorem we can write

$$p(z_1, \ldots, z_n) = \prod_{i=1}^{n} \sum_{j=1}^{n} M_{i,j} z_j \geq \prod_{i=1}^{n} \prod_{j=1}^{n} z_j^{M_{i,j}} = \prod_{j=1}^{n} z_j^{\sum_{i=1}^{n} M_{i,j}} = \prod_{j=1}^{n} z_j,$$

as desired. The inequality follows by (weighted) AM-GM and the fact that every row of M adds up to 1 and the last equality follows by the fact that every column of M adds up to 1. □

Before proving Theorem 3.1 let us first gain some intuition. The right side of the inequality is trivial and it holds for any polynomial with non-negative coefficients. This is because if we let $p = \sum_{\alpha \in \mathbb{N}_{\geq 0}} c_\alpha z^\alpha$, then

$$\inf_{z>0} \frac{p(z_1, \ldots, z_n)}{z_1 \ldots z_n} = c_{1,1,\ldots,1} + \inf_{z>0} \sum_{\alpha \in \mathbb{N}_{\geq 0}} c_\alpha \frac{z^\alpha}{z_1 \ldots z_n} \geq c_{1,1,\ldots,1},$$

as desired.

Now, the left side of the inequality crucially uses the fact that p is real stable. For example, consider the non-stable polynomial $z_1^2 + z_2^2$. Then,

$$\inf_{z>0} \frac{z_1^2 + z_2^2}{z_1 z_2} = \inf_{z>0} \frac{z_1}{z_2} + \frac{z_2}{z_1} \geq 2,$$

where as the coefficient of $z_1 z_2$ is 0.

We start the proof of Theorem 3.1 by proving a univariate version of the theorem. This not only will serve as the base of the induction, but will also give us some clues on how to approach the main theorem.

Lemma 3.4. *For any real rooted polynomial $f \in \mathbb{R}[t]$ with nonnegative coefficients,*

$$f'(0) \geq \frac{1}{e} \inf_{t>0} f(t)/t.$$

Proof. First, we show that f is a log-concave function.

Fact 3.5. *Let $f \in \mathbb{R}[t]$ be a real rooted polynomial with nonnegative coefficients then $\log f$ is a concave function of $\mathbb{R}_{\geq 0}$.*

This is because (dropping the constant term for convenience):

$$\log f = \log \prod_{i=1}^{n} (t + \alpha_i) = \sum_{i=1}^{n} \log(t + \alpha_i),$$

where $-\alpha_i$'s are roots of f. Since f has all positive coefficients, it is non-zero over the positive reals so $\alpha_i \geq 0$ for all i. The claim follows because log is a concave function and the sum of concave functions is concave.

Next, we prove the lemma. Since $\log f$ is concave, for any $t \geq 0$ we can write,

$$\log f(t) \leq \log f(0) + t(\log f(0))'$$

Therefore, for any $t \geq 0$ we get

$$\log \frac{f(t)}{t} \leq \log f(0) + t\frac{f'(0)}{f(0)} - \log t$$

Setting $t = f(0)/f'(0)$ in the RHS we get

$$\inf_{t>0} \log \frac{f(t)}{t} \leq \log f(0) + 1 - \log \frac{f(0)}{f'(0)} = 1 + \log f'(0)$$

as desired. □

Having proved the lemma we are ready to prove Theorem 3.1. We prove by induction on n. Let $q(z_1, \ldots, z_{n-1}) = \partial_{z_n} p|_{z_n=0}$. Note that by closure properties of real stable polynomials q is a real stable polynomial. Furthermore, it has non-negative coefficients. Then,

$$\partial_{z_1} \ldots \partial_{z_n} p|_{z=0} = \partial_{z_1} \ldots \partial_{z_{n-1}} q|_{z=0} \geq e^{-(n-1)} \inf_{z_1,\ldots,z_{n-1}>0} \frac{q(z_1, \ldots, z_{n-1})}{z_1 \ldots z_{n-1}} \tag{3.2}$$

where the inequality follows by IH. Say the infimum in the RHS is attained at a point $z_1^*, \ldots, z_{n-1}^*$[1]. Define $f(z_n) = p(z_1^*, \ldots, z_{n-1}^*, z_n)$. Then, observe that

$$q(z_1^*, \ldots, z_{n-1}^*) = f'(0) \geq \frac{1}{e} \inf_{z_n>0} \frac{f(z_n)}{z_n} = \frac{1}{e} \inf_{z_n>0} \frac{p(z_1^*, \ldots, z_{n-1}^*, z_n)}{z_n}, \tag{3.3}$$

[1] to be more precise we need to do a ϵ-δ argument but here we avoid that for the simplicity of the argument.

where the inequality follows by the fact that f is a real stable polynomial with non-negative coefficients. Therefore,

$$\partial_{z_1}\dots\partial_{z_n} p|_{z=0} \geq e^{-(n-1)} \frac{q(z_1^*,\dots,z_{n-1}^*}{z_1^*\dots z_{n-1}^*} \geq e^{-n} \inf_{z_n>0} \frac{p(z_1^*,\dots,z_{n-1}^*,z_n)}{z_1^*\dots z_{n-1}^* z_n} \geq e^{-n} \inf_{z>0} \frac{p(z_1,\dots,z_n)}{z_1\dots z_n}$$

where the first inequality follows by (3.2) and the second inequality follows by (3.3). This finishes the proof of Theorem 3.1.

Remark 3.6. *The above proof is fairly general; it mainly works for polynomial whose univariate restrictions are log-concave. Gurvits has studied many generalizations and extensions of the above proof such as using it to estimate mixed volume of convex bodies. In the following lectures we will discuss a generalization of this proof and applications to problems in Economics and Game Theory.*

3.2 Log Concavity of Real Stable Polynomials

An immediate generalization of the Lemma 3.4 is the following theorem:

Theorem 3.7. *Let $p \in \mathbb{R}[z_1,\dots,z_n]$ be a homogeneous real stable polynomial. Then, p is log-concave over $\mathbb{R}^n_{\geq 0}$.*

This theorem can be seen as a generalization of the well-known fact that $\det(P)$ is log-concave over the space of PSD matrices.

Proof. In the HW we will see that all coefficients of p must be non-negative so $\log p$ is well defined over $\mathbb{R}^n_{\geq 0}$. It is enough to show that $\log p$ is concave along any interval in the positive orthant. Let $a, b \in \mathbb{R}^n_{>0}$, $b \in \mathbb{R}^n$, and consider the line $a+tb$ where for any $t \in [0,1]$, $a+tb \in \mathbb{R}^n_{>0}$. We show that $\log p(a+tb)$ is concave. Say p is k-homogeneous, then

$$p(a+tb) = p(t(a/t+b)) = t^k p(a/t+b).$$

Since $a \in \mathbb{R}^n_{>0}$, and $p(.)$ is stable, $p(at+b)$ is real rooted. Write $p(at+b) = p(a)\prod_{i=1}^k (t-\lambda_i)$ where $\lambda_1,\dots,\lambda_k$ are the roots.

Then, we have

$$p(a/t+b) = p(a)\prod_{i=1}^k (1/t-\lambda_i).$$

So,

$$p(a+tb) = t^k p(a/t+b) = p(a)\prod_{i=1}^k (1-t\lambda_i).$$

We claim that for all $1 \leq i \leq k$, $\lambda_i < 1$. Otherwise, for some $t \in [0,1]$, $p(a+tb)=0$, but since $a+tb \in \mathbb{R}^n_{>0}$, $p(a+tb)>0$ which is a contradiction. Therefore,

$$\log p(a+tb) = \log p(a) + \sum_{i=1}^k \log(1-t\lambda_i).$$

The theorem follows by the fact that $\log(1-t\lambda)$ is a concave function of t for $t \in [0,1]$ when $\lambda < 1$. □

Remark 3.8. *If p is real stable but not homogeneous but it has positive coefficients, it is still log-concave over $\mathbb{R}^n_{\geq 0}$. It turns out that in this case one can* homogenize *p and apply the above theorem.*

3.3 Maximum Sub-determinant Problem

In the maximum sub-determinant problem we are given a PSD matrix $M \succeq 0$ and an integer k and the goal is to output a set $S \in \binom{n}{k}$ such that the determinant of the square-submatrix $M_{S,S}$ is maximized. Here, we give a simple proof of a recent result of Nikolov [**Nik16**] using the theory of real stable polynomials.

Theorem 3.9. *There is a polynomial time randomized algorithm that gives a e^k approximation to the sub-determinant maximization problem.*

First we need to construct a real stable polynomial.

Lemma 3.10. *For any $k \geq 0$, and any $M \succeq 0$, the following polynomial is real stable*

$$\sum_{S \in \binom{n}{k}} \det(M_{S,S}) z^S.$$

Proof. Let

$$Z = \begin{bmatrix} z_1 & & & \\ & z_2 & & \\ & & \ddots & \\ & & & z_n \end{bmatrix}$$

We use the following linear algebraic identity:

$$\det(Z + tM) = \sum_{k=0}^{n} \sum_{S \in \binom{n}{k}} t^k z^{\overline{S}} \det(M_{S,S}).$$

The identity is a generalization of the fact that the k-th coefficient of the characteristic polynomial of M is the sum of the determinants of all square $k \times k$ minors of M.

By the mother-of-all theorem, since $M \succeq 0$ and Z is diagonal, $\det(Z + tM)$ is real stable. Therefore,

$$\partial_t^{n-k} \det(Z + tM)|_{t=0} = k! \sum_{S \in \binom{n}{k}} z^{\overline{S}}$$

is real stable. The lemma simply follows by the closure of real stable polynomials under inversion. □

Since $\sum_{S \in \binom{n}{k}} z^S$ is homogeneous, it is log-concave over $\mathbb{R}^n_{\geq 0}$. To prove Theorem 3.9 we use the following convex program:

$$\begin{aligned} \max \quad & \log \sum_{S \in \binom{n}{k}} \det(M_{S,S}) x^S \\ \text{s.t.,} \quad & \sum_i x_i = k \\ & x_i \geq 0 \qquad\qquad \forall 1 \leq i \leq k. \end{aligned} \tag{3.4}$$

Note that the above program is a relaxation of the problem as $x = \mathbb{I}[OPT]$ is a feasible solution. It follows that the objective function is at least OPT.

To round we construct the following distribution μ over $\{1, \ldots, k\}$ where $\mu(i) = x_i/k$. We generate k samples from μ, $i_1, \ldots, i_k$. If these are all distinct we output the set $\{i_1, \ldots, i_k\}$ otherwise we output nothing. We claim that this algorithm receives at least e^{-k} fraction of OPT in expectation.

For every set $S = \{i_1, \ldots, i_k\}$ the probability that we output S is

$$\frac{x_{i_1}}{k} \cdot \frac{x_{i_2}}{k} \cdots \frac{x_{i_k}}{k} \cdot k! = x^S \frac{k!}{k^k} \approx x^S e^{-k},$$

where the $k!$ term comes from the fact that we can sample the k elements of S in any of the $k!$ possible orders. Therefore,

$$\mathbb{E}\left[ALG\right] = \sum_{S \in \binom{n}{k}} \mathbb{P}\left[S \text{ sampled}\right] \det(M_{S,S}) = \sum_{S \in \binom{n}{k}} x^S e^{-k} \det(M_{S,S}) \geq OPT e^{-k}$$

as desired.