

Lecture 11: Approximate Arrow's theorem using the hypercontractivity lemma

1 Approximate Arrow's theorem

In previous lectures, we investigated how to get a social welfare function for ranking 3 candidates given a social choice function for pairwise comparisons. Arrow's theorem tells us that dictators are the only functions which are guaranteed to give a non-cyclic ranking. More generally, given a boolean social choice function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, under the *Impartial Culture* assumption, we had obtained the exact probability for a rational outcome :

$$\begin{aligned} \Pr[\text{no cycles}] &= \Pr[\text{NAE test passes}] \\ &= \frac{3}{4} - \frac{3}{4} \sum_{S \subseteq [n]} \left(-\frac{1}{3}\right)^{|S|} \hat{f}(S)^2 \\ &\leq \frac{7}{9} + \frac{2}{9} W_1(f) \end{aligned}$$

where $W_1(f)$ is the weight of the first level of fourier coefficients, i.e. $W_1(f) = \sum_{|S|=1} \hat{f}(S)^2$.

Remark 1.1 *The above probability equals 1 if and only if $W_1(f) = 1$, which implies that f is either a dictator or an anti-dictator (Refer to Homework 1).*

In this lecture, we want to ask if the above probability is not required to be exactly 1, then do there exist functions “considerably different” from dictators which can also give rational outcomes with a reasonably high probability. To put it more formally, we want to ask the following :

Suppose we have $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ such that $\Pr[\text{no cycles}] = 1 - \epsilon$. Is f $O(\epsilon)$ -close to being a (anti-)dictator ? It turns out that the answer to the question is yes. This is due to the following theorem by Friedgut, Kalai and Naor from 2002.

Theorem 1.2 (FKN theorem) *If $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$ has $\sum_{|S|>1} \hat{f}(S)^2 < \epsilon$, then f is $O(\epsilon)$ -close to a 1-junta.*

Remark 1.3 *The constant in $O(\cdot)$ above is quite small ($\approx 2 - 4$).*

Corollary 1.4 *If $W_1(f) = 1 - \epsilon$ then f is $O(\epsilon)$ -close to a dictator or an anti-dictator.*

The above corollary shows that the answer to the question posed above is yes, because

$$\begin{aligned} \Pr[\text{no cycles}] &= 1 - \epsilon \\ \Leftrightarrow 1 - \epsilon &\leq \frac{7}{9} + \frac{2}{9} W_1(f) \\ \Leftrightarrow W_1(f) &\geq 1 - \frac{9}{2} \epsilon \end{aligned}$$

Lecture 11: Approximate Arrow's theorem using the hypercontractivity lemma

We now prove the Friedgut, Kalai, Naor theorem.

Proof: It suffices to prove the corollary above, because suppose we have f such that $\sum_{|S| \leq 1} \hat{f}(S)^2 = 1 - \epsilon$. Define another function $g : \{-1, 1\}^{n+1} \rightarrow \{-1, 1\}$ by $g(x_0, x) = x_0 \cdot f(x_0 x)$. If the Fourier expansion of f is :

$$f(x) = \hat{f}(\phi) + \hat{f}(\{1\})x_1 + \dots + \hat{f}(\{n\})x_n + \hat{f}(\{1, 2\})x_1x_2 + \dots$$

then

$$g(x) = \hat{f}(\phi)x_0 + \hat{f}(\{1\})x_1 + \dots + \hat{f}(\{n\})x_n + \hat{f}(\{1, 2\})x_0x_1x_2 + \dots$$

$\therefore W_1(g) = 1 - \epsilon$. Note that $\hat{f}(\phi)$ goes to level 1 in g .

Assuming the corollary, g is $O(\epsilon)$ -close to some dictator or some anti-dictator. $\therefore |\hat{g}(i)| \geq 1 - O(\epsilon)$ for some $0 \leq i \leq n$.

$\therefore |\hat{f}(S)| \geq 1 - O(\epsilon)$ for some S with $|S| \leq 1$. Hence f is $O(\epsilon)$ -close to a 1-junta.

Remark 1.5 In proving the corollary, we can assume f is balanced. Henceforth, we will assume $\hat{f}(\phi) = 0$.

We now prove the corollary. We express f as $f(x) = \sum_{i=1}^n \hat{f}(i)x_i + \sum_{|S| > 1} \hat{f}(S)x_S$. We denote the first term (with lower order coefficients) by $l(x)$ and the second term (with higher order coefficients) by $h(x)$. Note that $l, h : \{-1, 1\}^n \rightarrow \mathfrak{R}$, but when they are added together they always “magically” add upto 1 or -1 .

By the hypothesis, $\sum_{i=1}^n \hat{f}(i)^2 = 1 - \epsilon$, which implies $\|h\|_2^2 = \mathbf{E}[h(x)^2] = \sum_{|S| > 1} \hat{f}(S)^2 = \epsilon$

It is easy to see that

$$\begin{aligned} f^2 &\equiv 1 \text{ (the square of a boolean function is identically 1)} \\ (l+h)^2 &\equiv 1 \\ l^2 + h(2l+h) &\equiv 1 \\ l^2 + h(2f-h) &\equiv 1 \end{aligned}$$

- $l(x)^2 = (\sum_{i=1}^n \hat{f}(i)x_i)^2 = \sum_{i=1}^n \hat{f}(i)^2 + \sum_{i \neq j} \hat{f}(i)\hat{f}(j)x_ix_j$. If we let $q(x) = \sum_{i \neq j} \hat{f}(i)\hat{f}(j)x_ix_j$, then $l(x)^2 = 1 - \epsilon + q(x)$.
- Consider $h(2f-h)$. $\therefore \mathbf{E}[h(x)] = 0$, $\mathbf{E}[h(x)^2] = \epsilon$, by Chebyshev's inequality,

$$\Pr[|h(x)| \geq 10\sqrt{\epsilon}] \leq 0.01 \text{ (1\% of the time)}$$

When $h(x)$ is not too large, $h(2f-h) \leq 10\sqrt{\epsilon}(2+10\sqrt{\epsilon}) \leq 21\sqrt{\epsilon}$ (assuming ϵ is sufficiently small)

Substituting the two facts derived above into $l^2 + h(2f-h) \equiv 1$, we have

$$q(x) \leq 22\sqrt{\epsilon} \quad \text{with probability } \geq 99\%$$

Lecture 11: Approximate Arrow's theorem using the hypercontractivity lemma

The above fact implies that $q(x)^2 \leq 484\epsilon$ with probability at least 0.99. The next crucial idea in the proof is the intuition that if $q(x)$ is a "reasonable" random variable then probably $\mathbf{E}[q(x)^2] \leq 10^4\epsilon$ (some large constant times ϵ). In fact if we can prove that fact about $q(x)$ then we are done as shown below.

$$\begin{aligned}
 10^4\epsilon &\geq \mathbf{E}[q(x)^2] \\
 &= \sum_{i \neq j} \hat{f}(i)^2 \hat{f}(j)^2 \quad (\text{By Parseval}) \\
 &= [\sum_i \hat{f}(i)^2]^2 - \sum_i \hat{f}(i)^4 \\
 &= (1 - \epsilon)^2 - \sum_i \hat{f}(i)^4 \\
 \Rightarrow \sum_i \hat{f}(i)^4 &\geq 1 - O(\epsilon) \\
 \Rightarrow 1 - O(\epsilon) &\leq \max_i \hat{f}(i)^2 \sum_{i=1}^n \hat{f}(i)^2 \\
 &\leq \max_i \hat{f}(i)^2 \quad (\text{By Parseval})
 \end{aligned}$$

$\therefore \exists i$ such that $\hat{f}(i)^2 \geq 1 - O(\epsilon)$. \square

We now proceed to develop the tools we would need prove the claim made earlier about the $q(x)$.

2 Reasonable Random Variable Principle

Inspired by the SCS Reasonable Person Principle, we have the following definitions for a "reasonable" random variable. Say Y has $\mathbf{E}[Y] = 0$, $\mathbf{E}[Y^2] = 1$. We expect a "reasonable" random variable to satisfy the following :

- $\mathbf{E}[|Y|^3]$ and $\mathbf{E}[|Y|^4]$ should not be too large.
- $\Pr[Y \geq 10^6]$ should not be too large.
- $\Pr[Y \geq 0]$ should be at least some decent value.

So in some sense, we want the random variable to be "well-behaved" (analogously for the reasonable person in SCS). Some examples of very reasonable random variables are illustrated below.

- Y is a random ± 1 bit
- $Y \sim N(0, 1)$, i.e. Y is a Gaussian
- Y is uniform on $[-\sqrt{3}, \sqrt{3}]$
- $Y = \sum_{i=1}^n a_i x_i$ such that $\sum_{i=1}^n a_i^2 = 1$, and each x_i is a random bit. Two special cases are if all the a_i 's are small e.g. $\frac{1}{\sqrt{n}}$ then Y behaves similar to a Gaussian, and if one of the a_i 's is close to 1 then $Y \approx x_i$ (random bit).

Lecture 11: Approximate Arrow's theorem using the hypercontractivity lemma

An example of a random variable which is NOT reasonable is :

$$y = \begin{cases} 0 & \text{with prob } 1 - 2^{-m} \\ 1 & \text{with prob } 2^{-m} \end{cases} \quad \text{where } m \text{ is large}$$

To make sure that $\mathbf{E}[Y] = 0$, $\mathbf{E}[Y^2] = 1$ we need to make slight modifications - subtract a tiny bit and rescale. An easy example of such a random variable as a polynomial is $Y = (1 + x_1)(1 + x_2) \dots (1 + x_m)2^{m/2}$ where the x_i 's are random ± 1 bits.

We now come to the fabled "hypercontractivity" lemma which says that low degree polynomials over random bits are reasonable as defined above.

3 Hypercontractivity Lemma

We outlined in the previous section, the conditions a random variable Y should satisfy to be "reasonable".

Remark 3.1 Assuming $\mathbf{E}[Y^2] = 1$ then if $\mathbf{E}[Y^4] \leq C$ (where C is not too large), then Y has many of the "reasonable" properties.

Remark 3.2 The scale invariant way to say this is $\mathbf{E}[Y^4] \leq C^4 \mathbf{E}[Y^2]^2$.

We now make the notion of "reasonableness" more precise.

Definition 3.3 If Y is a random variable with $\mathbf{E}[Y^4] \leq C^4 \mathbf{E}[Y^2]^2$, we say that Y is $(2, 4, \frac{1}{C})$ -hypercontractive.

We now state the hypercontractivity lemma which says that low degree polynomials over random bits are hypercontractive (the constant C depends on the degree).

Theorem 3.4 (Hypercontractivity Lemma) If $Y = p(x_1, x_2, \dots, x_n)$, where p is a multilinear polynomial of degree d over independent random bits x_i , then Y is $(2, 4, (\frac{1}{\sqrt{3}})^d)$ -hypercontractive, i.e. $\mathbf{E}[Y^4] \leq 9^d \mathbf{E}[Y^2]^2$

Proof: By induction on n .

(Basis) If $n = 0$, then p is a constant. Clearly $d = 0$ and therefore $\mathbf{E}[p^4] = p^4 \leq 9^0 \mathbf{E}[p^2]^2$.

For $n \geq 1$, write $p(x_1, \dots, x_n) = r(x_1, \dots, x_{n-1}) + x_n s(x_1, \dots, x_{n-1})$

Note that $\text{deg}(r) \leq d$ and $\text{deg}(s) \leq d - 1$.

$$\begin{aligned} \mathbf{E}[p^4] &= \mathbf{E}[(r + x_n s)^4] \\ &= \mathbf{E}[r^4 + 4r^3 x_n s + 6r^2 x_n^2 s^2 + 4r x_n^3 s^3 + x_n^4 s^4] \\ &= \mathbf{E}[r^4] + \mathbf{E}[4r^3 x_n s] + \mathbf{E}[6r^2 x_n^2 s^2] + \mathbf{E}[4r x_n^3 s^3] + \mathbf{E}[x_n^4 s^4] \end{aligned}$$

We examine each of the five terms above.

- $\mathbf{E}[r^4] \leq 9^d \mathbf{E}[r^2]^2$ by the induction hypothesis

Lecture 11: Approximate Arrow's theorem using the hypercontractivity lemma

- $\mathbf{E}[4r^3x_n s] = 4\mathbf{E}[r^3s]\mathbf{E}[x_n] = 0$. We used the fact that x_n is independent of r, s and $\mathbf{E}[x_n] = 0$.
- $\mathbf{E}[6r^2x_n^2s^2] = 6\mathbf{E}[r^2s^2]\mathbf{E}[x_n^2]$. We will examine this term below.
- $\mathbf{E}[4rx_n^3s^3] = 4\mathbf{E}[rs^3]\mathbf{E}[x_n^3] = 0$. We again used that x_n is independent of r, s and $\mathbf{E}[x_n^3] = 0$.
- $\mathbf{E}[x_n^4s^4] = \mathbf{E}[x_n^4]\mathbf{E}[s^4] \leq 9^{d-1}\mathbf{E}[s^2]^2\mathbf{E}[x_n^4] \leq 9^dE[s^2]^2$. We used the induction hypothesis here again and the fact that s is a degree $d - 1$ polynomial and $\mathbf{E}[x_n^4] = 1 \leq 9$.

We now get back to the middle term.

$$\begin{aligned} \mathbf{E}[x_n^2]\mathbf{E}[r^2s^2] &= \mathbf{E}[r^2s^2] \\ &\leq \sqrt{\mathbf{E}[r^4]}\sqrt{\mathbf{E}[s^4]} && \text{By the Cauchy-Schwartz inequality} \\ &\leq 3^d\mathbf{E}[r^2]3^{d-1}E[s^2] && \text{By the induction hypothesis} \end{aligned}$$

\therefore , we ultimately get :

$$\begin{aligned} \mathbf{E}[p^4] &\leq 9^d\mathbf{E}[r^2]^2 + 6(3^d\mathbf{E}[r^2]3^{d-1}E[s^2]) + 9^d\mathbf{E}[s^2]^2 \\ &= 9^d\mathbf{E}[r^2]^2 + 9^d(2\mathbf{E}[r^2]\mathbf{E}[s^2]) + 9^d\mathbf{E}[s^2]^2 \\ &\leq 9^d[(\mathbf{E}[r^2] + \mathbf{E}[s^2])^2] \\ &\leq 9^d[(\mathbf{E}[r^2] + 2\mathbf{E}[x_nrs] + \mathbf{E}[x_n^2s^2])] && \text{Using the fact that } \mathbf{E}[x_n] = 0 \text{ and } \mathbf{E}[x_n^2] = 1 \\ &\leq 9^d\mathbf{E}[(r + x_n s)^2]^2 \\ &= 9^d\mathbf{E}[p^2]^2 \end{aligned}$$

□

Remark 3.5 All we used about x_i 's are :

- independence
- $\mathbf{E}[x_i] = 0, \mathbf{E}[x_i^2] = 1, \mathbf{E}[x_i^3] = 0$
- $\mathbf{E}[x_i^4] \leq 9$

4 Proving the claim about $q(x)$

We had made the following claim earlier during the proof of the FKN theorem.

Claim 4.1 If $q(x_1, \dots, x_n)$ is a degree 2 polynomial, such that $|q(x)| \leq 22\sqrt{\epsilon}$ with probability 99%, then $\mathbf{E}[q(x)^2] \leq 10^4\epsilon$

Proof: Assume $\mathbf{E}[q(x)^2] = K\epsilon$ for $K > 10^4$. Since $q^2 \leq 484\epsilon$ 99% of the time, then it must be that conditioned on 1% of the time $\mathbf{E}[q^2] \geq 95K\epsilon$, otherwise $\mathbf{E}[q^2] = 99\% \cdot 484\epsilon + 1\% \cdot 95K\epsilon = (0.95K + 500)\epsilon < K\epsilon$ using $K > 10^4$. But then

$$\mathbf{E}[q^4] \geq 1\%(95K\epsilon)^2 > 90K^2\epsilon^2$$

However, $\mathbf{E}[q^4] \leq 9^2\mathbf{E}[q^2]^2 = 81(K\epsilon)^2 < 90K^2\epsilon^2$ where the first inequality follows from the hypercontractivity lemma. □