80240233: Complexity of Computation | **Lecture 14**
ITCS, Tsinghua Univesity, Fall 2007 | 23 November 2007

Instructor: Andrej Bogdanov | **Notes by:** Andrej Bogdanov

1 The multilinearity test

The idea behind the multilinearity test is simple. The test is based on the observation that if f is a multilinear function, then the restriction of f on any "axis-parallel line" ℓ is an affine function in one variable (a function of the form $ax + b$). An *axis-parallel line* in direction i is a set of points $\ell \subseteq \mathbb{F}^m$ of the form:

$$\ell = \{(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_m) : x_i \in \mathbb{F}\}$$

where $a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_m$ are fixed elements of $\mathbb{F}$ that specify the line ℓ. This suggests the following procedure for testing multilinearity: Choose a random axis-parallel line ℓ and check if the restriction of f along ℓ is an affine function. How can we check that the restriction of f along ℓ is affine? The key property of affine functions that we will use is that the value of any affine function on a line is completely specified by values at any two distinct points; if f is affine, then for any three distinct points $a, b, c \in \ell$, the relation

$$\frac{f(c) - f(b)}{c_i - b_i} = \frac{f(b) - f(a)}{b_i - a_i} = \frac{f(a) - f(c)}{a_i - c_i} \tag{1}$$

must hold. In fact any two of these equalities are sufficient to check, as they imply the third one.

Since $\mathbb{F}$ is exponentially large in m, we cannot check this condition for *all* triplets of points $a, b, c \in \ell$, but we can carry out some random tests.

The multilinearity test. Given access to a function $f : \mathbb{F}^m \to \mathbb{F}$:

1. Choose a random axis-parallel line ℓ over $\mathbb{F}^m$.
2. Choose a random triple of distinct points $a, b, c \in \ell$.
3. If the values $f(a), f(b), f(c)$ satisfy (1), accept, otherwise reject.

By our discussion, if f is a multilinear function, then the multilinearity test accepts f with probability 1. Now we want to show that if f is far from multilinear, then the test rejects f with some nonnegligible probability. This probability can be increased by repeating the test several times.

There are several proofs of this fact, none of them easy. We will show the proof by Feige, Goldwasser, Lovasz, Safra, and Szegedy.

Theorem 1. *Let $\delta > 0$ be an arbitrary constant. Assume $|\mathbb{F}| \gg m/\delta$. Suppose that f is δ-far from multilinear. Then the multilinearity test rejects f with probability at least $\Omega(\delta/m)$.*

Let's introduce some notation. Let $\delta(f)$ denote the distance between f and the multilinear function closest to f, that is

$$\Delta(f) = \min_L \Pr_{x \sim \mathbf{F}^m}[f(x) \neq L(x)]$$

where L ranges over all multilinear functions over $\mathbb{F}^m$. Let $\tau(f)$ denote the probability that the multilinearity test rejects f. We say a triple of points (a, b, c) along some axis-parallel line is *f-affine* if the values $f(a), f(b), f(c)$ satisfy (1). Then the rejection probability of the test can be written as

$$\tau(f) = \mathrm{E}_\ell\big[\Pr_{a,b,c \sim \ell}[(a, b, c) \text{ is not } f\text{-affine}]\big]$$

where ℓ is chosen at random from the set of all axis-parallel lines in $\mathbb{F}^m$.

The analysis will split into two cases, depending on whether f is somewhat far or very far from multilinear.

Claim 2. *Suppose $|\mathbb{F}| \gg m/\delta$ and $\delta \leq \delta(f) \leq 9/10$. Then $\tau(f) = \Omega(\delta/m)$.*

Claim 3. *Suppose $|\mathbb{F}| \gg m$ and $\delta(f) > 9/10$. Then $\tau(f) = \Omega(1/m \cdot (1 - 1/|\mathbb{F}|)^{m-1})$.*

Putting the two claims together gives Theorem 1. Before we start the proofs we state two useful facts about affine functions that follow directly from (1):

Fact 4. *If (a, b, c) and (a, b, c') are both f-affine, then (a, c, c') is also f-affine.*

Fact 5. *If (a, b, c) is both f-affine and g-affine and $f(a) = g(a)$ and $f(b) = g(b)$, then $f(c) = g(c)$.*

2 Proof of Claim 2

Let L denote the closest multilinear function to f and let S be the set of points $x \in \mathbb{F}^m$ on which f and L match, that is

$$S = \{x \in \mathbb{F}^m : f(x) = L(x)\}.$$

Let N be the number of points among a, b and c that fall inside S. First we will show that $\Pr[N = 1 \text{ or } N = 2] = \Omega(\delta/m)$. Then we will show that the probability of a, b, c not being f-affine is roughly the same as the probability that $N = 1$ or $N = 2$.

To argue the first part, it is sufficient to show that

$$\Pr_{a,b}[a \in S \text{ and } b \notin S] = \Omega(\delta/m).$$

If a and b were independent, then this would be trivial, and in fact the probability would be $\delta(f)(1 - \delta(f)) = \Omega(\delta)$. However, they are not independent. Even so, we can think of a and b as being generated from the following distribution:

1. Choose two independent random points $p, q \sim \mathbb{F}^m$.
2. Choose a random index $i \sim \{1, \dots, m\}$.
3. Set $a = (p_1, \dots, p_{i-1}, p_i, q_{i+1}, \dots, q_m)$ and $b = (p_1, \dots, p_{i-1}, q_i, q_{i+1}, \dots, q_m)$.

Now observe that

$$\Pr_{a,b}[a \in S \text{ and } b \notin S] \geq \Pr_i[a \in S \text{ and } b \notin S \mid p \in S \text{ and } q \notin S] \cdot \Pr_{p,q}[p \in S \text{ and } q \notin S].$$

The second probability is $\Omega(1)$ since p and q are independent. The first probability is at least $1/m$: If $p \in S$ and $q \notin S$, there must be some index i for which $(p_1, \dots, p_{i-1}, p_i, q_{i+1}, \dots, q_m) \in S$ but $(p_1, \dots, p_{i-1}, q_i, q_{i+1}, \dots, q_m) \notin S$ and a, b "hit" this i with probability at least $1/m$.

Now we show that

$$\tau(f) \geq \Pr[N = 1 \text{ or } N = 2] - 3/|\mathbb{F}| \tag{2}$$

By the above calculation, this together with the fact $|\mathbb{F}| \gg m/\delta$ shows that $\tau(f) \geq \Omega(\delta/m)$ establishing the claim.

To show (2) we look at each line ℓ separately. We fix a line ℓ and consider a random triple (a, b, c) along ℓ. Let

$$S_\ell = S \cap \ell = \{x \in \ell : f(x) = L(x)\} \qquad \overline{S}_\ell = \{x \in \ell : f(x) \neq L(x)\}.$$

Also let N_ℓ be the number of points among a, b, c that fall inside S_ℓ.

- If $N_\ell = 2$, then the probability that (a, b, c) is f-affine is zero. Since (a, b, c) is always L-affine, by Fact 5, if f and L match at any two points, then they must also match at the third one.
- If $N_\ell = 1$, then it is possible that (a, b, c) could be f-affine, but we will show that this cannot happen "a lot". Suppose for instance that $a, b \notin S_\ell$. Then there is at most one $c \in S_\ell$ such that (a, b, c) is not f-affine: Suppose on the contrary that there are $c \neq c'$, $c, c \in S_\ell$ such that both (a, b, c) and (a, b, c') are f-affine. Then by Fact 4, (a, c, c') is also f-affine, but since f and L match at both c and c', by Fact 5 they must also match at a, a contradiction.

  In other words, if we fix $a, b \notin S_\ell$, and choose a random c from S_ℓ, the probability that (a, b, c) is f-affine is at most $1/|S_\ell|$.

By this discussion, we have that

$$\begin{aligned}
\Pr_{a,b,c\sim\ell}[(a, b, c) \text{ is not } f\text{-affine}] &\geq \Pr[(a, b, c) \text{ is not } f\text{-affine} \mid N_\ell = 2] \cdot \Pr[N_\ell = 2] \\
&\quad + \Pr[(a, b, c) \text{ is not } f\text{-affine} \mid N_\ell = 1] \cdot \Pr[N_\ell = 1] \\
&\geq 1 \cdot \frac{3 \cdot |S_\ell|^2 \cdot |\overline{S}_\ell|}{|\mathbb{F}|^3} + \left(1 - \frac{1}{|S_\ell|}\right) \cdot \frac{3 \cdot |S_\ell| \cdot |\overline{S}_\ell|^2}{|\mathbb{F}|^3} \\
&\geq \frac{3 \cdot |S_\ell|^2 \cdot |\overline{S}_\ell|}{|\mathbb{F}|^3} + \frac{3 \cdot |S_\ell| \cdot |\overline{S}_\ell|^2}{|\mathbb{F}|^3} - \frac{3}{|\mathbb{F}|} \\
&= \Pr[N_\ell = 1 \text{ or } N_\ell = 2] - 3/|\mathbb{F}|.
\end{aligned}$$

Taking expectation over axis-parallel ℓ we obtain (2).

3 Proof of Claim 3

We prove this claim by induction on m. When $m = 1$, we show that if $\delta(f) > 9/10$, then $\tau(f) > 9/10$. Suppose that $\tau(f) \leq 9/10$, so that

$$\Pr[(a, b, c) \text{ is } f\text{-affine}] > 1/10$$

where the randomness is over all distinct triples $a, b, c \in \mathbb{F}$. In particular we can fix a choice of a and b such that

$$\Pr_{c \neq a,b}[(a, b, c) \text{ is } f\text{-affine}] > 1/10.$$

Let L be the unique linear function such that $L(a) = f(a)$ and $L(b) = f(b)$. By Fact 5 we then have that $L(c) = f(c)$ for at least 1/10 fraction of points c, so f is 1/10-close to a linear function.

We now do the inductive step. For every $a_1 \in \mathbb{F}$, we define f_{a_1} to be the $m - 1$-variate function obtained when we restrict the first coordinate of f to a_1, that is

$$f_{a_1}(y) = f(a_1, y), \text{ where } a_1 \in \mathbb{F}, y \in \mathbb{F}^{m-1}.$$

Let's assume that $\delta(f) > 9/10$. We try to understand what could be the reason that f could be so far from multilinear. One option is that many of the restrictions f_{a_1}, where a_1 ranges over $\mathbb{F}$ are themselves far from multilinear. But it could also be the case that all restrictions f_{a_1} are multilinear, but they do not piece together into a single multilinear function f. Then we expect the test to detect this when the line ℓ is axis-parallel in direction 1.

Even for a fixed a_1, it may be reasonable to expect that either the function f_{a_1} is itself far from multilinear, or that the test "catches" f on some triple of the form $(a_1, y), (b_1, y), (c_1, y)$. Let τ_{a_1} denote this probability:

$$\tau_{a_1} = \Pr_{b_1,c_1,y}[((a_1, y), (b_1, y), (c_1, y)) \text{ is not } \mathbb{F}\text{-affine}]$$

where a_1, b_1, c_1 are conditioned to be all distinct. We would like to say that for all a_1, if $\delta(f)$ is large, then either $\delta(f_{a_1})$ is large or τ_{a_1} is large. This is not quite true, but a slightly weaker statement of a similar flavor can be shown.

Claim 6. *For every pair of distinct $s_1, t_1 \in \mathbb{F}$,*

$$\delta(f) \leq \delta(f_{s_1}) + \delta(f_{t_1}) + \tau_{s_1} + \tau_{t_1} + 6/|\mathbb{F}|.$$

Let us now assume that $\delta(f) > 9/10$. By this claim there can be at most one value $s_1 \in \mathbb{F}$ for which both $\delta(f_{s_1}) \leq 1/10$ and $\tau_{s_1} \leq 1/3$. Fix this value s_1 (if it doesn't exist, choose s_1 to be arbitrary).

Now consider a random triple (a, b, c) chosen by the multilinearity test. With probability $1 - 1/|\mathbb{F}|$, $a_1 \neq s_1$. Conditioned on this event, what is the probability of the test rejecting? We divide in three cases depending on the value a_1:

- If $\delta(f_{a_1}) \leq 1/10$, then it must be that $\tau_{a_1} > 1/3$. Conditioned on this value of a_1 and on ℓ being axis-parallel to direction 1, the probability of the test rejecting is 1/3, so the total probability of the test rejecting is at most $1/3m$.

- If $1/10 < \delta(f_{a_1}) \leq 9/10$, then conditioned on ℓ *not* being axis-parallel to direction 1, by Claim 2, the test rejects with probability $\Omega(1/(m-1))$.
- If $\delta(f_{a_1}) > 1/10$, then conditioned on ℓ *not* being axis-parallel to direction 1, by inductive hypothesis the test rejects with probability $\Omega(1/(m-1) \cdot (1-1/|\mathbb{F}|)^{m-2})$.

It follows that

$$\begin{aligned}\tau(f) &\geq \left(1-\frac{1}{|\mathbb{F}|}\right) \cdot \min\left(\frac{1}{3m}, \ \frac{m-1}{m} \cdot \Omega\Big(\frac{1}{m-1}\Big), \ \frac{m-1}{m} \cdot \Omega\Big(\frac{1}{m-1} \cdot \Big(1-\frac{1}{|\mathbb{F}|}\Big)^{m-2}\Big)\right) \\ &= \Omega\left(\frac{1}{m} \cdot \Big(1-\frac{1}{|\mathbb{F}|}\Big)^{m-1}\right)\end{aligned}$$

since the last term dominates the minimum. Since $|\mathbb{F}| \gg m$, the whole expression is bounded by $\Omega(1/m)$. It only remains to prove Claim 6.

Proof of Claim 6. Let L_{s_1}, L_{t_1} be the closest multilinear functions to f_{s_1} and f_{t_1}, respectively. These are functions in $m-1$ variables. We extend L_{s_1} and L_{t_1} to a multilinear function L in m variables via the formula

$$L(x_1, y) = L_{s_1}(y) + \frac{x_1 - s_1}{t_1 - s_1}(L_{t_1}(y) - L_{s_1}(y)).$$

Then $\delta(f) \leq \Pr_{x_1,y}[f(x_1,y) \neq L(x_1,y)]$. To upper bound this probability, we observe the following simple fact:

> Suppose $r_1 \in \mathbb{F}$ is such that s_1, t_1, x_1 and r_1 are all distinct. If $((s_1, y), (x_1, y), (r_1, y))$ is f-affine, $((t_1, y), (x_1, y), (r_1, y))$ is f-affine, $f_{s_1}(y) = L_{s_1}(y)$ and $f_{a_2}(y) = L_{a_2}(y)$ then $f(x_1, y) = L(x_1, y)$.

To see this, first note that if $((s_1, y), (x_1, y), (r_1, y))$ and $((t_1, y), (x_1, y), (r_1, y))$ are both f-affine, then by Fact 4 so is $((s_1, y), (t_1, y), (x_1, y))$. Since L is affine along the line ℓ specified by y by Fact 5 and L matches f at (s_1, y) and (t_1, y), it must be that $f(x_1, y) = L(x_1, y)$.

Therefore we have that

$$\begin{aligned}\Pr_{x_1,y}[f(x_1,y) \neq L(x_1,y)] \leq\ & \Pr_{x_1,r_1,y}[((s_1,y),(x_1,y),(r_1,y)) \text{ is not } f\text{-affine}] \\ &+ \Pr_{x_1,r_1,y}[((t_1,y),(x_1,y),(r_1,y)) \text{ is not } f\text{-affine}] \\ &+ \Pr_y[f_{s_1}(y) \neq L_{s_1}(y)] + \Pr_y[f_{t_1}(y) \neq L_{t_1}(y)] \\ &+ \Pr_{x_1,r_1}[x_1, r_1, s_1, t_1 \text{ are not all distinct}]\end{aligned}$$

where the first two terms are conditioned on x_1, r_1, s_1, t_1 being all distinct. The probabilities in this sum are $\tau_{s_1}, \tau_{t_1}, \delta(f_a), \delta(f_b)$ and $\leq 6/|\mathbb{F}|$, respectively. ☐