Correlation Based Testing

Elchanan Mossel   Per Austrin

UC Berkeley + Weizmann   U. Toronto

August 13, 2011

Testing predicates and PCPs

Tests are key in proving hardness of approximation

Long code analysis

Predicate Q

Joint distribution over inputs $X_1, \ldots, X_k$

Which functions $f_1, ..., f_k$ satisfy $Q(f_1(X_1), \ldots, f_k(X_k))$ with good probability?

Algebraic and Geometric Tests

Algebraic tests

Test additive (or algebraic) properties over finite fields

Closely related to additive combinatorics

Often invariant under the linear group

Algebraic and Geometric Tests

Algebraic tests

Test additive (or algebraic) properties over finite fields

Closely related to additive combinatorics

Often invariant under the linear group

Geometric tests

Often relate to $\mathbb{R}$-geometric questions

Not invariant under coordinate systems

Algebraic and Geometric Tests

Algebraic tests

Test additive (or algebraic) properties over finite fields

Closely related to additive combinatorics

Often invariant under the linear group

Geometric tests

Often relate to $\mathbb{R}$-geometric questions

Not invariant under coordinate systems

Motivation

What is the relation between the two types of tests?

Is there a unified framework to study both?

Algebraic test 1: BLR test / Roth Theorem (93 / 53)

Q: Is $f : \{-1,1\}^n \to \{-1,1\}^n$ linear over F_2 ?

$$\text{Distribution: } \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} X_1^1 & \dots & X_1^j & \dots & X_1^n \\ X_2^1 & \dots & X_2^j & \dots & X_2^n \\ X_3^1 & \dots & X_3^j & \dots & X_3^n \end{pmatrix}$$

Independent columns

$(X_1^j, X_2^j, X_3^j) \in \{-1,1\}^3$ uniform with $X_1^j X_2^j = X_3^j$

Test: $f(X_1)f(X_2)f(X_3)$ (parity predicate)

Algebraic test 1: BLR test / Roth Theorem (93 / 53)

Q: Is $f : \{-1,1\}^n \to \{-1,1\}^n$ linear over F_2 ?

$$\text{Distribution: } \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} X_1^1 & \dots & X_1^j & \dots & X_1^n \\ X_2^1 & \dots & X_2^j & \dots & X_2^n \\ X_3^1 & \dots & X_3^j & \dots & X_3^n \end{pmatrix}$$

Independent columns

$(X_1^j, X_2^j, X_3^j) \in \{-1,1\}^3$ uniform with $X_1^j X_2^j = X_3^j$

Test: $f(X_1)f(X_2)f(X_3)$ (parity predicate)

Equivalent formulation

μ uniform over $(X_1, X_2, X_3) \in \{-1,1\}^3$ with $X_1 X_2 = X_3$

Sample: (X_1, X_2, X_3) from μ^n and test: $f(X_1)f(X_2)f(X_3)$

Algebraic test 1: BLR/Roth test and Fourier Analysis

BLR test / Roth Theorem

Q: Is $f : \{-1,1\}^n \to \{-1,1\}$ linear over F_2 ?

Distribution: μ uniform over $(X_1, X_2, X_3) \in \{-1,1\}^3$ with $X_1 X_2 = X_3$

Sample: (X_1, X_2, X_3) from μ^n

Test: $f(X_1)f(X_2)f(X_3)$ (the parity predicate)

Analysis

$$\mathbb{E}[f(X_1)f(X_2)f(X_3)] = \sum_S \hat{f}^3(S) = 1 \text{ if } f \text{ linear}$$

$$|\mathbb{E}[f(X_1)f(X_2)f(X_3)]| = |\sum_S \hat{f}^3(S)| \leq \max_S |\hat{f}(S)|$$

Conclusion

ϵ bias in passing $\implies$ ϵ-correlation with an affine function

Algebraic test 2: Gowers norm tests

Gowers test

Q: Do $f_1, \ldots, f_k$ distinguish arithmetic progressions from the uniform measure ?

Distribution: μ uniform over arithmetic progressions $(X_1, \ldots, X_k) \in Z_p$

Sample: $(X_1, \ldots, X_k)$ from μ^n

Test: $\mathbb{E}[\prod_{i=1}^k f_i]$

Algebraic test 2: Gowers norm tests)

Gowers test

Q: Do $f_1, \ldots, f_k$ distinguish arithmetic progressions from the uniform measure ?

Distribution: μ uniform over arithmetic progressions $(X_1, \ldots, X_k) \in Z_p$

Sample: $(X_1, \ldots, X_k)$ from μ^n

Test: $\mathbb{E}[\prod_{i=1}^{k} f_i]$

Analysis (Gowers, 01):

$$|\mathbb{E}[\prod_{i=1}^{k} f_i]| \leq \min_{i=1}^{k} \|f_i\|_{U^{k-1}}$$

Conclusion

Functions with low Gowers norms cannot distinguish arithmetic progressions

Geometric test 1: Khot's Gaussian test

Khot's test

Q: Is $\max_i I_i(f)$ large?

Distribution: μ satisfies $\mu[X_1 X_2] = \rho$ over $\{-1, 1\}^2$

Sample: (X_1, X_2) from μ^n

Test: $f(X_1)f(X_2)$ (the (in)equality predicate)

Geometric test 1: Khot's Gaussian test

Khot's test

Q: Is $\max_i I_i(f)$ large?

Distribution: μ satisfies $\mu[X_1 X_2] = \rho$ over $\{-1,1\}^2$

Sample: (X_1, X_2) from μ^n

Test: $f(X_1)f(X_2)$ (the (in)equality predicate)

Analysis - "Majority is Stablest" (KKMO-04; MOO-05)

Let $g : \mathbb{R} \to \{-1,1\}$ with $\mathbb{E}[g] = \mathbb{E}[f]$ and g is increasing

Let $(N_1, N_2) \sim N(0,1)$ with $\mathbb{E}[N_1 N_2] = \rho$

If $\mathbb{E}[f(X_1)f(X_2)] > \mathbb{E}[g(N_1)g(N_2)]$

then $\max_i I_i(f)$ large

Geometric test 2: M's Gaussian test

Test

Q: Is $\max(f_j(i) : 1 \leq j \leq k, 1 \leq i \leq n)$ large?

μ: full support distribution on $\{-1, 1\}^k$

Sample: $(X_1, \ldots, X_k)$ from μ^n

Test: $P(f_1(X_1), \ldots, f_k(X_k))$ (P a general predicate)

Geometric test 2: M's Gaussian test

Test

Q: Is $\max(f_j(i) : 1 \leq j \leq k, 1 \leq i \leq n)$ large?

μ: full support distribution on $\{-1,1\}^k$

Sample: $(X_1, \dots, X_k)$ from μ^n

Test: $P(f_1(X_1), \dots, f_k(X_k))$ (P a general predicate)

Analysis - "Gaussian bounds" M-08

Let $g_1, \dots, g_k : \mathbb{R}^n \to \{-1,1\}$ with $\mathbb{E}[g_i] = \mathbb{E}[f_i]$

Let $(N_1^i, \dots, N_k^i)$ have the same first and second moments as μ

If $\mathbb{E}[P(f(X_1), \dots, f(X_k))] > \max_g \mathbb{E}[P(g_1(N_1), \dots, g_k(N_k))]$

then $\max f_j(i)$ large

Are the two approaches really different?

Distributions

Arithmetic tests: small support uniform on arithmetic structures

Geometric tests: general product distributions with full support

Conclusions

Arithmetic tests: correlation with arithmetic structure

Geometric tests: high influence variables

Common setup of two approaches

Distributions

Arithmetic tests: small support uniform on arithmetic structures

Geometric tests: general product distributions with full support

Conclusions

Arithmetic tests: correlation with arithmetic structure

Geometric tests: high influence variables

Common setup?

Pairwise independent distributions

w / w.o. full support

Question

What do the two approaches gives?

Håstad's Fourier test

The best of all worlds: Håstad's test (97)

μ: satisfies $\mu[x_1 x_2 x_3] = \rho$

Sample: (X_1, X_2, X_3) from μ^n

Test: $f(X_1)f(X_2)f(X_3)$ (the parity predicate)

Analysis

$$\mathbb{E}[f(X_1)f(X_2)f(X_3)] - \prod_{i=1}^{3} \mathbb{E}[f_i] = \sum_{S \neq \emptyset} \rho^{|S|} \hat{f}^3(S)$$

If large then correlated with a function of a small number of variables!

Very useful in PCP proof

Question

Can this be extended?

An algebraic extension of Håstad's test

Samorodnitsky and Trevisan

μ: $X_S = \prod_{i \in S} Y_i$ for $S \subset [k]$ where $(Y_1, \ldots, Y_k) \sim_{\text{Unif}} \{-1, 1\}^k$.

Sample: $(X_S : S \subset [k])$ from μ^n

Test: $\mathbb{E}[\prod_{S \subset [k]} f_S(x_S)] - \prod_{S \subset [k]} \mathbb{E}[f_S]$

ST Analysis (via Gowers norms, 05)

$$|\mathbb{E}[\prod_{S \subset [k]} f_S(x_S)] - \prod_{S \subset [k]} \mathbb{E}[f_S]| \le O\left(\sqrt{\max_{S \subset [k]} \max_{1 \le j \le n} I_j(f_S)}\right)$$

Note

Weaker than Håstad's test conclusion

Gives UCG hardness approximation resistance of predicate above

A geometric extension of Håstad's test

M-08

μ: A general pairwise independent distribution with full support

Sample: $X_1, \ldots, X_k$ from μ^n

Test: $\mathbb{E}[\prod_{i=1}^k f_i(X_i)] - \prod_{i=1}^k \mathbb{E}[f_i]$

Analysis (via Gaussian bounds)

$$\mathbb{E}[\prod_{i=1}^k f_i(X_i)] - \prod_{i=1}^k \mathbb{E}[f_i] \to 0 \text{ as } \max_{i,j} I_j(f_i) \to 0.$$

Note

Still only influences

Used in Austrin-M-09: pairwise independent predicates are approximation resistant (Also M-Håstad-10)

The ultimate extension of Håstad's test?

Tests - What we can hope for?

μ: A general pairwise independent distribution (with full support?)

Sample: $X_1, \ldots, X_k$ from μ^n

Test: $\mathbb{E}[\prod_{i=1}^k f_i(X_i)] - \prod_{i=1}^k \mathbb{E}[f_i]$

If pass the test then one of f_i is correlated with a function of a small number of variables

Hardness - What can we hope for?

NP-hardness of all predicates whose support supports a pairwise independent distribution

Results of Austrin-M-11

Thm: Bounded degree polynomials

If μ is a general pairwise independent distribution

If f_i are degree d polynomials:

$$|\mathbb{E}[\prod_{i=1}^{k} f_i(X_i)]| \leq C^d \|\widehat{f_1}\|_\infty \prod_{i=2}^{k} \|f_i\|_2$$

Corollary (Hatami): Noisy additive predicates

μ is given by k distinct noisy linear forms $X_1, \ldots, X_k$

If f_i are all bounded by 1:

$$|\mathbb{E}[\prod_{i=1}^{k} f_i(X_i)]| \leq H(\|\widehat{f_1}\|_\infty), \quad \lim_{x \to 0} H(x) = 0.$$

Noisy linear predicates

Example

Take μ to be the standard pairwise independent construction

$$(x_S = \prod_{i \in S} y_i : \emptyset \neq S \subset [k], y \in \{-1,1\}^r)$$

ν - take μ and flip each bit independently with probability ϵ or

ν - take μ and with probability ϵ flip all bits to a uniform random string

Challenge

Use result to prove approximation resistance of predicate

Proof of Hatami's corollary

Corollary (Hatami): Noisy additive predicates

μ is given by k distinct noisy linear forms $X_1, \ldots, X_k$; f_i bdd by 1

$$|\mathbb{E}[\prod_{i=1}^{k} f_i(X_i)]| \leq H(\|\widehat{f_1}\|_\infty), \quad \lim_{x \to 0} H(x) = 0.$$

Proof Sketch

Since predicate is noisy, may assume exists d so that $\|f_1^{>d}\|_2 < \epsilon/2$

By CS: $\mathbb{E}[f_1^{>d} f_2 \cdots f_k] \leq \|f_1^{>d}\|_2 \leq \epsilon/2$

By Gowers-CS $\mathbb{E}[f_1^{\leq d} f_2 \cdots f_k] \leq \|f_1^{\leq d}\|_{U(k-1)}$

By Theorem $\|f_1^{\leq d}\|_{U(k-1)} \to 0$ as $\|\widehat{f_1}\|_\infty \to 0$

Proof of Theorem

Thm: Bounded degree polynomials

If μ is a general pairwise independent distribution

If the sum of degree of f_i is at most D then

$$|\mathbb{E}[\prod_{i=1}^{k} f_i(X_i)]| \leq C^D \|\widehat{f_1}\|_\infty \prod_{i=2}^{k} \|f_i\|_2$$

Proof idea

Prove by induction on the number of variables

Use pairwise independence to show "second order terms" vanish

case $n = 0$ is fine

Proof of Theorem continued

Induction step sketch

Write $f_i = X_1^i g_i + h_i$ where g_i, h_i functions of $n-1$ variables:

$$\mathbb{E}[\prod_i f_i] = \sum_T \mathbb{E}[\prod_{i \in T} X_1^i] \mathbb{E}[\prod_{i \in T} g_i] \mathbb{E}[\prod_{i \notin T} h_i]$$

Proof of Theorem continued

Induction step sketch

Write $f_i = X_1^i g_i + h_i$ where g_i, h_i functions of $n-1$ variables:

$$\mathbb{E}[\prod_i f_i] = \sum_T \mathbb{E}[\prod_{i \in T} X_1^i] \mathbb{E}[\prod_{i \in T} g_i] \mathbb{E}[\prod_{i \notin T} h_i]$$

Pairwise ind. implies terms with $|T| = 1$ or $|T| = 2$ vanish:

$$|\mathbb{E}[\prod_i f_i]| \leq C^d \delta \prod_{i=2}^{k} \|h_i\|_2 + 2^k C^{d-3} \delta \max_{|T| \geq 3} \prod_{1 \neq i \in T} \|g_i\|_2 \prod_{1 \neq i \notin T} \|h_i\|_2$$

Proof of Theorem continued

Induction step sketch

Write $f_i = X_1^i g_i + h_i$ where g_i, h_i functions of $n-1$ variables:

$$\mathbb{E}[\prod_i f_i] = \sum_T \mathbb{E}[\prod_{i \in T} X_1^i]\mathbb{E}[\prod_{i \in T} g_i]\mathbb{E}[\prod_{i \notin T} h_i]$$

Pairwise ind. implies terms with $|T| = 1$ or $|T| = 2$ vanish:

$$|\mathbb{E}[\prod_i f_i]| \le C^d \delta \prod_{i=2}^{k} \|h_i\|_2 + 2^k C^{d-3} \delta \max_{|T| \ge 3} \prod_{1 \ne i \in T} \|g_i\|_2 \prod_{1 \ne i \notin T} \|h_i\|_2$$

Suffices to show that for every $T \subset \{2, \dots, k\}$ of size at least 2:

$$\prod_{i \in T} \|g_i\|_2 + 2^k C^{-3} \prod_{i \notin T} \|h_i\|_2 \le \prod_{i \in T} \|f_i\|_2$$

Proof of Theorem concluded

Suffices to show that for every $T \subset \{2, \ldots, k\}$ of size at least 2:

$$\prod_{i \in T} \|g_i\|_2 + 2^k C^{-3} \prod_{i \in T} \|h_i\|_2 \leq \prod_{i \notin T} \|f_i\|_2$$

Calculus: If $r \geq 2, \exists \epsilon(r)$ s.t. for all $a_i \geq 0, b_i \geq 0, 1 \leq i \leq r$:

$$\prod_{i=1}^{r} a_i + \epsilon \prod_{i=1}^{r} b_i \leq \prod_{i=1}^{r} \sqrt{a_i^2 + b_i^2}$$

A simpler calculus exercise

If $r \geq 2, \exists \epsilon(r)$ s.t. for all $x_i \geq 0, 1 \leq i \leq r$ it holds that:

$$1 + \epsilon \prod_{i=1}^{r} x_i \leq \prod_{i=1}^{r} \sqrt{1 + x_i^2}$$

Question

Suppose f has all small coefficients and is of small degree.

Is $D(f(X_1), \ldots, f(X_k)) \sim \prod D(f(X_i))$?

No!

Example

$$f(x) = (x_1 + 1)\frac{1}{n^{1/2}} \sum_{i=2}^{n+1} x_i, \quad X_3^i X_2^i X_1^i = -1$$

$(f(X_1), f(X_2), f(X_3))$ has 0 in at least one coordinate.

This is not true for product distribution.

Open problem 1

Hardness of Approximation

Prove NP-approximation resistance of pairwise independent predicates

Open Problem 2

problem

For which distributions the maximum of the Gaussian test is obtained in finite dimensions?

Open Problem 3

Vector linearity

Testing if $f : F_2^n \to F_2^n$ is linear

Pick X_1, X_2 with $X_3 = X_1 \oplus X_2$

Test if $f(X_1)f(X_2) = f(X_3)$

problem

If f passes the test with probability ϵ how correlated is it with linear functions?

Best results due to Sanders (almost polynomial in ϵ)

Thanks!

Thank you!

Any questions?