

Homomorphism Testing

Shpilka & Wigderson, 2006

Daniel Shahaf

Tel-Aviv University

January 2009

Table of Contents

Linearity
Testing

Preliminaries

The Problem

New Answers

Summary

1 Preliminaries

2 The Problem

3 New Answers

- 1 Preliminaries
 - Algebra
 - Definitions

Definition

A **group** is a set G and a binary function $+$: $G^2 \rightarrow G$ such that '+' has —

- identity: $g + 0 = 0 + g = g$.
- inverses: $g + g' = g' + g = 0$.
- associativity: $(g + g') + g'' = g + (g' + g'')$.

If '+' is commutative, then $(G, +)$ is called **Abelian**.

Definition

A **group** is a set G and a binary function $+$: $G^2 \rightarrow G$ such that '+' has —

- identity: $g + 0 = 0 + g = g$.
- inverses: $g + g' = g' + g = 0$.
- associativity: $(g + g') + g'' = g + (g' + g'')$.

If '+' is commutative, then $(G, +)$ is called **Abelian**.

Notational convention

In this work, G is a group written additively and H is a group written multiplicatively.

(Affine) Homomorphisms

Linearity
Testing

Preliminaries

Algebra

Definitions

The Problem

New Answers

Summary

Definition

A **homomorphism** from a group G to a group H is a function $f : G \rightarrow H$ such that $f(g_1) \cdot f(g_2) = f(g_1 + g_2)$ for any $g_1, g_2 \in G$.

Definition

If $f' = h \cdot f$ and $f \in \text{Hom}(G, H)$, then f' is an **affine homomorphism**.

- Is $f \cdot h$ **necessarily** an affine homomorphism?

(Affine) Homomorphisms

Definition

A **homomorphism** from a group G to a group H is a function $f : G \rightarrow H$ such that $f(g_1) \cdot f(g_2) = f(g_1 + g_2)$ for any $g_1, g_2 \in G$.

Definition

If $f' = h \cdot f$ and $f \in \text{Hom}(G, H)$, then f' is an **affine homomorphism**.

- Is $f \cdot h$ **necessarily** an affine homomorphism?

Lemma (Exercise)

A function f' is an affine homomorphism \iff the function $f'' := x \mapsto f'(0)^{-1} \cdot f'(x)$ is a homomorphism.

Definition

The (normalized) **Hamming distance** of two functions $f, g : D \rightarrow R$ is the fraction of D on which they disagree:

$$\text{dist}(f, g) := \text{Prob}_{x \leftarrow D} [f(x) \neq g(x)].$$

Definition

We say that a function $f : G \rightarrow H$ is **ϵ -close to linearity** if $\text{dist}(f, \text{Hom}(G, H)) \leq \epsilon$. Otherwise f is said to be **ϵ -far**.

- 2 The Problem
- Motivation
 - Formalization
 - Past Answers

Linear Functions

Linearity
Testing

Preliminaries

The Problem

Motivation

Formalization

Past Answers

New Answers

Summary

Many problems can be presented or solved in terms of questions about functions and linearity:

- Numerical computations.
- The PCP theorem.
- Dictator functions.

Linear Functions

Linearity
Testing

Preliminaries

The Problem

Motivation

Formalization

Past Answers

New Answers

Summary

Many problems can be presented or solved in terms of questions about functions and linearity:

- Numerical computations.
- The PCP theorem.
- Dictator functions.

Questions

- **Is** this function linear?
- How little **need** I change this function to make it linear?
- **How** do I change this function to make it linear?

Defining The Problem

Linearity
Testing

Preliminaries

The Problem

Motivation

Formalization

Past Answers

New Answers

Summary

Question

Given a function $f : G \rightarrow H$, is f a homomorphism, or ϵ -far from linearity?

Defining The Problem

Linearity
Testing

Preliminaries

The Problem

Motivation

Formalization

Past Answers

New Answers

Summary

Question

Given a function $f : G \rightarrow H$, is f a homomorphism, or ϵ -far from linearity?

- Yet another gap problem.
- Oracle-access model.

Defining The Solutions

We will seek probabilistic tests that classify **functions** with high **accuracy**.

- The tests will accept all **valid inputs** with probability 1.
- Need to show that they reject **invalid** (ϵ -far) **inputs** with high probability.

Defining The Solutions

We will seek probabilistic tests that classify **functions** with high **accuracy**.

- The tests will accept all **valid inputs** with probability 1.
- Need to show that they reject **invalid** (ϵ -far) **inputs** with high probability.
- A (δ, ϵ) -test for a **property** $P \subset \mathcal{D}$ is a algorithm(\mathcal{D}) that, on **input** that is ϵ -far from having P , accepts with probability at most δ .

Defining The Solutions

We will seek probabilistic tests that classify **functions** with high **accuracy**.

- The tests will accept all **valid inputs** with probability 1.
- Need to show that they reject **invalid** (ϵ -far) **inputs** with high probability.
- A (δ, ϵ) -test for a **property** $P \subset \mathcal{D}$ is an algorithm(\mathcal{D}) that, on **input** that is ϵ -far from having P , accepts with probability at most δ .

Properties of tests

- Query complexity.
- Probability–distance relation.
- Asymptotical running time.
- Number of random bits.

The BLR Test

Linearity Testing

Preliminaries

The Problem

Motivation

Formalization

Past Answers

New Answers

Summary

Linearity test

- Pick at random $x, y \in G$.
- Accept iff $f(x) \cdot f(y) = f(x + y)$.

The BLR Test

Linearity test

- Pick at random $x, y \in G$.
- Accept iff $f(x) \cdot f(y) = f(x + y)$.

Analysis

- If $f \in \text{Hom}(G, H)$, the test passes with probability 1.
- If the test passes with high probability, then f is ϵ -close.
 - The BLR test is a $(\delta, \underbrace{\delta/3 + O(\delta^2)}_{\epsilon})$ test.

The BLR Test

Linearity test

- Pick at random $x, y \in G$.
- Accept iff $f(x) \cdot f(y) = f(x + y)$.

Analysis

- If $f \in \text{Hom}(G, H)$, the test passes with probability 1.
- If the test passes with high probability, then f is ϵ -close.
 - The BLR test is a $(\delta, \underbrace{\delta/3 + O(\delta^2)}_{\epsilon})$ test.

Exercise

Extend the BLR linearity test to **affine** homomorphisms.

The BLR test picked random $x, y \in G$.

Generalization (Graph tests)

- Fix $E \subseteq G^2$.
 - E is viewed as a **graph** over vertices $V = G$.
 - BLR: $E = G^2$
- Pick random $(x, y) \in E$.

The BLR test picked random $x, y \in G$.

Generalization (Graph tests)

- Fix $E \subseteq G^2$.
 - E is viewed as a **graph** over vertices $V = G$.
 - BLR: $E = G^2$
- Pick random $(x, y) \in E$.

Question

Which graphs induce **good** tests?

Good Graphs

Linearity Testing

Preliminaries

The Problem

Motivation

Formalization

Past Answers

New Answers

Summary

- Problem: the BLR test requires $2 \log |G|$ random bits.
- New goal: minimize randomness.
- Lower bound: $\log |G| + O(1)$.

Good Graphs

Linearity
Testing

Preliminaries

The Problem

Motivation

Formalization

Past Answers

New Answers

Summary

- Problem: the BLR test requires $2 \log |G|$ random bits.
- New goal: minimize randomness.
- Lower bound: $\log |G| + O(1)$.
- Idea: use graph tests.
 - The graph test T_E (induced by $E \subseteq G$) requires only $\log |E|$ random bits.

Good Graphs

- Problem: the BLR test requires $2 \log |G|$ random bits.
- New goal: minimize randomness.
- Lower bound: $\log |G| + O(1)$.
- Idea: use graph tests.
 - The graph test T_E (induced by $E \subseteq G$) requires only $\log |E|$ random bits.

Theorem

For all but $e^{-|G|}$ -fraction of graphs E sized $C \cdot |G| \cdot \log |H|$, and for every $\delta > 0$, the test T_E is a $(\delta, \delta/3 + O(\delta^2) + e^{-|G|})$ -test for linearity.

- Need $\log G + \log \log H + O(1)$ randomness.
- Problem: which **explicit** such graphs are good/bad?

- 3 **New Answers**
- Preliminaries
 - Proposed Test
 - Analysis

Cayley Graphs

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Let $S \subset H$ be a **generating subset**.

Definition

The **Cayley graph** $\mathcal{G} = \text{Cay}(H; S)$ is the graph over vertices H with edges $\{ \langle h, h \cdot s \rangle \mid h \in H, s \in S \}$.

Cayley Graphs

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Let $S \subset H$ be a **generating subset**.

Definition

The **Cayley graph** $\mathcal{G} = \text{Cay}(H; S)$ is the graph over vertices H with edges $\{ \langle h, h \cdot s \rangle \mid h \in H, s \in S \}$.

Properties

- $|S|$ -regular.
- If S is **symmetric** then \mathcal{G} is '**undirected**'.
- \mathcal{G} is **vertex-transitive**: $\langle u, v \rangle \in \mathcal{G} \Rightarrow \langle h \cdot u, h \cdot v \rangle \in \mathcal{G}$.
- \mathcal{G} is **simple** iff $1 \notin S$.

Eigenvalues

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Definition

The **eigenvalues** of a graph \mathcal{G} are the eigenvalues of its adjacency matrix ($M_{ij} = [\langle i, j \rangle \in \mathcal{G}]$).

Facts about Cayley graphs

- The top eigenvalue is $\lambda_1 = d$.

Definition

The **eigenvalues** of a graph \mathcal{G} are the eigenvalues of its adjacency matrix ($M_{ij} = [\langle i, j \rangle \in \mathcal{G}]$).

Facts about Cayley graphs

- The top eigenvalue is $\lambda_1 = d$.
- Define the normalized second eigenvalue:

$$\lambda[\mathcal{G}] := \frac{1}{d} \cdot \max(\lambda_2, |\lambda_n|).$$

then $\lambda \ll 1$ (i.e., $\lambda_2 \ll \lambda_1$).

- If we remove **any** $2\delta dn$ edges from the graph, then remains a connected component of size at least $(1 - \epsilon)n$.
 - $\epsilon := 4\delta / (1 - \lambda[\mathcal{G}]) < 1/3$.

Constructive Tests

Linearity
Testing

Definition

For $\mathcal{G} = \text{Cay}(G; S)$, let $T_{\mathcal{G}}$ be the test that picks at random an edge $\langle x, x + s \rangle \in \text{Cay}(G; S)$ and checks whether

$$f(x) \cdot f(s) = f(x + s).$$

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Constructive Tests

Definition

For $\mathcal{G} = \text{Cay}(G; S)$, let $T_{\mathcal{G}}$ be the test that picks at random an edge $\langle x, x + s \rangle \in \text{Cay}(G; S)$ and checks whether

$$f(x) \cdot f(s) = f(x + s).$$

Theorem

For every G, H , and symmetric subset $S \subset G$, the test $T_{\mathcal{G}}$

- **accepts with probability 1** all homomorphisms $h : G \rightarrow H$;
- **rejects with probability δ** any function $f : G \rightarrow H$ which is ϵ -far from being an affine homomorphism;

where

- $\delta := \text{Prob}[f(g) \cdot f(s) \neq f(g + s)]$;
- $\epsilon := 4\delta / (1 - \lambda[\mathcal{G}]) < 1/3$.

Extracting a Homomorphism

Linearity
Testing

Take $f, G, H, S, \mathcal{G}, T_{\mathcal{G}}, \delta, \epsilon < 1/3$, and $\lambda := \lambda[\mathcal{G}]$ as before.
Denote $d := |S|$ and $n := |G|$.

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Conclusion

If $\delta \propto \epsilon$ is small, then f is close to an affine homomorphism.

Extracting a Homomorphism

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Take $f, G, H, S, \mathcal{G}, T_{\mathcal{G}}, \delta, \epsilon < 1/3$, and $\lambda := \lambda[\mathcal{G}]$ as before.
Denote $d := |S|$ and $n := |G|$.

Overview

- 1 For every x , almost all y agree on the value of

$$\varphi(x) := \text{Plurality}_{y \in G}[f(x + y) \cdot f(y)^{-1}].$$

- 2 φ is a homomorphism.
- 3 φ is close to an affine shift of f .

Conclusion

If $\delta \propto \epsilon$ is small, then f is close to an affine homomorphism.

The Plurality Function I

Linearity
Testing

Claim

For every $x \in G$,

$$\text{Prob}_{y \in G} [\varphi(x) = f(x + y) \cdot f(y)^{-1}] \geq 1 - \epsilon.$$

Proof

◀ go back

▶▶ skip proof

The Plurality Function I

Claim

For every $x \in G$,

$$\text{Prob}_{y \in G} [\varphi(x) = f(x + y) \cdot f(y)^{-1}] \geq 1 - \epsilon.$$

Proof

- Fix $x \in G$.
- By definition, $\delta = \text{Prob}_{y \in G, s \in S} [f(x + y)f(s) \neq f(x + y + s)]$.
- Number of **bad edges**, $\langle y, y + s \rangle \in \mathcal{G}$ having $f(y) \cdot f(s) \neq f(y + s)$ or $f(x + y)f(s) \neq f(x + y + s)$, is at most $2\delta dn$.

The Plurality Function II

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Let $H_x \subset \mathcal{G}$ be obtained by removing all (undirected) bad edges, and let $C_x \subset H_x$ be a connected component of size $(1 - \epsilon)n$.

- Exists because there are at most $2\delta dn$ bad edges.

The Plurality Function II

Let $H_x \subset \mathcal{G}$ be obtained by removing all (undirected) bad edges, and let $C_x \subset H_x$ be a connected component of size $(1 - \epsilon)n$.

- Exists because there are at most $2\delta dn$ bad edges.

Lemma

For any $u, v \in C_x$, we have

$$f(x + v) \cdot f(v)^{-1} = f(x + u) \cdot f(u)^{-1}.$$

Proof

Take a path $v = v_1, \dots, v_t = u$ in C_x . Write $v_{i+1} = \langle v_i, v_i + s_i \rangle$. Observe that

$$f(v_i)^{-1} \cdot f(v_{i+1}) = f(s_i) = f(x + v_i)^{-1} \cdot f(x + v_{i+1}).$$

The Plurality Function III

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Conclusion

- Thus, $f(x + v) \cdot f(v)^{-1}$ is constant for $v \in C_x$.
- $|C_x| \geq (1 - \epsilon)n > \frac{1}{2} |G|$.

The Homomorphism I

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries
Proposed Test

Analysis

Summary

Claim

φ is a homomorphism.

Proof

◀ go back

▶▶ skip proof

The Homomorphism I

Linearity
Testing

Claim

φ is a homomorphism.

Proof

Fix $x, y \in G$.

Define:

- $p := \text{Prob}_{h \in G} [\varphi(x) \cdot \varphi(y) = \varphi(x + y)]$.
- $A(x, y) := \{ \varphi(x) = f(y) \cdot f(-x + y)^{-1} \}$.

Then:

- $p \in \{0, 1\}$.
- $p \geq \text{Prob}_{h \in G} [A(x, x + h) \wedge A(y, h) \wedge A(x + y, x + h)]$.

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

The Homomorphism II

From Claim 1:

$$\text{Prob}_{h \in G} [A(x, x + h)] \geq 1 - \epsilon;$$

$$\begin{aligned} \text{Prob}_{h \in G} [A(y, h)] &= \text{Prob}_{h \in G} [A(y, y + (-y + h))] \\ &= \text{Prob}_{h' \in G} [A(y, y + h')] \\ &\geq 1 - \epsilon; \end{aligned}$$

$$\text{Prob}_{h \in G} [A(x + y, x + h)] \geq 1 - \epsilon.$$

The Homomorphism II

From Claim 1:

$$\text{Prob}_{h \in G} [A(x, x + h)] \geq 1 - \epsilon$$

$$\text{Prob}_{h \in G} [A(y, h)] \geq 1 - \epsilon$$

$$\text{Prob}_{h \in G} [A(x + y, x + h)] \geq 1 - \epsilon$$

Therefore:

- $p \geq \text{Prob}_{h \in G} [A(x, x + h) \wedge A(y, h) \wedge A(x + y, x + h)]$.
- $p \geq 1 - 3\epsilon > 0$.

The Homomorphism II

From Claim 1:

$$\text{Prob}_{h \in G} [A(x, x + h)] \geq 1 - \epsilon$$

$$\text{Prob}_{h \in G} [A(y, h)] \geq 1 - \epsilon$$

$$\text{Prob}_{h \in G} [A(x + y, x + h)] \geq 1 - \epsilon$$

Therefore:

- $p \geq \text{Prob}_{h \in G} [A(x, x + h) \wedge A(y, h) \wedge A(x + y, x + h)].$

- $p \geq 1 - 3\epsilon > 0.$

- $p \in \{0, 1\}.$

$\Rightarrow p = 1.$

$\Rightarrow \varphi$ is a homomorphism.

The Affine Shift

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries
Proposed Test

Analysis

Summary

Claim

There is some $\gamma \in H$ such that $\text{dist}(\varphi, f \cdot \gamma) \leq \epsilon$.

Proof

◀ go back

▶▶ skip proof

The Affine Shift

Claim

There is some $\gamma \in H$ such that $\text{dist}(\varphi, f \cdot \gamma) \leq \epsilon$.

Proof

- $G_{x,y} := \{ \varphi(x) = f(x+y) \cdot f(y)^{-1} \}$
- $\forall x. |\{ y : G_{x,y} \}| \geq |C_x| \geq (1 - \epsilon) |G|$

The Affine Shift

Claim

There is some $\gamma \in H$ such that $\text{dist}(\varphi, f \cdot \gamma) \leq \epsilon$.

Proof

- $G_{x,y} := \{ \varphi(x) = f(x+y) \cdot f(y)^{-1} \}$
- $\forall x. |\{ y : G_{x,y} \}| \geq |C_x| \geq (1 - \epsilon) |G|$
- $\forall x. \text{Prob}_y [G_{x,y}] \geq 1 - \epsilon$
- $\text{Prob}_{x,y} [G_{x,y}] \geq 1 - \epsilon$
- $\exists y_0. \text{Prob}_x [G_{x,y_0}] \geq 1 - \epsilon$

The Affine Shift

Claim

There is some $\gamma \in H$ such that $\text{dist}(\varphi, f \cdot \gamma) \leq \epsilon$.

Proof

- $G_{x,y} := \{ \varphi(x) = f(x+y) \cdot f(y)^{-1} \}$
- $\forall x. |\{ y : G_{x,y} \}| \geq |C_x| \geq (1 - \epsilon) |G|$
- $\forall x. \text{Prob}_y [G_{x,y}] \geq 1 - \epsilon$
- $\text{Prob}_{x,y} [G_{x,y}] \geq 1 - \epsilon$
- $\exists y_0. \text{Prob}_x [G_{x,y_0}] \geq 1 - \epsilon$
- $\text{Prob}_x [\varphi(x) = f(x+y_0) \cdot f(y_0)^{-1}] \geq 1 - \epsilon$
- $\text{Prob}_{x' \in G} [\varphi(x' + (-y_0)) = f(x') \cdot f(y_0)^{-1}] \geq 1 - \epsilon$

Extracting a Homomorphism

Linearity
Testing

Preliminaries

The Problem

New Answers

Preliminaries

Proposed Test

Analysis

Summary

Take $f, G, H, S, \mathcal{G}, T_{\mathcal{G}}, \delta, \epsilon < 1/3$, and $\lambda := \lambda[\mathcal{G}]$ as before.
Denote $d := |S|$ and $n := |G|$.

Overview

- 1 For every x , almost all y agree on the value of

$$\varphi(x) := \text{Plurality}_{y \in G}[f(x + y) \cdot f(y)^{-1}].$$

- 2 φ is a homomorphism.
- 3 φ is close to an affine shift of f .

Conclusion

If $\delta \propto \epsilon$ is small, then f is close to an affine homomorphism.

Summary

Linearity
Testing

Preliminaries

The Problem

New Answers

Summary

- 1 Preliminaries
- 2 The Problem
- 3 New Answers

The answer is soup.

The End.