max planck institut informatik

Lecture 8:Testing Sparse images - Monotonicity

Themis Gouleakis

June 1, 2021

Definitions

Images:

- An image will be represented by a $0/1$-valued $n \times n$ matrix M.
  - Dense if it contains $\Omega(n^2)$ 1-entries/pixels.

Access models:

- Dense image model: (analog to dense graph model)
  - Query access to entries
- Sparse image model: (analog to sparse graph model)
  - Query access to entries
  - Sample access to 1-entries

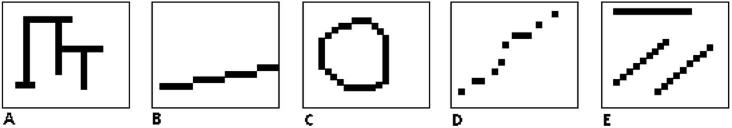
Distance:

- Dense image model: $\delta(M, M') = \frac{d_H(M,M')}{n^2}$
- Sparse image model: $\delta(M, M') = \frac{d_H(M,M')}{w(M)}$

where $w(M)$ is the number of 1-pixels in M

Example properties

- Connectivity: Graph of M is connected
- Line imprint: $\exists$ a line segment such that $M(i, j) = 1$ iff the line intersects the pixel.
- Convexity: Similar for a convex shape
- Monotonicity: $\forall$ (i_1, j_1) and (i_2, j_2) 1-pixels it holds: $i_1 < i_2 \Rightarrow j_1 \leq j_2$.



Assumptions and warmup

- The algorithm is given an estimate of $w(M)$
  - Can be obtained using $\tilde{O}(\min\{\sqrt{w(M)}, \frac{n^2}{w(M)}\})$ queries.

Theorem (sampling only tester)

There exists a sampling-only property tester for monotonicity the requires an estimate $\hat{w} = \Theta(w(M))$ and has sample complexity and running time $O(\sqrt{w(M)/\epsilon})$

The algorithm is the following:

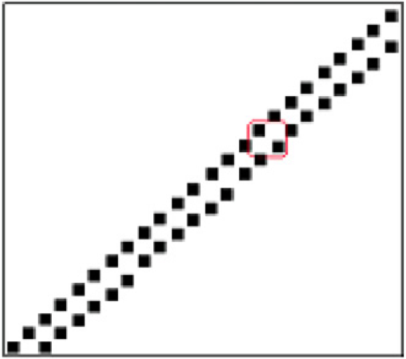
1. Take $\Theta(\sqrt{\hat{w}/\epsilon})$ samples of 1-pixels u.a.r
2. If some pair violates monotonicity REJECT, otherwise ACCEPT

Proof:

Assumptions and warmup

Theorem (sampling lower bound)

There exists NO sampling-only property tester for monotonicity using $o(\sqrt{w(M)})$ samples and no queries.



Improved algorithm using queries

Algorithm with queries:

(1) *Take a sample* S_1 *of* $t_1 = \Theta(g_1 \log n/\epsilon^2)$ *1-pixels. If* S_1 *contains a violating pair, then* REJECT; *otherwise, continue.*

(2) *Take a sample* S_2 *of* $t_2 = \Theta(1/\epsilon)$ *1-pixels. If there is a violating pair in* $S_1 \cup S_2$*, then* REJECT. *Otherwise, for each of the 1-pixels* (a, b) *in* S_2 *perform the following subtest:*
—*For* $\ell = 1$ *to* g_2*, where* ℓ *increases by a multiplicative factor of 2 in each iteration, uniformly select* $t_3(\ell) = \Theta(\ell \cdot (n/w(M)) \cdot \log(n)/\epsilon^2)$ *entries in the submatrix of dimensions* $\ell \times \ell$ *that* (a, b) *is the bottom-right corner of, and similarly for the* $\ell \times \ell$ *submatrix that* (a, b) *is the top-left corner of, and perform queries on all these pixels. If any is answered by '1', then* REJECT.

(3) *If no step caused rejection, then* accept.

Remarks:

- We try to decompose into a set of submatrices with the properties:
  - Captures most 1-pixels in *M*
  - No cross-violations
- Structural result: Either we detect a violation or the above holds
  - Latter case: Use queries to detect violations within submatrices.

Improved algorithm using queries

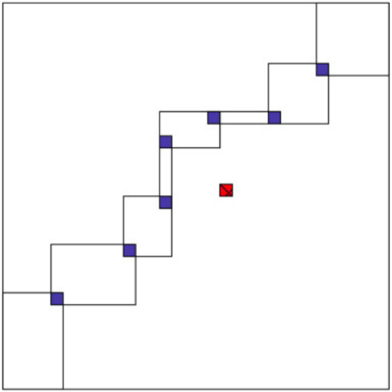
We will show the following theorem for the above algorithm:

Theorem (sampling and query tester)

There exists an one-sided error property tester for monotonicity the requires an estimate $\hat{w} = \Theta(w(M))$ and has sample and query complexity as well as its running time is $\tilde{O}\left(\frac{n^{2/3}}{w(M)^{1/3}\epsilon^2}\right)$

Improved algorithm using queries

- Let S be a subset of 1-pixels obeying monotonicity.
- The picture below shows the corresponding set of submatrices $\mathcal{M}(S)$.



Improved algorithm using queries

Lemma *With high constant probability over the choice of the sample S_1^1, either S_1^1 contains a violating pair or there exists a subset of S_1^1 of size at most $6g_1$ $(= 6n^{2/3}/w(M)^{1/3})$, denoted $\widetilde{S}_1^1$ such that at most an $(\epsilon/16)$-fraction of the 1-pixels in M belong to heavy submatrices in $\mathcal{M}(\widetilde{S}_1^1)$.*

Proof: