STOS LECTURE 13 3/19/2008

Note Title 3/19/2008

TODAY: LIST DECODING ALGORITHM FOR REED SOLOMON CODES

[Reminder: No Lecture Monday 3/31]

Recall from last time

Johnson Bound: $(n, k, d)_q$ code is

$\left(1 - \sqrt{1 - \frac{d}{n}}, \text{poly}\right)$ – list decodable

↑ fraction errors, ↑ list size

Specializing to RS code: $[n, k, n-k]_q$ code

is combinatorially list decodable from $n - \sqrt{kn}$ errors.

Algorithmic Problem / Challenge

Given: F, $\alpha_1 \ldots \alpha_n$, k, $y_1 \ldots y_n$

& agreement parameter t

Find: list of all poly p of degree k s.t. $p(\alpha_i) = y_i$ for at least t values of $i \in [n]$.

[Johnson Bound $\Rightarrow$ should be doable in poly time for $t \geq \sqrt{kn}$]

Today: [Sudan '96] $t \geq \sqrt{2kn}$

[Guruswami + S. '98] $t \geq \sqrt{kn}$

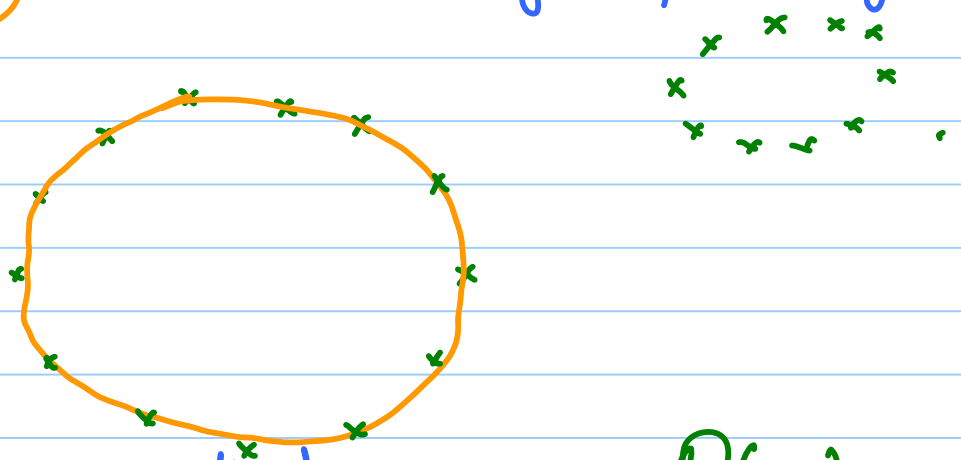
Algebra ⟷ Geometry

Idea: ① Get algebraic "explanation" of n points $\{(\alpha_i, y_i)\}_{i=1}^{n}$

② Throw away points & "stare" at algebraic "explanation"

Algebraic Explanation?

Example ◯ below is an alg. explanation of



In general some small degree curve $Q(x,y) \neq 0$ s.t. $Q(\alpha_i, y_i) = 0 \quad \forall i, j$

Lemma 1: Given n points a low-degree ($\deg_x(Q), \deg_y(Q) \leq \sqrt{n}$) algebraic explanation exists, and can be found in poly time.

Proof: Let $Q(x,y) = \sum_{i,j} q_{ij} x^i y^j$

$\{q_{ij}\}$ unknown. # unknown's $= (\sqrt{n}+1)^2 > n$.

Constraints: ① $Q(a_\ell, y_\ell) = 0$

$$\Rightarrow \sum_{i,j} q_{ij} a_\ell^i y_\ell^j = 0$$

↑

homogenous, linear, equation

② $Q \not\equiv 0 \Rightarrow$ want non-trivial solution.

#variables > # equations $\Rightarrow$ solution exists, can be computed

□

Question: Isn't this overfitting the points?

($> n$ q_{ij}'s explaining $y_1, \dots y_n$)

Answer: YES & NO.

YES ⇐ This is why Lemma 1 holds.

NO Doesn't overfit nice data ...

Example: if $\exists$ $\sqrt{n}$ deg poly p s.t.

$y_i = p(\alpha_i)$ $\forall i$, then $Q(x,y)$

$= y - p(x)$.

But if $\exists$ $\sqrt{n} - 1$ deg p s.t.

$y_i = p(\alpha_i)$ $\forall i$, then

$Q(x,y)$ can be $(ax+by+c)(y-p(x))$

for any (a,b,c). Is this all that can go wrong?

Lemma 2: if $Q \not\equiv 0$,

① $Q(\alpha_i, y_i) = 0 = y_i - p(\alpha_i)$ for t points $\{(\alpha_i, y_i)\}_{i=1}^{t}$, $\deg(p) \le k$

② $\deg_x Q,\ \deg_y Q \le D$

③ $t > (k+1)D$

then $y - p(x)$ divides $Q(x,y)$

Proof: • How to prove $y - p(x) \mid Q(x,y)$?

• Think $Q(x,y) \in \underbrace{F(x)}_{\text{field of rational functions in } x}[y]$

- Trying to prove β root of $Q_x(y)$,
  where $\beta = p(x) \in \mathbb{F}(x)$.
- Holds iff $Q_x(\beta) = 0$. Is it?
- $Q_x(\beta) = Q(x, p(x)) = g(x)$
  ↑
  some poly in x
  $\deg(g) \leq (k+1)D$
- $g(\alpha_i) = Q(\alpha_i, p(\alpha_i))$
  $= Q(\alpha_i, y_i)$ (since $y_i = p(\alpha_i)$)
  $= 0$ (since $Q(\alpha_i, y_i) = 0$)

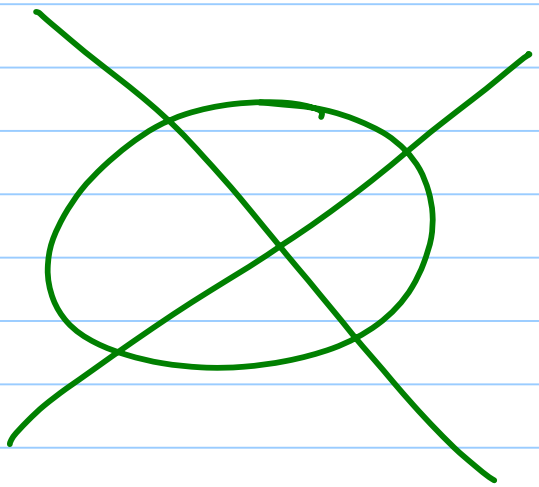
$\Rightarrow$ g has $t > (k+1)D$ roots $\Rightarrow$ $g \equiv 0$.

"Staring at Algebraic Explanation"

E.g. $Q(x,y) = x^4 - y^4 - x^2 + y^2$

Plot all zeroes .....

Zeroes =



Split naturally into three sets ....

Algebraically ... factor

$$x^4 - y^4 - x^2 + y^2 = (x^2+y^2-1)(x+y)(x-y).$$

Lemma 2 ⇒ Such factors have useful information.

[Kaltofen, von zur Gathen, Chistov, Grigoriev, Landau] ⇒ factoring in polytime.

Putting it together

Algorithm 1:

1. find $Q \not\equiv 0$, $\deg_x Q \leq \sqrt{n}$, $\deg_y Q \leq k$,

$$Q(\alpha_i, y_i) = 0 \quad \forall i$$

2. Factor $Q(x,y)$; report all p s.t. $y - p(x)$ divides $Q(x,y)$

Theorem: Algorithm 1 recovers all poly with agreement $t > (k+1)\sqrt{n}$

(follows from Lemma 1 + Lemma 2;)

(not great if $k > \sqrt{n}$;

great if $k = o(\sqrt{n})$.)

Easy Improvements

① let $\deg_x Q \le \sqrt{kn}$, $\deg_y Q \le \sqrt{\frac{n}{k}}$

Lemma 1′ : Such Q also exists.

Lemma 2′ : $t > 2\sqrt{kn} \Rightarrow$

$y - p(x) \mid Q(x, y)$.

Theorem′ : Agreement $2\sqrt{kn}$ suffices.

(Great if $k = o(n)$; but doesn't dominate [Peterson, ... Welch-Berlekamp])

(2) Choosing degree of Q even more carefully..

$$Q(x,y) = \sum_{i,j} q_{ij}\, x^i y^j$$

$$\underbrace{i + kj \le \sqrt{2kn}}$$

$(1,k)$- weighted degree of $Q \le \sqrt{2kn}$ ..

Lemma 1' } --- ⇒ Agreement $t > \sqrt{2kn}$ suffices.
Lemma 2' }

Using minimal wtd. degree of Q leads to algorithm that dominates W-B.

Theorem [S '96]: Gen list-decode RS from $t > \sqrt{2kn}$ agreement.

Going Beyond?

$t \to \sqrt{kn}$? (no factor of 2)

Example: Take ℓ lines in general position.

$n = \binom{\ell}{2}$ points of intersection.

Then n points have $\approx \sqrt{2n}$ lines passing through $\sqrt{2n}$ points each.

Each point has 2 lines through it (by construction).

Can we hope to find all such lines by our approach?

- Careful calculations ⇒ just miss all lines.

- Uncareful explanation ⇒

  if we had found "all lines" or even "most lines"

  ...

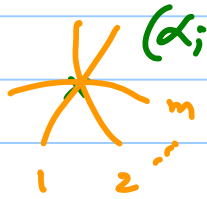
  - would have asked Q to pass through every point <u>once</u>
  - would have found Q that passes through most points <u>twice</u>
  - We're counting on too much <u>luck</u>!

  ... lets not ... & lets require Q to pass through every point <u>twice</u>.

Multiplicity of Zeroes

$Q(x,y)$ has zero of multiplicity m at (α_i, y_i) if

(1) Pictorially



Q goes through pt. m times.

(2) Analytically $Q(\alpha_i, y_i) = 0$

$$\frac{\partial Q}{\partial x}(\alpha_i, y_i) = 0$$

$$\vdots$$

$$\frac{\partial^{j+l} Q}{\partial x^j \partial y^l}(\alpha_i, y_i) = 0 \quad \forall\, j+l \le m$$

Correct definition for us.

③ Algebraically

$$\overline{Q}(x,y) \triangleq Q(x+\alpha_i, y+y_i)$$

has no support on monomials of degree $\leq m$.

[Guruswami + S' 98]

Algorithm 2: /* Set parameters m, ℓ */

① find $Q \neq 0$, $(1,k)$-wtd. deg $(Q) \leq \ell$

s.t. Q has zeroes of mult. m at $\{\alpha_i, y_i\}_{i=1}^{n}$

② Factor Q; report p s.t. $y - p(x) \mid Q(x,y)$

Theorem: As $m \to \infty$, needs $t \to \sqrt{2n}$.

Proof Omitted

Rough Idea:

$m=1$	vs.	$m=2$
n constraints		$3n$ constraints
$\deg = l$		$\Rightarrow \deg \approx \sqrt{3l}$
require $l+1$ zeroes to find p		require $\frac{\sqrt{3l}}{2}$ zeroes to find p (Every zero counts twice).