LECTURE 21

4/30/2012

Note Title 4/28/2012

Today: Algebra in Coding Theory

- Reed-Solomon Codes
- List-decoding algorithms
- Ideal - Error-Correcting Codes & Decoding

Error-Correcting Codes

DVD Motivation:

- Wish to store $m \in M$ as a sequence of symbols $(x_1, \ldots x_n) \in \Sigma^n$ s.t. even after t symbols are corrupted arbitrarily $m / (x_1 \ldots x_n)$ are uniquely determined.
- For simplicity $\underset{\substack{\uparrow \\ \text{message}}}{M} = \Sigma^k$

[Hamming]

Definition: $E : \Sigma^k \xrightarrow{\text{Encoding}} \Sigma^n$

$$m \longmapsto (x_1 \ldots x_n)$$

Code $\rightarrow$ $C = \text{Image}(E)$; $\Delta(\bar{x}, \bar{y}) = |\{i \mid x_i \neq y_i\}|$

$$\Delta(C) = \min_{\bar{x} \neq \bar{y} \in C} \{\Delta(\bar{x}, \bar{y})\} \leftarrow \text{distance}$$

Proposition:

Code of distance $2t+1$ corrects t errors.

[Notation]: Code $C \subseteq \Sigma^n$, $|C| = |\Sigma|^k$,

$\Delta(C) = d$, $|\Sigma| = q$,

denoted an $[n, k, d]_q$ code.

[Singleton] **Bound:** $\forall$ $[n, k, d]_q$ codes, $k + d \leq n+1$

Proof: Let $\pi : \Sigma^n \to \Sigma^{k-1}$ map $(x_1 \ldots x_n) \longmapsto (x_1 \ldots x_{k-1})$.

By PHP $\exists$ $x \neq y \in C$ s.t. $\pi(x) = \pi(y)$.

$\Rightarrow$ $\Delta(x, y) \leq n - (k-1)$ $\Rightarrow$ $d \leq n - k + 1$ $\boxtimes$

Reed-Solomon Codes

Defn: $\Sigma = \mathbb{F}_q$; $n \leq q$, $\{\alpha_1 \ldots \alpha_n\} \subseteq \mathbb{F}_q$

distinct

$$RS = RS_{n,k,\mathbb{F}_q,\{\alpha_1 .. \alpha_n\}}$$

$$= \{ (p(\alpha_1), \ldots p(\alpha_n)) \mid p \in \mathbb{F}_q[x], \deg(p) < k \}$$

= evaluations of univ. polys of deg $< k$.

Proposition: $\Delta\left(RS_{n,k,\mathbb{F}_q,\{\alpha_1..\alpha_n\}}\right) = n - k + 1$

Proof: Consider $f, g \in \mathbb{F}_q[x]$, $\deg(f), \deg(g) < k$.

Let $S = \{ i \mid f(\alpha_i) = g(\alpha_i) \}$. $\Delta(\bar{f}, \bar{g}) = n - |S|$

$$|S| \leq \deg(f-g) \leq k-1$$

$$\Rightarrow \Delta(\bar{f}, \bar{g}) \geq n - (k-1)$$

$\boxtimes$

Note: Meets Singleton Bound !!

The (list) decoding problem for RS codes

Given: $\alpha_1 \dots \alpha_n \in \mathbb{F}_q$

$\beta_1 \dots \beta_n \in \mathbb{F}_q$

Find: The/all polynomials p with

1. $\deg(p) < k$
2. $\left| \{ i \mid p(\alpha_i) = \beta_i \} \right| \geq n - t \triangleq a$.

- [Hamming bound]

  $a > \frac{n+k}{2} \Rightarrow$ unique p

- Inclusion-Exclusion Counting

  $a > \sqrt{2kn} \Rightarrow$ # p's small, $< 2\sqrt{\frac{n}{k}}$.

- [Johnson bound]

  $a > \sqrt{kn} \Rightarrow$ # p's small, $< n^2$

But can we find them?

Main Idea:

- Need an "algebraic description" of points

$$\{(\alpha_i, \beta_i) \mid i = 1 \dots n\}$$

- Should have low "algebraic complexity" if

$$\beta_i = p(\alpha_i) \quad \forall i$$

- Complexity should degrade nicely if we add $\underbrace{\text{random points}}_{\text{errors}}$ (α_i, β_i).

- Classical approach (effectively)
  [Peterson, Berlekamp, Massey, Welch-Berlekamp, Gemmell-S.]
  Use rational functions

- [S.'97, Guruswami + S '98]
  Use Ideal / Variety Correspondence.

find $Q \neq 0$ s.t. $Q(\alpha_i, \beta_i) = 0 \quad \forall i$.

Decoding Algorithm

Step 1: Find $Q(x,y)$, $\deg Q \le D$, $Q \not\equiv 0$
$\forall i \quad Q(\alpha_i, \beta_i) = 0$

Step 2: Factor Q into irreducibles;
report all p s.t. $y - p(x) \mid Q(x,y)$.

Analysis :

Step 1: ① Finding Q if it exists : linear system.
② Solution exists if # monomials in Q
$> n$.
[e.g. if $D > \sqrt{2n}$]

Step 2: Obviously solution exists;

Lemma: $Q(x,y)$ & $y - p(x)$ have too many common zeroes $\Rightarrow$ common factor.
(Bezout's theorem in plane). $\boxtimes$

Conclusion: • Setting $D = \sqrt{2n}$, get algorithm that works if $a > k\sqrt{2n}$.

• Better choice of monomials:

$$\left(\deg_x Q + (k-1)\deg_y Q < \sqrt{2kn}\right)$$

yields $a > \sqrt{2kn}$

(meets inclusion - exclusion bound)

Ideals & Error-Correcting Codes

- **Messages:** $\mathcal{M} \subseteq R \leftarrow$ ring, likely infinite

  $\uparrow$ finite

- **Coordinates:** $I_1, I_2 \ldots I_n$ ideals in R

- **Encoding:** $m \longmapsto (m \pmod{I_1}, \ldots \; m \pmod{I_n})$

Reed-Solomon

- $R = \mathbb{F}_q[x]$
- $\mathcal{M} = \mathbb{F}_q^{<k}[x]$
- $I_j = (x - \alpha_j)$

Chinese Remainder Code

- $R = \mathbb{Z}$
- $m = \{0, \ldots M\}$
- $I_j = (p_j)$

- So message is big (say n-bit) number.
- Encoding = residues modulo small (poly(n) large) primes.

Works almost as well as Reed-Solomon.

Other examples: almost all "algebraic" codes

esp. "Algebraic-Geometry Codes"

Ideal Decoding

Given: $R, \; I_1 \ldots I_n, \; M$

$\beta_1 \ldots \beta_n$

Find: all $m \in M$ s.t.

$$|\{ i \mid m - \beta_i \in I_i \}| \geq a$$

Algorithm Idea:

- Set up polynomial $Q \in R[y]$ s.t. $(y - m)$ is a factor of Q
- Let $J_i = I_i + (y - \beta_i)$
- $Q \in \bigcap_{i=1}^{n} J_i$

- Notion of "size" of elements of R
- $\text{size}(a+b) \leq \text{size}(a) + \text{size}(b)$

  $\text{size}(a \cdot b) \leq \text{size}(a) \cdot \text{size}(b)$
- Need: if $a \in \bigcap_{i \in S} I_i$

  thus $\text{size}(a) = \text{large}$.
- Need: lots of "small" elements.
- All the above imply $\exists Q$ with "small" coefficients, small degree st

$$Q = \bigcap_{i \in [n]} J_i$$

- $Q(m) \in \bigcap_{i \in A} I_i$, $Q(m)$ is small

$\Rightarrow$ $Q(m) = 0$ $\boxtimes$

Algorithmic Needs

1. finding small Q.
   - linear codes $\Rightarrow$ linear algebra
   - CRT codes $\Rightarrow$ LLL

2. Factorization over $R[y]$
   - RS codes $\Rightarrow$ Bivariate factorization
   - CRT $\Rightarrow$ LLL
   - AG codes $\Rightarrow$ Factorization over function fields.

Other Ideas

1. Multiplicities : $Q \in \left(\prod_{i=1}^{n} J_i \right)^m$

   gives better results.

2. Best known results for RS decoding

   $\#$ errors $\rightarrow$ $n - \sqrt{kn} < n-k$

3. [Parvaresh-Vardy], [Guruswami Rudra]

   codes where

   $\#$ errors $\rightarrow (1-\epsilon)(n-k)$

   [GR]: Folded Reed-Solomon Codes.

FRS Codes

$$\Sigma = \mathbb{F}_q^{\ell} \; ; \qquad n = \left\lfloor \frac{q-1}{\ell} \right\rfloor ; \quad \gamma \text{ primitive in } \mathbb{F}_q$$

- $m = (c_0, \ldots c_{k-1}) \in \mathbb{F}_q^k$: message
- Encoding : let $m(x) = \sum c_i x^i$

$$m^{(\ell)}(\alpha) \triangleq \langle m(\alpha), m(\gamma \cdot \alpha), \ldots m(\gamma^{\ell-1}\alpha) \rangle$$

$$m \longmapsto \left\langle m^{(\ell)}(\gamma^{i\ell}) \right\rangle_{i=0}^{n-1}$$

- List-decodability :

$$\alpha_1 \ldots \quad \alpha_n \qquad\qquad \alpha_i = \gamma^{i\ell}$$

Received $\leftarrow \left(\begin{pmatrix} \beta_{11} \\ \vdots \\ \beta_{1\ell} \end{pmatrix} \ldots \begin{pmatrix} \beta_{n1} \\ \vdots \\ \beta_{n\ell} \end{pmatrix} \right)$

- Let $Q(x, y_1 \dots y_\ell) \neq 0$ be s.t.

$$Q(\alpha_i, \beta_{i1} \dots \beta_{i\ell}) = 0 \quad \forall i$$

- As with RS codes: $Q(x, m(x), m(\gamma x) \dots m(\gamma^{\ell-1} x)) \equiv 0$

for m with large enough agreement.

- Let $R(y_1 \dots y_\ell) = Q(x, y_1 \dots y_\ell) \pmod{\underbrace{x^{q-1} - \gamma}_{\text{irreducible.}}}$

- Claim: $R(m, m^q, m^{q^2}, \dots m^{q^{\ell-1}}) = 0$

for m with large agreement

Proof: $m^q = \sum c_i x^{iq} = \sum c_i (\gamma x)^i \pmod{x^{q-1} - \gamma}$

$$= m(\gamma x)$$

- m is a root of $\Delta(y) = R(y, y^q, \dots y^{q^{\ell-1}})$

- $\deg(\Delta)$ small $\Rightarrow$ # roots small.