Order domains, Sakata's algorithm and majority voting

Michael E. O'Sullivan

San Diego State University

July 11, 2008, S^{3}CM, Soria, Spain

Reed-Solomon codes

Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ and consider the evaluation map.

$$\begin{aligned} ev : \mathbb{F}_q[x] &\longrightarrow \mathbb{F}_q^n \\ f &\longmapsto (f(\alpha_1), \ldots, f(\alpha_n)) \end{aligned}$$

Let L_m be the set of polynomials of degree m.
Define the codes:

$$\begin{aligned} E(\bar{\alpha}, m) &= ev(L_m) \\ C(\bar{\alpha}, m) &= E(\bar{\alpha}, m)^{\perp} \end{aligned}$$

One-point codes

- $\mathcal{X}$, a projective (smooth absolutely irreducible) curve over $\mathbb{F}$.
- Q, an $\mathbb{F}$-point on $\mathcal{X}$.
- $K(\mathcal{X})$ the function field of $\mathcal{X}$.
- Let R be the ring of functions with poles only at Q.

$$R = \{f \in K(\mathcal{X}) : v_P(f) \geq 0 \text{ for all } P \neq Q\}$$

- $L(mQ)$ elements of R with pole order at most m at Q.
- For $\mathcal{X}$ the projective line, R is a polynomial ring in one variable. The space $L(mQ)$ contains polynomials of degree at most m.

One-point codes

Let $P_1, P_2, \ldots, P_n$ be $\mathbb{F}$-points of $\mathcal{X}$ and $D = P_1 + \cdots + P_n$
We have the evaluation map

$$\begin{aligned} ev : R &\longrightarrow \mathbb{F}_q^n \\ f &\longmapsto (f(P_1), \ldots, f(P_n)) \end{aligned}$$

Define:

$$\begin{aligned} E(D, m) &= ev(L(mQ)) \\ C(D, m) &= E(D, m)^{\perp} \end{aligned}$$

This gives a nice family of codes generalizing RS codes amenable to Sakata's generalization of Berlekamp-Massey.

Order domains

The natural setting for Sakata's algorithm is order domains.

Definition

Let $\mathbb{F}$ be a field and let R be an $\mathbb{F}$-algebra. An *order function* on R is a map

$$\rho : R \longrightarrow \mathbb{N}_{-1}$$

which satisfies the following.

O1. The set $L_m = \{f \in R \mid \rho(f) \leq m\}$ is an $m+1$ dimensional vector space over $\mathbb{F}$.

O2. If $f, g, z \in R$ and z is nonzero then $\rho(f) > \rho(g) \Longrightarrow \rho(zf) > \rho(zg)$

The pair R, ρ is called an *order domain*.

Examples

- $\mathbb{F}[x, y]$ with grevlex.

$x^i y^j$	1	x	y	x^2	xy	y^2	...
ρ	0	1	2	3	4	5	...

- $\mathbb{F}[x, y]$ with lex is NOT.
  The space of elements smaller than y is infinite dimensional.
- **Proposition** For f, g with $\rho(f) > 1$, there exists n such that $\rho(f^n) > \rho(g)$.
- **Proof:**

$$\rho(1) < \rho(f) < \rho(f^2) \cdots < \rho(f^n)$$

Example from a curve, point

- Let $R = L(\infty Q)$ be the ring of functions with poles only at Q.
- Enumerate the Weierstrass semigroup $\Lambda = \{-v_Q(f) : f \in R\}$. $0 = \lambda_0 < \lambda_1 < \lambda_2 < \lambda_3, \ldots$ are the elements of Λ.
- Define an order function by

$$\begin{aligned} \rho : R &\to \mathbb{N}_{-1} \\ 0 &\to -1 \\ f &\mapsto i \text{ such that } -v_Q(f) = \lambda_i \end{aligned}$$

Equivalent formulation

Proposition

Property **O1** *is equivalent to all of the following being true.*

1. ρ *is surjective.*
2. $\rho(a) = -1$ *if and only if* $a = 0$.
3. $\rho(\alpha f) = \rho(f)$ *for all* $\alpha \in \mathbb{F}$.
4. $\rho(f + g) \leq \max(\rho(f), \rho(g))$.
5. *If* $f, g \neq 0$ *and* $\rho(f) = \rho(g)$, *there exists some* $\alpha \in \mathbb{F}$ *such that* $\rho(f - \alpha g) < \rho(f)$.

Alternative definition of order domain:
Replace **O1** by properties 2-5 above. ρ is not necessarily surjective.

Observations

- $\rho^{-1}(0) = \mathbb{F}$.
- R must be a domain.
- **Proposition** ρ induces a semigroup structure on $\mathbb{N}_0$ in which 0 is the identity, and there is a well defined operation $\oplus$.

$$\rho(f) \oplus \rho(g) = \rho(fg)$$

- **Proposition** ρ induces a partial order on $\mathbb{N}_0$.
  $a \preccurlyeq b$ when there exists a c such that $a \oplus c = b$.

Example

Let $R = \mathbb{F}[x, y]$ with glex.
We have

$$\begin{aligned} 1 \oplus 1 &= 3 \\ 1 &\not\preccurlyeq 2 \\ 1 &\preccurlyeq 3 \\ 2 \oplus 2 &= 5 \\ 2 &\not\preccurlyeq 3 \\ 2 &\preccurlyeq 4, 5 \end{aligned}$$

This is most easily seen using the isomorphism of the semigroup $\mathbb{N}_0, \oplus$ with $\mathbb{N}_0^2, +$ that is induced by ρ.

Back to valuations

Theorem

ρ determines a unique valuation on $K(R)$.
The residue field of this valuation is $\mathbb{F}$.

Proof.

- Let $S = \{f/g : \rho(f) \leq \rho(g)\}$.
- This is a local ring with maximal ideal $\mathfrak{n} = \{f/g : \rho(f) < \rho(g)\}$.
- If $f/g \in K(R)$ is not in S then g/f is.
- Therefore S is a valuation ring.
- Equivalently, there is a totally ordered group Γ and a map $v : K(R)^* \longrightarrow \Gamma$ such that $S = v^{-1}(\Gamma_{\geq 0})$.

□

Surface examples

Valuations on surfaces are interesting! See Zariski, "Reduction of singularities of an algebraic surface."

- $\mathcal{X}$ an algebraic surface over $\mathbb{F}$.
- C a smooth curve on $\mathcal{X}$ defines a valuation, but
  - The residue field is not $\mathbb{F}$.
  - The codimension $L(mC) \subseteq L((m+1)C)$ can grow without bound.
- So, let Q be a point on C.
- Q and C together define a valuation with residue field $\mathbb{F}$.
- There are quirkier examples!

Examples on the affine plane

- glex on $\mathbb{F}[x, y]$ is from $C = L_\infty$ and $Q = [0 : 1 : 0]$.
- weighted lex orders on $\mathbb{F}[x, y]$ come from blowing up Q and points above Q to obtain some exceptional curve E and a point Q on this curve, which define the valuation.
- Let $p < q$ be coprime positive integers.

  What is the valuation ring for the monomial order $\begin{bmatrix} p & q \\ 0 & 1 \end{bmatrix}$?

  Describe the geometry.
- Let $\tau > 1$ be irrational.
  What is the valuation ring for the monomial order defined by $[1, \tau]$?
  Describe the geometry.

Order domains: recall

Definition

Let $\mathbb{F}$ be a field and let R be an $\mathbb{F}$-algebra. An order function on R is a map

$$\rho : R \longrightarrow \mathbb{N}_{-1}$$

which satisfies the following.

O1. The set $L_m = \{f \in R \mid \rho(f) \leq m\}$ is an $m + 1$ dimensional vector space over $\mathbb{F}$.

O2. If $f, g, z \in R$ and z is nonzero then
$\rho(f) > \rho(g) \Longrightarrow \rho(zf) > \rho(zg)$

The pair R, ρ is called an order domain.

Let $z_b \in R$ satisfy $\rho(z_b) = b$.
This is a basis for R.

Properties of order domains: recall

- **Proposition** ρ induces a semigroup structure on $\mathbb{N}_0$ in which 0 is the identity, and there is a well defined operation $\oplus$.

$$\rho(f) \oplus \rho(g) = \rho(fg)$$

- **Proposition** ρ induces a partial order on $\mathbb{N}_0$. $a \preccurlyeq b$ when there exists a c such that $a \oplus c = b$.

Grobner bases

- Let I be an ideal in R.
- **Definitions:**

$$\begin{aligned}
&\Sigma(I) = \{\rho(f) : f \in I\},\\
&\sigma(I) = \min_{\preccurlyeq} \Sigma(I),\\
&F(I) = \{f_a : \rho(f_a) = a, f_a \in I\}_{a \in \sigma(I)}\\
&\Delta(I) = \mathbb{N}_0 \setminus \Sigma(I).
\end{aligned}$$

Theorem

$F(I)$ is a Grobner basis for I.

1. *$F(I)$ generates I.*
2. *Given any $h \in I$, $\rho(h) \preccurlyeq a$ for some $a \in \sigma(I)$.*
   So $\rho(f - \beta f_a z_b) < \rho(h)$ for some $\beta \in \mathbb{F}$ and $b \in \mathbb{N}_0$.
3. *$\{z_c : c \in \Delta(I)\}$ is a basis for R/I.*

Codes from order domains

- Let $P_1, \ldots, P_n$ be $\mathbb{F}$-points on the variety defined by R. Equivalently, maximal ideals of R with residue field $\mathbb{F}$.
- The evaluation map:

$$\begin{aligned} ev : R &\longrightarrow \mathbb{F}^n \\ f &\longmapsto (f(P_1), \ldots f(P_n)) \end{aligned}$$

- Let $E_m = ev(L_m)$ and let $C_m = E_m^{\perp}$.
- A check matrix for C_m is

$$H = \begin{bmatrix} ev(1) \\ ev(z_1) \\ ev(z_2) \\ ev(z_3) \\ \cdots \\ ev(z_m) \end{bmatrix}$$

The decoding problem

- Send $c \in C_{\bar{m}}$.
- Receive $v \in \mathbb{F}^n$.
- Error is $e = v - c$.
- The *error locator ideal* is $I^e = \{f \in R : f(P_k) = 0 \text{ for all } k \text{ such that } e_k \neq 0\}$.
- Decode by finding a Grobner basis for I^e.
- Notation: $\Sigma^e = \Sigma_{I^e}$ and similarly, σ^e, Δ^e, $\delta^e = \max_{\preccurlyeq}\{c \in \Delta^e\}$.

The syndrome as a function

- Let $s = Hv = H(c + e) = He$.
- Extend the notion of syndrome to a function:

$$\begin{aligned} S^e : R &\longrightarrow \mathbb{F} \\ h &\longmapsto ev(h) \cdot e \end{aligned}$$

- Then z_m maps to s_m for $m \leq \bar{m}$.
- Define $s_m = S^e(z_m)$ for all $m \in \mathbb{N}_0$.

Two cooperative algorithms for decoding

- Berlekamp-Massey-Sakata: Process sequence $s_0, \ldots, s_{\bar{m}}, \ldots$ to get a Grobner basis for I^e.
- Feng-Rao/Duursma majority voting: Compute s_{m+1} from $s_m, s_{m-1}, s_{m-2}, \ldots$ and data from the mth iteration of the algorithm.
- If the error vector is "not too bad" we can compute s_m for enough $m > \bar{m}$ to find a Grobner basis for I^e.
- Majority voting gives get better decoding capability than BMS alone.

Crucial concepts

- Notice: For $f \in I^e$, $S^e(fg) = 0$ for all g.
- For $f \notin I^e$, define

$$\begin{aligned} \text{span}(f) &= \min\{c \in \mathbb{N}_0 : S^e(fz_c) \neq 0\} \\ \text{fail}(f) &= \rho(f) \oplus \text{span}(f) \end{aligned}$$

- An f with large span is "pretending" to be in I^e.

Approximations to I^e, etc.

Definitions:

$$I^m = \{f : \text{fail}(f) > m\}$$
$$\Sigma^m = \{\rho(f) : f \in I^m\}$$
$$\sigma^m = \min_{\preccurlyeq} \Sigma^m$$
$$\Delta^m = \mathbb{N}_0 \setminus \Sigma^m$$
$$\delta^m = \max_{\preccurlyeq} \Delta^m$$

Proposition

$\Delta^m = \{\text{span}(f) : \text{fail}(f) \leq m\}$.

Berlekamp-Massey-Sakata

Given $s_m = S^e(z_m)$.

Data σ^m and δ^m and sets of functions:

$$F^m = \{f_a : \rho(f_a) = a, \text{fail}(f_a) > m\}_{a \in \sigma^m}$$
$$G^m = \{g_c : \text{span}(g_c) = c, \text{fail}(g_c) \leq m\}_{c \in \delta^m}$$

Initialize For $m = -1$, $\sigma^{-1} = \{0\}$, $F^{-1} = \{1 \in R\}$ $\delta^{-1} = \emptyset$.

For $m = 0$ to m large enough, compute Data(m) from Data$(m-1)$.

How to compute Data(m)?

- Test each $f_a \in F^{m-1}$ to see if fail $f_a > m$.
- If f_a fails and $m - a \notin \Delta^{m-1}$ then $m - a \in \delta^m$ and f_a becomes $g_{m-a} \in G^m$ (case ($\star$)).
  Compute new δ^m, G^m from δ^{m-1}, G^{m-1}
  and failures of such f_a.
- Compute σ^m using $\Sigma^m = \mathbb{N}_0 \setminus \Delta^m$.
- Compute F^m using combinations like

$$\begin{array}{ll} z_i f_a + \mu g_c & \text{in case } (\star) \\ f_a + \mu z_i g_c & \text{else} \end{array}$$

Stopping criteria

Proposition

Let c_{max} be the largest integer in Δ^e. Then, for $m \geq c_{max} \oplus c_{max}$, $\Delta^m = \Delta^e$.

Proposition

Let s_{max} be the largest integer in σ^e and let $M = c_{max} \oplus \max\{c_{max}, s_{max}\}$.

For any $m \geq M$, if $F^m m$ is a Gröbner subset of I^m, then F^m is a Gröbner basis of I^e.

Majority voting: preliminaries

Consider the change in the set Δ.

- Notice that $\Delta^m \supsetneq \Delta^{m-1}$ iff for some $a \in \sigma^{m-1}$, $\text{fail}(f_a) = m$ and $\text{span}(f_a) \notin \Delta^{m-1}$. In this case $a \preccurlyeq m$.
- Set $N_m = \{a \in \mathbb{N}_0 : a \leq m\}$ (defined just by arithmetic of $\mathbb{N}_0, \oplus$).
- Then $\Delta^m \setminus \Delta^{m-1} \subseteq N_m \bigcap \Sigma^{m-1}$.
- Let $\Gamma^m = N_m \bigcap \Sigma^{m-1}$.

Main theorem for majority voting

- Suppose $s_0, s_1, \ldots, s_{m-1}$ are known, but not s_m.
- Elements of Γ^m will vote for the value of s_m.

Theorem

If $|N_m| > 2|N_m \bigcap \Delta^e|$ *then* $|\Sigma^m \bigcap \Gamma^m| > |\Delta^m \cap \Gamma^m|$.

- That is: More than half of Γ^m is in Σ^m.

Algorithm

- For each $f_a \in F^{m-1}$, find α_a such that $s_m = \alpha_a$ implies $\text{fail}(f) > m$.
- For each $b \in \Gamma^m$ choose some $a \in \sigma^{m-1}$ such that $a \preccurlyeq b$.
- b votes for α_a.
- If the conditions of the proposition are satisfied, a majority will vote for the correct value for s_m.

A bound on the minimimum distance

Here is the *order bound*, also called the *Feng-Rao bound*.

Proposition

The minimum distance of $C_{\bar{m}}$ is at least

$$d_{\bar{m}} = \min\{|N_m| : m > \bar{m}\}$$

One can also improve on the codes C_m by designing a code to have a specified minimum distance.
Let $M = \{m : |N_m| > d\}$ and let C be the code orthogonal to the space spanned by $\{ev(z_m) : m \in M\}$. The minimum distance of C is at least d.

Correction beyond the minimum distance bound

- The main theorem allows us to show that decoding well beyond half the minimum distance is possible for high rate Hermitian codes.
- Some examples:

# Check Symbols	Code [n,k,d]	Correction
10	[64, 54, 5]	3
36	[512, 476, 9]	10
48	[512, 464, 24]	13, 14
126	[4096,3970, 16]	34
192	[4096,3904, 80]	53
225	[4096,3871,112]	64

- An overwhelming proportion of vectors with weights less than the right hand column are correctable.

Generic points

- Suppose $\mathbb{F}$ is algebraically closed.
  A set V of t points will almost always have
  $\Delta_{I(V)} = \{0, 1, \ldots, t-1\}$ (this is an open condition).
  Call this "generic."
- For general $\mathbb{F}$, we may expect that "most" sets of t points will be generic.
- Experiments with Hermitian curves over $\mathbb{F}_{q^2}$ suggest the proportion sets of t points which are non-generic is $1/(q-1)$.
- The worst case scenario for t errors—those for which majority voting requires many check symbols—are exceedingly rare.
- Minimum distance is less important than the capability of decoding algorithm!

References

An extensive exposition of the subject is in

- Tom Høholdt, Jacobus H. van Lint, and Ruud Pellikaan. "Algebraic geometry of codes." In *Handbook of coding theory*, pages 871–961. North-Holland, Amsterdam, 1998. Vol. I.

Some recent results are found in

- Bras-Amorós, O'Sullivan, "The Correction Capability of the Berlekamp-Massey-Sakata Algorithm with Majority Voting," Applicable Algebra in Engineering, Communications and Computing 17 (2006), no. 5, 315–335.
- Geil, Pellikaan, "On the structure of order domains," Finite Fields Applic. 8 (2002) no. 3, 369-396.
- Little, The ubiquity of order domains for the construction of error control codes. Adv. Math. Commun. 1 (2007), no. 1, 151–171.
- O'Sullivan,"New codes for the Berlekamp-Massey-Sakata algeorithm," Finite Fields and Their Applications, vol. 7, pp. 293-317, 2001.

Original discoveries

- Sakata, "Finding a minimal set of linear recurring relations capable of generating a given finite two-dimensional array." J. Symbolic Comput. 5 (1988), no. 3, 321–337.
- Feng, Rao, Decoding algebraic-geometric codes up to the designed minimum distance. IEEE Trans. Inform. Theory 39 (1993), no. 1, 37–45.
- Feng, Rao, "Improved geometric Goppa codes. I. Basic theory." Special issue on algebraic geometry codes. IEEE Trans. Inform. Theory 41 (1995), no. 6, part 1, 1678–1693.
- Kirfel, Pellikaan, "The minimum distance of codes in an array coming from telescopic semigroups." IEEE Trans. Inform. Theory 41 (1995), no. 6, part 1, 1720–1732.