Linear Error Correcting Algebraic Geometry Codes

Abstract

In modern communication protocols, error correcting codes are used to compensate for loss of data in a noisy channel. A message is encoded, not to keep it secret (encryption), but to be able to recover the message in case a part of the received message is erroneous. We will see that Algebraic Geometry can be used to construct linear error correcting codes with good properties. We show how the theory of nonsingular projective algebraic curves, or equivalently their algebraic function fields, is used for constructing these codes and computing good lower bounds for their parameters.

Linear Codes

Definition 1 A *linear error correcting code over* $\mathbf{F}_q$ of length n and dimension k is a $\mathbf{F}_q$-linear subspace of $\mathbf{F}_q^n$ of dimension k.

Definition 2 Let $v \in \mathbf{F}_q^n$. We define the weight $w(v)$ of v as the number of nonzero coordinates of v. In other words

$$w(v) = |\{1 \leq i \leq n \quad v_i \neq 0\}.$$

Definition 3 Let C be a linear error correcting code over $\mathbf{F}_q$ of length n and dimension k. The *minimum distance* d of C is defined by

$$d = \min\{w(v) \mid v \in C \text{ and } v \neq 0\}|.$$

We will call C a linear $[n, k, d]_q$-code.

Example 4 The space $\mathbf{F}_q^n$ is a $[n, n, 1]_q$-code. The space with basis $(1, 1, 1, 1)$ is a $[4, 1, 4]_q$-code. This is also called the repetition code of length four.

An $[n, k, d]_q$-code C contains exactly q^k different code words. Hence it can be used to transmit q^k different messages. We identify the message space with $\mathbf{F}_q^k$. Encoding of a message is done by a linear isomorphism $Enc : \mathbf{F}_q^k \to C \subset \mathbf{F}_q^n$. The elements of $C = \mathrm{Im}(Enc)$ are called code words. After transmitting and receiving these code words a number of t errors may have occurred. If t is not too large, all of these errors can be corrected. In other words, from the received message containing not too many errors, a unique code word can be computed. This is called decoding.

Proposition 5 *Let C be a $[n,k,d]_q$-code. A received message $m \in \mathbf{F}_q^n$, containing t errors, can be uniquely decoded to a message of C, if and only if $t \leq \frac{d-1}{2}$.*

Proof. We define the distance of two code words x and y as $d(x,y) = w(x-y)$. It is easy to check that this is indeed a metric on $\mathbf{F}_q^n$. Because the code is linear, we have that for all $x, y \in C$ also $x - y \in C$. Therefore

$$\min\{d(x,y) \mid x \neq y \in C\} = \min\{w(x) \mid 0 \neq x \in C\} = d.$$

This says that with the above notion of distance, all the spheres centered at vectors in C with radius less than $\frac{d-1}{2}$ are disjoint.

Suppose m contains t errors and can be uniquely decoded to a message $x \in C$. Then $x = m - e$ for some error vector e of weight $w(e) = t$. Because m can be uniquely decoded to x, it must lie in the sphere of radius $\frac{d-1}{2}$ around x. Therefore $t = w(e) = w(m-x) = d(m,x) \leq \frac{d-1}{2}$.

Conversely, $x = m - e$ for some error vector e of weight $w(e) = t$. If $t \leq \frac{d-1}{2}$, then $d(m,x) = w(m-x) = w(e) = t \leq \frac{d-1}{2}$. Therefore m lies in the sphere of radius $\frac{d-1}{2}$ around x. Therefore it can be uniquely decoded to a message $x \in C$. □

Generalized Algebraic Geometry Codes

To define a generalized algebraic geometry code we need the following:

- A finite field $\mathbb{F}_q$ and an algebraic function field $F/\mathbb{F}_q$ of genus g.
- A natural number N. Distinct places $P_1, \cdots P_N$ of F, each P_i has degree k_i. We define the divisor $D = P_1 + P_2 + \ldots P_N$.
- A set of linear codes $C_1, \ldots, C_N$ over $\mathbb{F}_q$, each C_i has parameters $[n_i, k_i, d_i]$. We set $n = \sum_{i=1}^{N} n_i$.
- For every place P_i a $\mathbb{F}_q$-linear isomorphism $\phi_i : F_{P_i} \to C_i \subset \mathbb{F}_q^{n_i}$, where $F_{P_i} = \mathcal{O}_{P_i}/P_i$ is the residue class field of P_i.
- A divisor G, such that $\mathrm{supp}(G) \cap \mathrm{supp}(D) = \emptyset$.

We define a map $\phi : \mathcal{L}(G) \to \mathbb{F}_q^n$ by setting

$$\phi(z) = (\phi_1(z(P_1)), \phi_2(z(P_2)), \ldots, \phi_N(z(P_N))) \ \in \mathbf{F}_q^n \tag{1}$$

for all $z \in \mathcal{L}(G)$.

Definition 6 The image of ϕ is called a *generalized algebraic geometry code* and we denote it by $C(D, G, \phi)$.

Parameters of a generalized algebraic geometry code

Proposition 7 *The code $C(D, G, \phi)$ is a linear $[n, k, d]$-code over $\mathbb{F}_q$ with*

- $n = \sum_{i=1}^N n_i$
- $k = \dim \mathcal{L}(G) \geq \deg G + 1 - g \qquad$ *if* $\deg G < \sum_{i=1}^N k_i$
- $d \geq \min\{\sum_{i \notin S} d_i \mid S \subseteq \{1, \ldots, N\}$ *is such that* $\sum_{i \in S} k_i \leq \deg G\}$.

Proof. Since every ϕ_i has image in $\mathbb{F}_q^{n_i}$, the image of ϕ is in $\prod_{i=n}^N \mathbb{F}_q^{n_i} \cong \mathbb{F}_q^n$, so the code has length n. Furthermore, ϕ is a linear map, because ϕ_i is linear for all i. So the code is linear.

For $z \in \mathcal{L}(G)$, $\phi(z)$ can only be zero if $\phi_i(z(P_i)) = 0$ for every i. Since ϕ_i is injective, this is the same as saying that $v_{P_i}(z) \geq 1$ for all $P_i \in \operatorname{supp} D$. Hence $\operatorname{Ker} \phi = \mathcal{L}(G - D)$. If $\deg G < \sum_{i=1}^N k_i = \deg D$, then $\deg(G - D) < 0$, hence ϕ is injective and therefore its image has the same dimension as $\mathcal{L}(G)$.
The inequality $\dim \mathcal{L}(G) \geq \deg G + 1 - g$ is a consequence of the Riemann-Roch Theorem.

The last part asks us to find a lower bound for d. Because the code is linear $d = \min\{w(\phi(z)) \mid z \in \mathcal{L}(G)\}$. Now $w(\phi(z)) = w((\phi_1(z), \ldots, \phi_N(z))) = \sum_{i \in R_z} w(\phi_i(z(P_i))) \geq \sum_{i \in R_z} d_i$, where $R_z = \{i \in \{1, \ldots, N\} \mid z(P_i) \neq 0\}$. Here the last equality follows from the fact that for $i \notin R_z$, $\phi_i(z(P_i))$ has weight zero and does not contribute to the sum. The inequality holds, because ϕ_i maps invectively into a $[n_i, k_i, d_i]$-code and therefore any nonzero code word has weight at least d_i. Now we have that $d \geq \min\{\sum_{i \in R_z} d_i \mid z \in \ \mathcal{L}(G) \setminus \{0\}\}$.
We set $S = \{1, \ldots, N\} \setminus R_z = \{i \in \{1, \ldots, N\} \mid z(P_i) = 0\}$ and we will show that $\sum_{i \in S} k_i \leq \deg G$. For all $i \in S$, $v_{P_i}(z) \geq 1$, so $z \in \mathcal{L}(G - \sum_{i \in S} P_i)$, so $\dim(G - \sum_{i \in S} P_i) \neq 0$. Now proposition I.4.2b implies that $\deg(G - \sum_{i \in S} P_i) \geq 0$, hence we can conclude that $\deg G - \sum_{i \in S} \deg P_i \geq$

0. Using $k_i = \deg P_i$ we find $\sum_{i \in S} k_i \leq \deg G$.
Finallly we note that the set $\{S \subseteq \{1, \ldots, N\},\ S = \{1, \ldots, N\} \setminus R_z \mid z \in \mathcal{L}(G)\}$ is a subset of the set $\{S \subseteq \{1, \ldots, N\} \mid S$ is such that $\sum_{i \in S} k_i \leq \deg G\}$. Since the minimum will not increase, when S runs through a bigger set, we get the result. □

We can find another lower bound for the minimum distance of generalized algebraic geometry code.

Lemma 8 *Let d be the minimum distance of a generalized algebraic geometry code. Then*

$$d \geq \min\{\sum_{i=1}^{N} d_i - \deg G - \sum_{i \in S}(d_i - k_i)\}$$

where the minimum is taken over all sets $S \subseteq \{1, \ldots, N\}$ such that $\sum_{i \in S} k_i \leq \deg G$.

Proof. Assume $S \subseteq \{1, \ldots, N\}$ is such that $-\sum_{i \in S} k_i \geq -\deg G$. Then $\sum_{i \notin S} d_i = \sum_{i=1}^{N} d_i - \sum_{i \in S} d_i + \sum_{i \in S}(k_i - k_i) = \sum_{i=1}^{N} d_i - \sum_{i \in S} k_i - \sum_{i \in S}(d_i - k_i) \geq \sum_{i=1}^{N} d_i - \deg G - \sum_{i \in S}(d_i - k_i)$. The result now follows from proposition 7. □

Corollary 9 *If $d_i = k_i$ for all i, then $d \geq \deg D - \deg G$.*

Proof. This follows from substituting $\sum_{i=1}^{N} d_i = \sum_{i=1}^{N} k_i = \deg D$ in the above lemma. □

Corollary 10 *Let $C(D, G, \phi)$ be a generalized algebraic geometry code with $k_u = \min\{k_i \mid 1 \leq i \leq N\}$ and $d_v = \min\{d_i \mid 1 \leq i \leq N\}$ and minimum distance d. Then*

$$d \geq d_v(N - \frac{\deg G}{k_u}).$$

Proof. By proposition 7 $d \geq \min\{\sum_{i \notin S} d_i\} \geq (N - |S|)d_v$, where the minimum is taken over all sets S for which $\sum_{i \in S} k_i \leq \deg G$. The sum is minimal when $|S|$ is maximal. We find an upperbound for $|S|$. We have $|S|k_u = \sum_{i \in S} k_u \leq \sum_{i \in S} k_i \leq \deg G$. So $|S| \leq \frac{\deg G}{k_u}$ and therefore $d \geq (N - |S|)d_v \geq d_v(N - \frac{\deg G}{k_u})$. □

Lemma 11 *Let C be a generalized algebraic geometry code with $k_1 = \min\{k_i \mid 1 \leq i \leq N\}$. Let $z \in \mathcal{L}(G)$ and let d' be the number of indices i for which $\phi_i(z(P_i))$ is nonzero. Then $d' \geq N - \frac{\deg G}{k_1}$.*

Proof. From the proof of proposition 7 we have $d' = |R_z|$ and $d \geq \min_z\{\sum_{i \in R_z} d_i\}$. Therefore taking $d_i = 1$ for all i we find that $d' \geq \min\{\sum_{i \notin S} 1 \mid S \subseteq \{1, \ldots, N\}$ is such that $\sum_{i \in S} k_i \leq \deg G\}$. Now corollary 10 with $d_1 = 1$ gives the result. □

Relation with Goppa codes

Let us recall the definition of a Goppa code. Let $F/\mathbf{F}_q$ be a function field, $P_1, \ldots, P_N$ distinct rational places of F of degree one. $D = P_1 + \cdots + P_N$. Let G be a divisor of F such that $\mathrm{Supp}\ G \cap \mathrm{Supp}\ D = \emptyset$.

Definition 12 The *geometric Goppa code* associated with the divisors D and G is defined by

$$C(D, G) = \{(z(P_1), z(P_2), \ldots, z(P_N)) \mid z \in \mathcal{L}(G)\} \subset \mathbf{F}_q^n.$$

We will show that generalized algebraic geometry codes generalize the definition of a Goppa code. Let F, D, G be as above. For each i we define the linear code C_i to be the $[1, 1, 1]_q$-code $\mathbb{F}_q$. For every place P_i the residue field $F_{P_i} = \mathbf{F}_q$. Therefore we can take $\phi_i : F_{P_i} \to C_i$ the identity map. The generalized algebraic geometry code $C(D, G, \phi)$ is equal to the Goppa code $C(D, G)$. From proposition 7 this is a code of length $n = N$ and if $\deg G < N$ the dimension is equal to the dimension of $\mathcal{L}(G)$. Corollary 9 tells us that the minimum distance is d with $d \geq N - \deg G$.