MAD 4471: Cryptography and Coding Theory **Fall 2020**

Lecture 10: Linear Codes

Lecturer: Jean-François Biasse *TA: William Youmans*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

We gave bounds on the theoretical possibility of detecting/decoding errors, but we did not address the problem of performing this task efficiently. In this section, we introduce a secial class of codes for which error detection/correction is easy: The linear codes.

10.1 Basic properties of linear codes

Definition 10.1 (Linear codes) *A (m, k) code over a field $\mathbb{F}$ is a linear subspace of $\mathbb{F}^m$ of dimension k.*

Example 1 *Let*

$$G = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

and $\mathbb{F} = \mathbb{F}_2$. Then the linear span of the rows of G (i.e., all the possible linear combinations of the rows of G) is a (6,2) code. Its codewords are

$$(1, 0, 1, 0, 1, 0)$$

$$(0, 1, 0, 1, 0, 1)$$

$$(1, 1, 1, 1, 1, 1)$$

$$(0, 0, 0, 0, 0, 0)$$

The important parameters quantifying how many errors we can handle is $d(\mathcal{C})$.

Definition 10.2 (Hamming weight) *The Hamming weight $w(v)$ of $v \in \mathbb{F}^m$ is the number of non-zero entries of v. Incidentally, it is also $w(v) = d(v, (0, ..., 0))$.*

Proposition 10.3 *Let $\mathcal{C}$ be a linear code, then*

$$d(\mathcal{C}) = \min\{w(v) \mid v \neq 0 \in \mathcal{C}\}$$

Proof: Let $u, v \in \mathcal{C}$, $u \neq v$ such that $d(u, v) = d(\mathcal{C})$. Then $u - v \in \mathcal{C}$ and

$$d(u, v) = d(u - v, (0, ..., 0)) = w(u - v) \geq \min\{w(x) \mid x \neq 0 \in \mathcal{C}\}$$

So $d(\mathcal{C}) \geq \min\{w(v) \mid v \neq 0 \in \mathcal{C}\}$. Also, let $u \in \mathcal{C}$ such that $w(u) = \min\{w(x) \mid x \neq 0 \in \mathcal{C}\}$. Then $w(u) = d(u, (0, ..., 0)) \geq d(\mathcal{C})$ because $u \in \mathcal{C}$, $(0, ..., 0) \in \mathcal{C}$. So $d(\mathcal{C}) \leq \min\{w(v) \mid v \neq 0 \in \mathcal{C}\}$. ∎

Example 2 *The code defined by the span of the rows of*

$$G = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

has distance 3.

Definition 10.4 *The matrix G such that $\mathcal{C}$ is the span of the rows of G is called the generating matrix of $\mathcal{C}$. If $G = (I_h\ P)$ where I_h is the $h \times h$ identity matrix: 1's on the diagonal, 0's everywhere else. P is a $h \times (m-h)$ matrix. $\mathcal{C}$ is said to be systematic.*

Let $G_1, ..., G_h$ be the rows of G. A codeword is of the form

$$c = \lambda_1 G_1 + \cdots \lambda_h G_h = (\lambda_1, ..., \lambda_h, ?\,, ..., ?\,)$$

where $\lambda_1, ..., \lambda_h$ are information symbols and ?,...,? are check symbols.

Definition 10.5 *A matrix H such that*

$$vH^T = 0 \iff v \in \mathcal{C}$$

is a parity check matrix for $\mathcal{C}$.

Proposition 10.6 *If $G = (I_h\ P)$ is a generating matrix for $\mathcal{C}$, then $H = (-P^T\ I_{m-h})$ is a parity check matrix for $\mathcal{C}$.*

Proof: We start by proving that each row G_i of G satisfies $G_i H^T = 0$. H^T has the form:

$$\begin{pmatrix} -P \\ I_{n-h} \end{pmatrix}$$

The coordinate j of $G_i \cdot H^T$ is the dot product of $G_i = (0, \cdots, 0, 1, 0, \cdots, 0, P_{i,1}, \cdots, P_{i,n-k})$ where 1 in i th position, $H_j = (-P_{1,j}, \cdots, -P_{i,j}, \cdots, -P_{h,j}, 0, \cdots, 1, 0, \cdots, 0)$ where H_j is j th column of H^T, i.e. j th row of H, it equals $-P_{il} + P_{ij} = 0$.

Moreover, the column rank of H^T is $n-k$ while the sign of its columns is n. So by the rank-nullity theorem, the dimension of the left nullspace $Null(H^T)$ of H^T is $dim(Null(H^T)) = n - (n-k) = k$.

We found k independent vectors G_i in $Null(H^T)$, therefore they are a basis of $Null(H^T)$.
So

$$\begin{aligned} v \cdot H^T = 0 &\Leftrightarrow v \in Null(H^T) \\ &\Leftrightarrow v \in Span(G_i) \\ &\Leftrightarrow v \in \mathcal{C} \end{aligned}$$

■

Example 3 *With*

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

,

$$H^T = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and we see that $G_1 \cdot H^T = 0, G_2 \cdot H^T = 0$.

Use for decoding: We can use this property to determine if $\mu \in \mathbb{K}^n$ is in $\rfloor$, but we can also find the nearest $c \in \mathcal{C}$ to the received message $r \in \mathbb{K}^n$!

Before we show how to decode, we need to introduce the notion of coset.

Definition 10.7 (coset) *Let* $\mathcal{C}$ *be a linear code, and let* $\mu \in \mathbb{K}^n$.

- *The sets of the form* $\mu + \mathcal{C} := \{\mu + c | c \in \mathcal{C}\}$ *are called cosets of* $\mathcal{C}$.
- *Any* $\mu^{'} \in \mu + \mathcal{C}$ *is a member of the coset, and the one with the smallest Hamming weight is called a coset leader.*

Definition 10.8 *The syndrome of* $\mu \in \mathbb{K}^n$ *is* $S(\mu) = \mu H^T$.

Lemma 10.9 *Two vectors* μ *and* v *belongs to the same coset if and only if they have the same syndrome.*

Proof: Clearly, $c \in \mathcal{C} \Leftrightarrow S(c) = 0$. moreover, μ and v belongs to the same coset if and only if $\mu - v \in \mathcal{C}$. So:

$$\begin{aligned}(\mu \text{ and } v \text{ belongs to the same coset}) &\Leftrightarrow \mu - v \in \mathcal{C} \\ &\Leftrightarrow S(\mu - v) = 0 \\ &\Leftrightarrow S(\mu) - S(v) = 0 \\ &\Leftrightarrow S(\mu) = S(v)\end{aligned}$$

■

To decode r, we simply observe that we want to find the smallest $\mu \in \mathbb{K}^n$ such that $r - \mu \in \mathcal{C}$ (Otherwise stated: the smallest pertubation of r leading to a code word).
This is by definition the coset leader of $r + \mathcal{C}$.
So we keep a lookup table of all the coset leaders with their syndromes and we follow the procedure:

- Calculate the syndrome $S(r) = rH^T$ of the received vector r.
- Find the coset leader G with the same syndrome as $S(r)$
- Return $r - c_0 \in \mathcal{C}$.

Example 4 *For*

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

,

$$H = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

There are four different cosets and we have the lookup table

Coset leader	*syndrome*
$(0,0,0,0)$	$(0,0)$
$(1,0,0,0)$	$(1,1)$
$(0,1,0,0)$	$(1,0)$
$(0,0,0,1)$	$(0,1)$

Suppose we receive $r = (0,1,0,1)$*, we first calculate its syndrome* $S(r) = rH^T = (1,1)$*. Then we decode* $r - \underbrace{(1,0,0,0)} = (1,1,0,1) \in \mathcal{C}$*.*
coset leader with syndrome $(1,1)$*.*

10.2 Hamming codes

The Hamming codes are a class of linear codes with distance $d = 3$, which means that the decoding procedure works under the assumption that the number of errors does not exceed 1.
In the case of a single error, the decoding procedure can be simplified. Supposed H is the parity check matrix of $\mathcal{C}$, and suppose we receive $r \in \mathbb{K}^n$.

1. If $rH^T = 0$, then return $r \in \mathcal{C}$.
2. Otherwise $r = c + e_i$ where $c \in \mathcal{C}, e_i = (0,0,\cdots,0,1,0,\cdots,0)$ where 1 in i th position.
   $rH^T = cH^T + e_iH^T = 0 + e_iH^T = \text{i th column of} H$

Flip the coefficient i of r and return the corresponding vector. Now, how do we construct a Hamming code? We start by building H. The two parameters are m and $n = 2^m - 1$.
We first build the $m \times n$ matrix whos columns are the bits of $1, 2, \cdots, 2^m - 1$.

Example 5 *Ex. for* $m = 3, n = 7$*:*

$$M = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Then we reorder the columns such that the last block be the $m \times m$ *identity matrix, and that gives us* H*.*

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} = (-p^T, I_m)$$

From that, we deduce the generality matrix $G = (I_h \ P)$ *where* $h = 2^n - 1 - m$*.*

Proposition 10.10 *The distance of a Hamming code is* 3*.*

Proof: The generating matrix is of the form $G = (I_h \ P)$. Each row of P is the bit vector of an integer that has between 2 and m no zero bits. Therefore some of the rows of G have Hamming weight 3 (Since the rows of I_h have Hamming weight 1). Moreover, we cannot obtain a linear combination of the rows of G with weight less than 3 because rows of P cannot add to the 0 vector. Any linear combination of more rows of G will have at least weight 3 on the first k coordinates.
So $d(\mathcal{C}) = min\{W(\mu), \mu t \ \mathcal{C}\} = 3$. ∎