Coding Theory: Linear Error-Correcting Codes

Anna Dovzhik

April 23, 2014
1. Coding Theory
   - Basic Definitions
   - Error Detection and Correction

2. Finite Fields

3. Linear Codes
   - Hamming Codes
   - Finite Fields Revisited
   - BCH Codes
   - Reed-Solomon Codes

4. Conclusion
Basic Definitions

Definition
If $A = a_1, a_2, \ldots, a_q$, then $A$ is a code alphabet of size $q$. 
Basic Definitions

Definition
If $A = a_1, a_2, \ldots, a_q$, then $A$ is a code alphabet of size $q$.

Definition
A $q$-ary word $w = w_1 w_2 w_3 \ldots w_n$ is a vector where $w_i \in A$. 
Basic Definitions

Definition
If \( A = a_1, a_2, \ldots, a_q \), then \( A \) is a code alphabet of size \( q \).

Definition
A q-ary word \( \mathbf{w} = w_1 w_2 w_3 \ldots w_n \) is a vector where \( w_i \in A \).

Definition
A q-ary block code is a set \( C \) over an alphabet \( A \), where each element, or codeword, is a q-ary word of length \( n \).
Definition

For two codewords, $w_1, w_2$, over the same alphabet, the \textbf{Hamming distance}, denoted $d(w_1, w_2)$, is the number of places where the two vectors differ.
Basic Definitions

Definition
For two codewords, $w_1, w_2$, over the same alphabet, the Hamming distance, denoted $d(w_1, w_2)$, is the number of places where the two vectors differ.

Definition
For a code $C$, the minimum distance is denoted $d(C) = \min\{d(w_1 w_2) : w_1, w_2 \in C, w_1 \neq w_2\}$. 
Basic Definitions

Definition
For two codewords, \(w_1, w_2\), over the same alphabet, the Hamming distance, denoted \(d(w_1, w_2)\), is the number of places where the two vectors differ.

Definition
For a code \(C\), the minimum distance is denoted \(d(C) = \min\{d(w_1w_2) : w_1, w_2 \in C, w_1 \neq w_2\}\).

Definition
For a codeword \(w\), the Hamming weight of \(w\), or \(wt(w)\), is the number of nonzero places in \(w\). That is, \(wt(w) = d(w, 0)\).
Example

**Notation:** A $q$-ary $(n, M, d)$-code
Notation: A $q$-ary $(n, M, d)$-code

Example

- A binary (3,4,2)-code
- $A = F_2 = \{0, 1\}$
- $C = \{000, 011, 110, 101\}$
Example

**Notation:** A $q$-ary $(n, M, d)$-code

**Example**

- A binary (3,4,2)-code
- $A = \mathbb{F}_2 = \{0, 1\}$
- $C = \{000, 011, 110, 101\}$

The main coding theory problem: optimizing one parameter when others are given.
Errors

- vector received is not a codeword
- $x$ is sent, but $y$ is received $\rightarrow e = x + y$
- To detect $e$, $x + e$ cannot be a codeword
Errors

- vector received is not a codeword
- \( \mathbf{x} \) is sent, but \( \mathbf{y} \) is received \( \rightarrow \mathbf{e} = \mathbf{x} + \mathbf{y} \)
- To detect \( \mathbf{e} \), \( \mathbf{x} + \mathbf{e} \) cannot be a codeword

Example

Binary \((3,3,1)\)-code \( C = \{001, 101, 110\} \)

- \( \mathbf{e}_1 = 010 \) can be detected \( \rightarrow \) for all \( \mathbf{x} \in C, \mathbf{x} + \mathbf{e}_1 \not\in C \)
- \( \mathbf{e}_2 = 100 \) cannot be detected \( \rightarrow 001 + 100 = 101 \in C \)
Error Detection

Definition

A code is **u-error-detecting** if when a codeword incurs between one to $u$ errors, the resulting word is not a codeword.
Error Detection

Definition
A code is **u-error-detecting** if when a codeword incurs between one to $u$ errors, the resulting word is not a codeword.

Theorem
*A code is $u$-error-detecting if and only if $d(C) \geq u + 1$.***
Error Detection

**Definition**

A code is *u-error-detecting* if when a codeword incurs between one to *u* errors, the resulting word is not a codeword.

**Theorem**

*A code is u-error-detecting if and only if* $d(C) \geq u + 1$.

**Proof.**

($\Leftarrow$) Any error pattern of weight at most *u* will alter a codeword into a non-codeword.

($\Rightarrow$) Suppose that for $x, y \in C$, $d(x, y) \leq u$. Let $e = x + y$, $wt(e) \leq u$, and $x + e = x + x + y = y$, which is a codeword. Therefore, $e$ cannot be detected. ($\Rightarrow$)($\Leftarrow$)
Error Correction

- $e + x$ is closer to $x$ than any other codeword
- evaluate minimum distances
Error Correction

- $e + x$ is closer to $x$ than any other codeword
- evaluate minimum distances

Definition

A code is **v-error-correcting** if $v$ or fewer errors can be corrected by decoding a transmitted word based on minimum distance.
Error Correction

- $e + x$ is closer to $x$ than any other codeword
- evaluate minimum distances

**Definition**

A code is **v-error-correcting** if $v$ or fewer errors can be corrected by decoding a transmitted word based on minimum distance.

**Theorem**

A code is $v$-error-correcting if and only if $d(C) \geq 2v + 1$. That is, if $C$ has a distance $d$, it corrects $\frac{d-1}{2}$ errors.
Finite Fields

Definition

A field is a nonempty set \( F \) of elements satisfying:

- operations addition and multiplication
- eight axioms
  - closure under addition and multiplication
  - commutativity of addition and multiplication
  - associativity of addition and multiplication
  - distributivity of multiplication over addition
  - additive and multiplicative identities
  - additive and multiplicative inverses
Finite Fields

Definition

A **field** is a nonempty set $F$ of elements satisfying:

- operations addition and multiplication
- eight axioms
  - closure under addition and multiplication
  - commutativity of addition and multiplication
  - associativity of addition and multiplication
  - distributivity of multiplication over addition
  - additive and multiplicative identities
  - additive and multiplicative inverses

Binary field - arithmetic mod 2

<table>
<thead>
<tr>
<th>+</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\cdot$</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Finite Fields

Theorem

$\mathbb{Z}_p$ is a field if and only if $p$ is a prime.
Finite Fields

Theorem

\( \mathbb{Z}_p \) is a field if and only if \( p \) is a prime.

Definition

Denote the multiplicative identity of a field \( F \) as 1. Then characteristic of \( F \) is the least positive integer \( p \) such that 1 added to itself \( p \) times is equal to 0. This characteristic must be either 0 or a prime number.
Finite Fields

**Theorem**

\( \mathbb{Z}_p \) is a field if and only if \( p \) is a prime.

**Definition**

Denote the multiplicative identity of a field \( F \) as 1. Then characteristic of \( F \) is the least positive integer \( p \) such that 1 added to itself \( p \) times is equal to 0. This characteristic must be either 0 or a prime number.

**Theorem**

A finite field \( F \) of characteristic \( p \) contains \( p^n \) elements for some integer \( n \geq 1 \).
A linear \((n, k, d)\)-code \(C\) over a finite field \(\mathbb{F}_q\) is a subspace of the vector space \(\mathbb{F}_q^n\).
Linear Codes

- A linear \((n, k, d)\)-code \(C\) over a finite field \(\mathbb{F}_q\) is a subspace of the vector space \(\mathbb{F}_q^n\)
- Codewords are linear combinations \(\left(q^k\right)\) distinct codewords
Linear Codes

- A linear \((n, k, d)\)-code \(C\) over a finite field \(\mathbb{F}_q\) is a subspace of the vector space \(\mathbb{F}_q^n\)
- Codewords are linear combinations \((q^k\) distinct codewords)

**Definition**

A matrix whose rows are the basis vectors of a linear code is a **generator matrix**.
Linear Codes

- A linear \((n, k, d)\)-code \(C\) over a finite field \(\mathbb{F}_q\) is a subspace of the vector space \(\mathbb{F}_q^n\)
- Codewords are linear combinations \((q^k\) distinct codewords)

**Definition**

A matrix whose rows are the basis vectors of a linear code is a **generator matrix**.

**Definition**

Two \(q\)-ary codes are **equivalent** if one can be obtained from the other using a combination of the operations

- permutation of the positions of the code (column swap)
- multiplication of the symbols appearing in a fixed position (row operation)
Definition

If $C$ is a linear code in $\mathbb{F}_q^n$, then the dual code of $C$ is $C^\perp$.

Definition

A **parity-check matrix** is a generator matrix for the dual code.
Definition

If $C$ is a linear code in $\mathbb{F}_q^n$, then the dual code of $C$ is $C^\perp$.

Definition

A parity-check matrix is a generator matrix for the dual code.

- $C$ is a $(n, k, d)$-code $\rightarrow$ generator matrix $G$ is $k \times n$ and parity-check matrix $H$ is $(n - k) \times n$.
- The standard form of $G$ is $(I_k|A)$ and the standard form of $H$ is $(B|I_{n-k})$. 
Theorem

If $C$ is a $(n, k)$-code over $\mathbb{F}_p$, then $\mathbf{v}$ is a codeword of $C$ if and only if it is orthogonal to every row of the parity-check matrix $H$, or equivalently, $\mathbf{v}H^T = \mathbf{0}$.

This also means that $G$ is a generator matrix for $C$ if and only if the rows of $G$ are linearly independent and $GH^T = \mathbf{0}$.

Proof: orthogonality
Theorem

If \( G = (I_k|A) \) is the standard form of the generator matrix for a \((n, k, d)\)-code \( C \), then a parity-check matrix for \( C \) is
\[
H = (-A^T | I_{n-k}).
\]

Note that if the code is binary, negation is unnecessary.
Theorems

Theorem

For a linear code $C$ and a parity-check matrix $H$,

- $C$ has distance $\geq d$ if and only if any $d - 1$ columns of $H$ are linearly independent
- $C$ has distance $\leq d$ if and only if $H$ has $d$ columns that are linearly dependent.

So, when $C$ has distance $d$, any $d - 1$ columns of $H$ are linearly independent and $H$ has $d$ columns that are linearly dependent.

Proof: orthogonality
Recall the \textit{main coding theory problem}
Bounds

Recall the main coding theory problem

Definition

A $q$-ary code is a **perfect code** if it attains the Hamming, or sphere-packing bound. For $q > 1$ and $1 \leq d \leq n$, this is defined as having

$$\frac{q^n}{\sum_{i=0}^{\lfloor(d-1)/2\rfloor} \binom{n}{i} (q - 1)^i}$$

codewords.
Theorem

When \( q \) is a prime power, the parameters \((n, k, d)\) of a linear code over \( \mathbb{F}_q \) satisfy \( k + d \leq n + 1 \). This upper bound is known as the **Singleton bound**.
Bounds

**Theorem**

When $q$ is a prime power, the parameters $(n, k, d)$ of a linear code over $\mathbb{F}_q$ satisfy $k + d \leq n + 1$. This upper bound is known as the **Singleton bound**.

**Definition**

A $(n, k, d)$ code where $k + d = n + 1$ is a **maximum distance separable code** (MDS) code.
**Bounds**

**Theorem**
When $q$ is a prime power, the parameters $(n, k, d)$ of a linear code over $\mathbb{F}_q$ satisfy $k + d \leq n + 1$. This upper bound is known as the *Singleton bound*.

**Definition**
A $(n, k, d)$ code where $k + d = n + 1$ is a maximum distance separable code (MDS) code.

**Theorem**
If a linear code $C$ over $\mathbb{F}_q$ with parameters $(n, k, d)$ is MDS, then:
$C^\perp$ is MDS, every set of $n - k$ columns of $H$ is linearly independent, every set of $k$ columns of $G$ is linearly independent.
Hamming Codes

- single error-correcting
- double error-detecting codes
- easy to encode and decode
Hamming Codes

- single error-correcting
- double error-detecting codes
- easy to encode and decode

Definition

The **binary Hamming code**, denoted $\text{Ham}(r, 2)$, has a parity-check matrix $H$ whose columns consist of all nonzero binary codewords of length $r$

For a non-binary finite field $F_q$, the $q$-ary Hamming code is denoted as $\text{Ham}(r, q)$
Properties for both Ham($r, 2$) and Ham($r, q$):

- perfect code
- $k = 2^r - 1 - r$, where $k$ denotes dimension
- more generally, $k = \frac{q^r - 1}{q - 1}$
- $d = 3$, where $d$ denotes distance
- exactly single-error-correcting
Hamming Codes

Ham(3, 2) code
Constructing the parity-check matrix

Decoding Hamming
Decoding Hamming

Ham(3, 2) code

Constructing the parity-check matrix

- all binary Hamming codes of a given length are equivalent
- arrange the columns of $H$ in order of increasing binary numbers

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

**Ham(3, 2) code**
Suppose \( y = (1101011) \) is received
Decoding Hamming

**Ham(3, 2) code**
Suppose $y = (1101011)$ is received

$$yH^T = (1101011) \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = (110)$$

- error is in the sixth position of $y$
- $y$ is corrected to $(1101001)$
Decoding Hamming

**Ham(3, 2) code**

Suppose \( y = (1101011) \) is received

\[
\begin{pmatrix}
0 & 0 & 1 \\
0 & 1 & 0 \\
0 & 1 & 1 \\
1 & 0 & 0 \\
1 & 0 & 1 \\
1 & 1 & 0 \\
1 & 1 & 1 \\
\end{pmatrix}
\begin{pmatrix}
1101011 \\
\end{pmatrix} = (110)
\]

- error is in the sixth position of \( y \)
- \( y \) is corrected to \( (1101001) \)
Encoding Hamming

To derive $G$, recall that if $H = (-A^T|I_{n-k})$, $G = (I_k|A)$
Encoding Hamming

To derive $G$, recall that if $H = (-A^T | I_{n-k})$, $G = (I_k | A)$

To encode $x = 1101$:

$xG = (1101) \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1
\end{pmatrix} = (1101001)$

- encoded vector is $n+k$ digits long
- first $k$ digits (message digits) are the original vector
- last $n-k$ digits (check digits) represent redundancy
To derive $G$, recall that if $H = (-A^T | I_{n-k})$, $G = (I_k | A)$

To encode $x = 1101$:

$$xG = (1101) \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} = (1101001)$$

- encoded vector is $n + k$ digits long
- first $k$ digits (message digits) are the original vector
- last $n - k$ digits (check digits) represent redundancy
Finite Fields Revisited

Definition

For \( n \) polynomials in \( F_q[x] \), denoted \( f(x_1), f_2(x), \ldots, f_n(x) \), the least common multiple, denoted \( \text{lcm}(f(x_1), f_2(x), \ldots, f_n(x)) \) is the lowest degree monic polynomial that is a multiple of all the polynomials.
Finite Fields Revisited

**Definition**
For $n$ polynomials in $\mathbb{F}_q[x]$, denoted $f(x_1), f_2(x), \ldots, f_n(x)$, the **least common multiple**, denoted $\text{lcm}(f(x_1), f_2(x), \ldots, f_n(x))$ is the lowest degree monic polynomial that is a multiple of all the polynomials.

**Definition**
A **minimal polynomial** of an element in a finite field $\mathbb{F}_p$ is a nonzero monic polynomial of the least degree possible such that the element is a root.
Finite Fields Revisited

**Definition**
For $n$ polynomials in $\mathbb{F}_q[x]$, denoted $f(x_1), f_2(x), \ldots, f_n(x)$, the least common multiple, denoted $\text{lcm}(f(x_1), f_2(x), \ldots, f_n(x))$ is the lowest degree monic polynomial that is a multiple of all the polynomials.

**Definition**
A minimal polynomial of an element in a finite field $\mathbb{F}_p$ is a nonzero monic polynomial of the least degree possible such that the element is a root.

**Definition**
A primitive element or generator of $\mathbb{F}_p$ is an $\alpha$ such that $\mathbb{F}_q = \{0, \alpha, \alpha^2, \ldots, \alpha^{p-1}\}$ Every finite field has at least one primitive element, and primitive elements are not unique.
BCH Codes

- Generalization of Hamming codes for multiple-error correction
- Eliminate certain codewords from Hamming code
- Can be determined from a generator polynomial
BCH Codes

- Generalization of Hamming codes for multiple-error correction
- Eliminate certain codewords from Hamming code
- Can be determined from a generator polynomial

**Definition**

Suppose $\alpha$ is a primitive element of a finite field $\mathbb{F}_q^m$ and $M_i(x)$ is the minimal polynomial of $\alpha^i$ with respect to $\mathbb{F}_q$. Then a primitive BCH code over $\mathbb{F}_q$ of length $n = q^m - 1$ and distance $d$ is a $q$-ary cyclic code that is generated by the polynomial defined as $\text{lcm}(M_a(x), M^{a+1}(x), \ldots, M^{a+d-2}(x))$ for some $a$. 
Codewords and Polynomials

- One way to represent a codeword $c$ is with a binary polynomial $c(x)$, where $\alpha$ is a primitive element and $c(\alpha^k) = 0$.
- Given a codeword $c$ of length $n$, let the digits of $c$ be denoted $c = c_{n-1}, \ldots, c_1, c_0$, and define the polynomial $c(x)$ as

$$c(x) = \sum_{i=0}^{n-1} c_i x^i$$

Example

The BCH code of length 15, $00001\ 11011\ 00101$, corresponds to the polynomial $x^{10} + x^9 + x^8 + x^6 + x^5 + x^2 + 1$. 
Codewords and Polynomials

- One way to represent a codeword $c$ is with a binary polynomial $c(x)$, where $\alpha$ is a primitive element and $c(\alpha^k) = 0$.
- Given a codeword $c$ of length $n$, let the digits of $c$ be denoted $c = c_{n-1}, \ldots, c_1, c_0$, and define the polynomial $c(x)$ as
  
  $$c(x) = \sum_{i=0}^{n-1} c_i x^i$$

Example
The BCH code of length 15, 00001 11011 00101, corresponds to the polynomial $x^{10} + x^9 + x^8 + x^6 + x^5 + x^2 + 1$
Reed-Solomon Codes

- Subclass of BCH codes that can handle error-bursts
- MDS codes
Reed-Solomon Codes

- Subclass of BCH codes that can handle error-bursts
- MDS codes

**Definition**

A \( q \)-ary Reed-Solomon code is a \( q \)-ary BCH code of length \( q - 1 \) that is generated by 
\[
g(x) = (x - \alpha^{a+1})(x - \alpha^{a+2}) \ldots (x - \alpha^{a+d-1}), \text{ where } a \geq 0, \ 2 \leq d \leq q - 1, \text{ and } \alpha \text{ is a primitive element of } \mathbb{F}_q.
\]

Since the length of a binary RS code would be \( 2 - 1 = 1 \), this type of code is never considered.
Reed-Solomon Codes

Example

For a 7-ary RS code of length 6 and generator polynomial $g(x) = (x - 3)(x - 3^2)(x - 3^3) = 6 + x + 3x^2 + x^3$, 

$$G = \begin{pmatrix} 6 & 1 & 3 & 1 & 0 & 0 \\ 0 & 6 & 1 & 3 & 1 & 0 \\ 0 & 0 & 6 & 1 & 3 & 1 \end{pmatrix}$$

$$H = \begin{pmatrix} 1 & 4 & 1 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 & 1 \end{pmatrix}$$
Applications

Any case where data is transmitted through a channel that is susceptible to noise

- digital images from deep-space
- compact disc encoding
- radio communications
References


