Linear Codes

Linear Codes

In the V[n,q] setting, an important class of codes are the ***linear codes***, these codes are the ones whose code words form a sub-vector space of V[n,q]. If the subspace of V[n,q] is k dimensional then we talk about the subspace as an **[n,k]*-code***. (Note that the square brackets indicate a linear code).

In the V[n,q] setting, the terms “word” and “vector” are interchangeable.

Linear codes, because of their algebraic properties, are the most studied codes from a mathematical point of view. Our text (and many others) is devoted almost exclusively to linear codes.

Linear Codes

There are several consequences of a code being linear.

1) The sum or difference of two codewords is another codeword.

2) The zero vector is always a codeword.

3) The number of codewords in an [n,k]-code C of V[n,q] is q^k.

There are k vectors in a basis of C. Every codeword is expressible as a unique linear combination of basis vectors. Thus, to count the number of codewords, we just have to count the number of linear combinations. There are q choices for a scalar multiple of each basis vector and therefore q^k linear combinations in total.

Since the number of codewords of a linear code is determined by the dimension of the subspace, the (n, M, d) notation for general codes is generally replaced by **[n, k, d]** for linear codes.

Linear Codes

In general, finding the minimum distance of a code requires comparing every pair of distinct elements. For a linear code however this is not necessary.

Proposition 4: *In a linear code the minimum distance is equal to the minimal weight among all non-zero code words.*

Proof: Let x and y be code words in the code C, then $x - y$ is in C since C is linear. We then have $d(x,y) = d(x-y,0)$ which is the weight of $x-y$. ❑

(Notice that this proposition is actually valid in a larger class of codes ... one only requires that the alphabet permits algebraic manipulation and that the code is “closed” under subtraction.)

Generator Matrix

We shall now look at two ways of describing a linear code C.

The first is given by a ***generator matrix*** G which has as its rows a set of basis vectors of the linear subspace C. If C is an [n,k]-code, then G will be a k × n matrix.

The code C is the set of all linear combinations of the rows of G, or as we usually call it, the row space of G.

Given the matrix G, the code C is obtained by multiplying G on the left by all possible 1 × k row vectors (this gives all possible linear combinations):

$$C = \{xG \mid x \in V[k,q]\}.$$

Equivalence of Linear Codes

The general concept of equivalence of codes does not necessarily preserve the property of a code being linear. That is, linear codes may be equivalent to non-linear codes. In order to preserve the linear property we must limit the types of operations allowed when considering equivalence.

Two linear codes are ***equivalent*** if one can be obtained from the other by a series of operations of the following two types:

1) an arbitrary permutation of the coordinate positions, and
2) in any coordinate position, multiplication by any non-zero scalar.

(Note that this second operation preserves the 0 entries.)

Generator Matrix

Due to this definition of equivalence, elementary row and column operations on the generator matrix G of a linear code produce a matrix for an equivalent code.

Since G has rank k, by elementary row operations we can transform G to a matrix with a special form. Thus, we see that every linear code has an equivalent linear code with a generator matrix of the form $G = [I_k\ P]$, where I_k is the $k \times k$ identity matrix and P is a $k \times n\text{-}k$ matrix. We call this the ***standard form of G***.

Example

Let C be the [7,4]-code of V[7,2] generated by the rows of G (in standard form):

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

We get the 16 code words by multiplying G on the left by the 16 different binary row vectors of length 4.

So for instance we get code words:

$$(1,1,0,0)\ G = (1,1,0,0,1,0,1)$$
$$(1,0,1,1)\ G = (1,0,1,1,1,0,0)$$
$$(0,0,0,0)\ G = (0,0,0,0,0,0,0).$$

Example

The list of all the codewords is:

0 0 0 0 0 0 0
0 0 0 1 1 0 1
1 1 1 1 1 1 1
1 1 1 0 0 1 0
1 1 0 1 0 0 0
1 0 0 0 1 1 0
0 0 1 0 1 1 1
0 1 1 1 0 0 1
0 1 1 0 1 0 0
0 1 0 0 0 1 1
1 0 0 1 0 1 1
1 0 1 1 1 0 0
0 0 1 1 0 1 0
1 0 1 0 0 0 1
1 1 0 0 1 0 1
0 1 0 1 1 1 0

Notice that there are 7 codewords of weight 3, 7 of weight 4, 1 of weight 7 and 1 of weight 0. Since this is a linear code, the minimum distance of this code is 3 and so it is a 1-error correcting code.

This [7,4,3] code is called the **[7,4] –** ***Hamming Code***. It is one of a series of codes due to Hamming and Golay.

Parity Check Matrix

We now come to the second description of a linear code C.

The orthogonal complement of C, i.e. the set of all vectors which are orthogonal to every vector in C [orthogonal = standard dot product is 0], is a subspace and thus another linear code called the ***dual code*** of C, and denoted by $C^\perp$. If C is an [n,k]-code then $C^\perp$ is an [n, n-k] code.

A generator matrix for $C^\perp$ is called a ***parity check*** matrix for C. If C is an [n,k]-code then a parity check matrix for C will be an n-k × n matrix. If H is a parity check matrix for C, we can recover the vectors of C from H because they must be orthogonal to every row of H (basis vectors of $C^\perp$).

Parity Check Matrix

Thus the code C is given by a parity check matrix H as follows:

$$C = \{ x \in V[n,q] \mid Hx^T = \mathbf{0} \}$$

since the entries of this product are just the dot products of x with the rows of H.

Example

A parity check matrix for the [7,4]-Hamming code is given by:

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

Recall from the earlier example that 0001101 is a codeword and notice that

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

Parity Check Matrices

Theorem 1 : *Let H be a parity-check matrix for an* [n,k]-*code C in V[n,F]. Then every set of s-1 columns of H are linearly independent if and only if C has minimum distance at least s*.

Proof: First assume that every set of s-1 columns of H are linearly independent over F. Let c = (c_1 c_2 ... c_n) be a non-zero codeword and let $\mathbf{h_1}$, $\mathbf{h_2}$, ... , $\mathbf{h_n}$ be the columns of H. Then since H is the parity check matrix, Hc^T = **0**. This matrix-vector product may be written in the form

$$H\, c^T = \sum_{i=1}^{n} c_i h_i = 0\,.$$

The weight of c, wt(c) is the number of non-zero components of c.

Parity Check Matrices

If wt(c) ≤ s - 1, then we have a nontrivial linear combination of less than s columns of H which sums to **0**. This is not possible by the hypothesis that every set of s - 1 or fewer columns of H are linearly independent. Therefore, wt(c) ≥ s, and since c is an arbitrary non-zero codeword of the linear code C it follows that the minimum non-zero weight of a codeword is ≥ s. So, since C is linear (Prop. 4), the minimum distance of C is ≥ s.

To prove the converse, assume that C has minimum distance at least s. Suppose that some set of t < s columns of H are linearly dependent. Without loss of generality, we may assume that these columns are $\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_t$.

Parity Check Matrices

Then there exist scalars λ_i in F, not all zero, such that

$$\sum_{i=1}^{t} \lambda_i h_i = 0 .$$

Construct a vector c having λ_i in position i, $1 \leq i \leq t$, and 0's elsewhere. By construction, this c is a non-zero vector in C since $Hc^T = \mathbf{0}$. But wt(c) = t < s. This is a contradiction since by hypothesis, every non-zero codeword in C has weight at least s. We conclude that no s-1 columns of H are linearly dependent. ■

Parity Check Matrices

It follows from the theorem that a linear code C with parity-check matrix H has minimum distance (exactly) d if and only if *every* set of d-1 columns of H are linearly independent, and *some* set of d columns are linearly dependent. Hence this theorem could be used to determine the minimum distance of a linear code, given a parity-check matrix.

Parity Check Matrices

It is also possible to use this theorem to *construct* single-error correcting codes (i.e., those with a minimum distance of 3). To construct such a code, we need only construct a matrix H such that no 2 or fewer columns are linearly dependent. The only way a single column can be linearly dependent is if it is the zero column. Suppose two non-zero columns $\mathbf{h}_i$ and $\mathbf{h}_j$ are linearly dependent. Then there exist non-zero scalars a ,b $\in$ F such that

$$a\,\mathbf{h}_i + b\mathbf{h}_j = 0.$$

This implies that

$$\mathbf{h}_i = -a^{-1}b\,\mathbf{h}_j,$$

meaning that $\mathbf{h}_i$ and $\mathbf{h}_j$ are scalar multiples of each other. Thus, if we construct H so that H contains no zero columns and no two columns of H are scalar multiples, then H will be the parity-check matrix for a linear code having distance at least 3.

Example

Over the field F = GF(3) = $\mathbb{Z}_3$ (integers mod 3), consider the matrix

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

The columns of H are non-zero and no column is a scalar multiple of any other column.

Hence, H is the parity-check matrix for a [5,2]-code in V[5,3] with minimum distance at least 3.

Converting Representations

When working with linear codes it is often desirable to be able to convert from the generator matrix to the parity-check matrix and vice-versa. This is easily done.

Theorem 2: *If G* = $[I_k\ A]$ *is the generator matrix (in standard form) for the* [n,k]-*code C, then H* = $[-A^T\ I_{n-k}]$ *is the parity check matrix for C*.

Proof: We must show that H is a generator matrix for $C^\perp$. Now $GH^T = I_k\,(-A) + A\,I_{n-k} = 0$, implying that the rows of H are orthogonal to the rows of G, thus *span*(H) = {row space of H} is contained in $C^\perp$.

Converting Representations

Consider any $\mathbf{x} \in C^{\perp}$ where $\mathbf{x} = (x_1\ x_2\ \ldots x_n)$ and let

$$\boldsymbol{y} = \boldsymbol{x} - \sum_{i=1}^{n-k} x_{i+k} \boldsymbol{r_i}$$

where $\mathbf{r_i}$ is the ith row of H. Since $\mathbf{x} \in C^{\perp}$ and we have just proven that $\mathbf{r_i} \in C^{\perp}$, $1 \le i \le k$, it follows that $\mathbf{y} \in C^{\perp}$. We now examine the structure of **y**. By construction, components k + 1 through n are 0, so $\mathbf{y} = (y_1\ y_2\ \ldots\ y_k\ 0\ 0\ \ldots\ 0\ )$. But since $\mathbf{y} \in C^{\perp}$, $G\mathbf{y}^T = 0$, which implies that $y_i = 0$, $1 \le i \le k$. Therefore, $\mathbf{y} = \mathbf{0}$ and

$$\boldsymbol{x} = \sum_{i=1}^{n-k} x_{i+k} \boldsymbol{r_i}.$$

Hence, $\mathbf{x} \in$ *span*(H) and we have $C^{\perp} \subseteq$ *span*(H). Thus, *span*(H) = $C^{\perp}$ and so, H is a generator matrix of $C^{\perp}$. ■

Example (Cont.)

To look at the code we have previously constructed, it would be convenient to have the generator matrix. Since H is the generator matrix for $C^{\perp}$, if we apply the last theorem we can get the parity-check matrix for $C^{\perp}$ which is the generator matrix for C. To this end we perform row operations on H to put it into standard form H'.

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \rightarrow H' = \begin{pmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} = (I_3\, A),\, so\, A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

$$G = (-A^T\, I_2) = \begin{pmatrix} 2 & 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Example (Cont.)

We can now take all linear combinations (over GF(3)) of the rows to write out the 9 codewords of C. With their weights they are

Codeword	Weight
00000	0
20210	3
11001	3
10120	3
22002	3
01211	4
21121	5
12212	5
02122	4

$$G = \begin{pmatrix} 2 & 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

And we see that we have indeed generated a code of minimum distance 3.

Singleton Bound

As a corollary to Theorem 1 we can derive a relationship between the parameters of a linear code which is known as the ***Singleton bound***.

Corollary: *For any* [n,k,d]- *linear code we have* $n - k \geq d - 1$.

Proof: Let H be the parity check matrix for the code. By Theorem 1, any d-1 columns of H are linearly independent, so the column rank of H $\geq d - 1$. But since the column rank = row rank of H, and H has row rank = n - k, we obtain the desired inequality. ■

(While our statement and proof above only works for linear codes, a generalization of the result is actually true for all codes.)

Hamming Codes

Hamming Codes

A ***Hamming Code*** of order r over GF(q) is an [n,k]-code where n = $(q^r-1)/(q-1)$ and k = n - r, with parity check matrix H_r an r × n matrix such that the columns of H_r are non-zero and no two columns are scalar multiples of each other.

Note that $q^r - 1$ is the number of non-zero r-vectors over GF(q) and that q - 1 is the number of non-zero scalars, thus n is the maximum number of non-zero r-vectors no two of which are scalar multiples of each other. It follows immediately from Theorem 1 that the Hamming codes all have minimum distance exactly 3 and so are 1-error correcting codes.

Hamming Codes are Perfect

Since the number of codewords in a Hamming code is q^k, a direct computation shows that sphere packing bound is met, so:

Theorem 3 : *The Hamming codes of order r over GF(q) are perfect codes*.

Proof: With M = q^k, and d = 3 = 2e + 1, ie. e = 1 we have:

$$M\sum_{i=0}^{e}\binom{n}{i}(q-1)^i=q^k(1+n(q-1))=q^k(1+\frac{q^r-1}{q-1}(q-1))=q^{k+r}=q^n.$$

Example

The Hamming code of order r = 3 over GF(2) is given by the parity-check matrix

$$H_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

This is the [7,4]-code with distance 3. Re-ordering the columns of H_3 would define an equivalent Hamming code.

Example

The [13,10]-Hamming code of order 3 over GF(3) is given by the parity-check matrix

$$H_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 2 & 0 & 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 2 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 2 & 1 & 1 & 2 \end{pmatrix}.$$