Martin Cochran

Error Correcting Codes.

April 13, 2002

Definition: V is the space of all n-tuples of 0's and 1's with addition of vectors component wise mod-2.

Definition: An $[n, k]$ linear binary code is the set of all linear combinations of k independent vectors in V. Another way to think: a k-dimensional subspace of V.

Definition: An $[n, k]$ code over $GF(q)$ is a k-dimensional subspace of F^n, the space of all n-tuples with components from $GF(q)$, where $GF(q)$ stands for a field of order q.

Definition: A Generator Matrix is a matrix whos rows form a basis for an $[n, k]$ code.

Definition: Since a code can be described by parity checks, a parity check matrix can be formed, in which all rows are orthagonal to all elements from the code. That is, in order for a vector to be in the code, it must be orthagonal to every row in the parity check matrix.

Definition: We say that a generator matrix G of an $[n, k]$ code C is in standard form if $G = (I, A)$, where I is the $k \times k$ identity matrix and A is a $k \times (n - k)$ matrix.

Examples Here is a generator matrix for the $[7, 4]$ hamming code:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

In terms of parity equations, this code can be written as $a_5 = a_2 + a_3 + a_4$, $a_6 = a_1 + a_3 + a_4$, $a_7 = a_1 + a_2 + a_4$. And the parity check matrix can be written as

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Definition: An $[n, k]$ code C is specified by a parity check matrix H is every vector in C is orthongonal to the rows of a H of rank $n - k$ with n columns.

Definition: The weight of a vector u is the number of non-zero components it has and is denoted by $wt(u)$.

Definition: The minimum weight of a code C, denoted by d, is the weight of the non-zero vector of smallest weight in the code. An $[n, k]$ code with minimum weight d is often called an $[n, k, d]$ code.

Definition: The distance between two vectors u and v is the number of positions in which they differ. This is denoted by $d(u, v)$. It is easy to see that $d(u, v) = wt(u - v)$.

Theorem 1. *The distance function is a metric. Thus the following three properties hold:*

$$(i)\ d(u, u) = 0.$$

$$(ii)\ d(u, v) = d(v, u).$$

$$(iii)\ d(u, w) \leq d(u, v) + d(v, w).$$

(Triangle inequality).

Error Correcting

Definition: A sphere of radius r about a vector u, denoted by $S_r(u)$ as

$$S_r(u) = \{v \in V \mid d(u, v) \leq r\}$$

Theorem 2. *If d is the minimum weight of a code C, then C can correct $t = [(d-1)/2]$ or fewer errors and conversely. (Proof on page 11. All page numbers refer to Vera Pless: Error Correcting Codes 3rd ed).*

Proof: a) We prove that spheres of radius $t = [(d-1)/2]$ are disjoint.

b) Suppose that they are not.

c) Let u and w be distinct vectors in C, and assume that $S_r(u) \cap S_r(w)$ is non-empty.

d) Then $\exists$ a vector $v \in S_r(u) \cap S_r(w)$.

e) Then $d(u, w) \leq d(u, v) + d(v, w)$ by the Triangle Inequality.

f) but $d(u, v) + d(v, w) \leq 2t$ from sphere of radius t.

g) and $2t \leq d - 1$

h) but $d(u, w) = wt(u - w)$ which must have $wt \geq d$ because $u - w \in C$. i) Contradiction.

Definition: If C is a code, we let $C^Y = \{u \in V \mid v\dot{w} = 0 \ \forall w \in C\}$. It is known that if C is k-dimensional, then C^Y is $(n-k)$-dimensional. C^Y is called the dual or orthagonal code of C.

Syndrome Decoding

Definition: The standard array of a code is a table of vectors whose rows are the cosets of C arranged as follows. The first row is C itself with the zero vector in the first column. The first entry of any other row (i.e., any other coset) contains a coset leader, and the remainder of the row is constructed by adding this leader to the codewords in the first row to obtain the other vectors in the coset. Each element in the coset is placed in the column of the codeword it came from.

For example, given the binary code with generator matrix

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

and H the parity check matrix,

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

The standard array is as follows, with the top row being the code itself

$$\begin{bmatrix} (0,0,0,0) & (1,0,1,0) & (0,1,1,1) & (1,1,0,1) \\ (1,0,0,0) & (0,0,1,0) & (1,1,1,1) & (0,1,0,1) \\ (0,1,0,0) & (1,1,1,0) & (0,0,1,1) & (1,0,0,1) \\ (0,0,0,1) & (1,0,1,1) & (0,1,1,0) & (1,1,0,0) \end{bmatrix}$$

As an implementation, storing a standard array for a code would not be efficient. A code of length n would require a matrix with 2^n entries. This is where syndrome decoding comes in.

Definition: Let H be the parity check matrix of an $[n,k]$ code C with rows $h_1,\ldots,h_{n-k}$. If y is any vector in V, the syndrome of y is defined to be the column vector

$$syn(y) = \begin{bmatrix} yh_1 \\ . \\ . \\ . \\ yh_{n-k} \end{bmatrix}$$

of height $n-k$.

Theorem 3. *Every vector in a fixed coset has the same syndrome. Vectors in different cosets have different syndromes. All possible q^{n-k} syndromes occur as syndromes of some vectors.*

Proof: 1) a) Let $a+C$ be a coset of C.

b) Then two elements of $a+C$ can be written as $a+c_1$ and $b+c_2$ with $c_1,c_2 \in C$.

c) Then $(a+c_1)h_i = ah_i = (a+c_2)h_i$ for each row h_i in H.

2) a) Suppose that $a+c_1$ and $a+c_2$ are in distinct cosets but they have the same syndrome.

b) Then $(ah_i) = (bh_i)$ for all h_i in H c) $\Rightarrow a-b$ is orthogonal to all rows of H and thus $a-b$ is in the same coset.

d) This means a and b are in the same coset.

3) Since there are q^{n-k} distinct cosets, there are q^{n-k} distinct syndromes. There are all possible vectors with $n-k$ components from $GF(q)$.

Theorem 4. *If C is a binary code and e is any vector, the syndrome of e is the sum of those columns of H where e has nonzero components. The proof of this follows directly from the definition of a syndrome.*

Theorem 5. *Syndrome decoding is a maximum-likelihood decoding scheme.*

If $t = [(d-1)/2]$ where d is the minumum of weight of the code, then we can decode all vectors with coset leaders of weight t or less, and detect otherwise.

Algorithm: Decode a $v \in V$ by computing its syndrome $syn(v) = e$. Now subtract the coset leader with syndrome e from v. That's it.

Definition: A vector of smallest weight (there can be more than one of smallest weight) in a coset is called a coset leader. **Definition:** The weight of a coset is the weight of its coset leader. The code itself has weight 0.

Perfect Codes

In Theorem 2 we showed that spheres of radius $t = [(d-1)/2]$ about codewords in a code of minimum weight d are disjoint. It is possible that there are vectors in V that are not contained in any of these spheres. Often this is the case.

Definition: A code of minimum weight d in called perfect if all the vectors in V are contained in the spheres of radius $t = [(d-1)/2]$ about the codewords. In this case the spheres are said to cover the space.

The trivial perfect codes are the whole space or a binary repetition code of odd length. Some other perfect codes are the binary Hamming $[7,4,3]$ code, the binary Golay $[23,12,7]$ code, and the ternary Golay $[11,6,5]$ code.

Definition: For each non-trivial positive integer r there is a general binary Hamming code, denoted by $Ham(r,2)$, whose parity check matrix has as columns all nonzero binary r-tuples. Note: any ordering gives an equivalent code. The general Hamming code can easily be shown to be a $[2^r-1, 2^r-1-r, 3]$ code.

General Hamming Codes

Note, the columns of the parity check matrix of the Hamming code shown earlier were all the non-zero triples. We can construct an infinite family of single-error correcting perfect codes in this manner whose parity check matricies consist of all $n-k$ tuples. That is, for each r, there is a $Ham(r,2)$ code where the whose parity check...r-tuples.

Theorem 6. *The general binary Hamming* $[2^r-1, 2^r-1-r, 3]$ *codes are perfect single-error-correcting codes. (proof on pg 22).*

*Proof:*Each general Hamming code is a single-error correcting code because every vector of weight 1 is in a distinct coset. The codes are perfect because $(n(q-1)+1)(q^{n-r}) = q^n$ where $n(q-1)+1$ is the number of vectors in a sphere of radius 1 about a codeword and q^{n-r} is the number of spheres.

Theorem 7. *In order for a perfect t-error-correcting binary $[n,k]$ code to exist, the numbers n, k, and t must satisfy the following equation.*

$$\left(\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t}\right) 2^k = 2^n.$$

In order for a perfect t-error-correcting $[n,k]$ code over $GF(q)$ to exist, the numbers n, k, and t must satisfy the following equation.

$$\left(\binom{n}{0} + (q-1)\binom{n}{1} + \cdots + (q-1)^t\binom{n}{t}\right) q^k = q^n$$

Theorem 8 (Sphere Packing Bound)**.** *If C is an $[n,k,d]$ code over $\mathbb{Z}_2$, then*

$$\left(\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t}\right) 2^k \leq 2^n.$$

Hence, given n and k, this equation bounds t and so bounds d. (proof on page 23).

Theorem 9. *The only nontrivial multiple-error-correcting perfect codes are equivalent to either the binary* $[23, 12, 7]$ *code or the ternary* $[11, 6, 5]$ *code. The only non-trivial single-error-correcting perfect codes have the same parameters of the Hamming codes.*

Definition: The packing radius t is the largest among the numbers s where the spheres of radius s about codewords are distinct.

Theorem 10. *The packing radius t has the following properties.*
(i) If C has minimum weight d, $t = [(d-1)/2]$.
(ii) t is the largest among the numbers s so that each vector of weight $\leq s$ is a unique coset leader.

Proof: (i) Let $t' = [(d-1)/2]$.
We know by a previous theorem that spheres of radius t' are disjoint. We want to show that spheres of larger radius are not. It suffices to show that spheres of radius $t'+1$ are not disjoint.
Let u be a vector in C of weight d. Suppose first that d is even and let x be a vector with only $d/2$ non-zero components that agree with $d/2$ non-zero components of u. Then $x \in S_(t'+1)(u) \cap S_(t'+1)(0)$. If d is odd, let x have $(d+1)/2$ non-zero components in common with u then $d(u,x) = (d-1)/2 = t'$ and $d(u,0) = (d+1)/2 = t'+1$ so $x \in S_(t'+1)(u) \cap S_(t'+1)(0)$.
(ii) (work this proof out)
Definition: The covering radius r is the smallest number s such that spheres of radius s about codewords cover V.

Theorem 11. *The covering radius has the following properties.*
(i) r is the weight of the coset of largest weight.
(ii) r is the smallest among the numbers s such that every syndrome is a combination of s or fewer columns of any parity check matrix.

Proof: (i) Suppose that x is a coset leader of weight greater than r. Then $d(c,x) = wt(x-c) > r\ \forall c \in C$. Thus x cannot be in any sphere of radius r about a codeword. Now let a be the weight of the coset leader with greatest weight. Assume that there is a vector y whose distance from all codewords is greater than a. But y is in a coset with leader w and $y = w + c$ for some $c \in C$. Then $d(y,c) = d(w+c) = wt(w)$. But from the minimality argument of a, we have a contradiction.
(ii) Note that the syndrome corresponding to a coset leader of weight i is a combination of i columns of any parity check matrix. So (ii) naturally follows from (i).

Perfect Codes: If $t = r$ then the code is perfect.