Error-Correcting Codes: Solutions #3

Alfred Menezes (cryptography101.ca)

---

1. G_1 is a $k_1 \times n_1$ matrix of rank k_1, and G_2 is a $k_2 \times n_2$ matrix of rank k_2. Hence G is a $(k_1 + k_2) \times (n_1 + n_2)$ matrix. One can perform row operations on the first k_1 rows of G to convert G_1 to reduced echelon form E_1, and then perform row operations on the last k_2 rows of the resulting matrix to convert G_2 to reduced echelon form E_2. The resulting matrix

$$E = \left[\begin{array}{c|c} E_1 & 0 \\ \hline 0 & E_2 \end{array}\right]$$

and has $k_1 + k_2$ leading 1's, so its rank is $k_1 + k_2$. Since G is row equivalent to E, the rank of G is also $k_1 + k_2$. Hence C is a linear code of length $n = n_1 + n_2$ and dimension $k = k_1 + k_2$.

Since the codewords in C are the linear combinations of rows of G, C can be described as the set

$$\{(c_1, c_2) \ : \ c_1 \in C_1, \ c_2 \in C_2\}.$$

Hence

$$\begin{aligned} d(C) &= w(C) = \min_{(c_1,c_2)\neq 0} w((c_1, c_2)) \\ &= \min_{(c_1,c_2)\neq 0} \{w(c_1) + w(c_2)\} \\ &= \min\left(\min_{c_1\neq 0} w(c_1), \ \min_{c_2\neq 0} w(c_2)\right) \\ &= \min\left(w(C_1), w(C_2)\right) \\ &= \min(d_1, d_2). \end{aligned}$$

Thus the distance of C is $d = \min(d_1, d_2)$.

2. (a) A parity-check matrix H for C is an $(n-k) \times n$ matrix every $d-1$ columns of which are linearly independent over $GF(q)$. Hence the column rank of H is at least $d-1$. Since the column and row ranks of H are equal, H must have row rank at least $d-1$. Finally, since H has $n-k$ rows, we must have $d - 1 \leq n - k$, and so $d \leq n - k + 1$.

   (b) Let H be a parity-check matrix for C, and let $h_1, h_2, \ldots, h_n$ denote the columns of H. Without loss of generality, suppose that $S = \{1, 2, \ldots, d\}$. Now, since H has rank $n-k$ and $d = n-k+1$, the columns $h_1, h_2, \ldots, h_d$ must be linearly dependent over F. Hence there exists $a_1, a_2, \ldots, a_d \in F$, not all zero, such that $a_1h_1 + a_2h_2 + \cdots + a_dh_d = 0$, from which it follows that $c = (a_1, a_2, \ldots, a_d, 0, \ldots, 0)$ is a nonzero codeword. Now, since C has distance d, we must have $w(c) = d$, and so $a_1, a_2, \ldots, a_d$ are all nonzero. Hence, the nonzero coordinate positions of c are precisely the elements of S.

   (c) For each set S of d coordinate positions, let c be a codeword whose nonzero coordinate positions are precisely the elements of S. Then, for each nonzero λ in F, λc is also a codeword of weight d whose nonzero coordinate positions are precisely the elements of S. Altogether this gives $(q-1)\binom{n}{d}$ codewords of weight d.

   Finally, we need to show that there are no other weight-d codewords. Let c_1 and c_2 be any two weight-d codewords having the same set S of nonzero coordinate positions. Let λ_1 be the first nonzero component of c_1, and let λ_2 be the first nonzero component of c_2. Then $c = \lambda_1^{-1}c_1 - \lambda_2^{-1}c_2$ is a codeword of weight at most $d-1$. Since C has distance d, it must be the case that $c = 0$. Hence $c_1 = \lambda_1\lambda_2^{-1}c_2$, and so c_1 and c_2 are scalar multiples of each other. This shows that all weight-d codewords were accounted for in the previous paragraph.

3. (a) $n = 10$.
   H has rank 4 since columns 1, 7, 8 and 9 are linearly independent. Hence $n - k = 4$, and so $k = 6$.
   Since the columns of H are nonzero and distinct, $d(C) \geq 3$. However, the sum of columns 1 and 2 of H' equals column 7 of H. Hence $d(C) \not\geq 4$. It follows that $d(C) = 3$.

   (b) There are $2^{n-k} = 2^4 = 16$ cosets. Note that every vector of weight $\leq \lfloor \frac{d-1}{2} \rfloor = 1$ must be a coset leader. For the remaining 5 coset leaders, we choose arbitrary vectors of weight 2. Here is one 1-1 correspondence between syndromes and coset leaders.

Coset leader	Syndrome	Coset leader	Syndrome
0000000000	0000	0000000100	0010
1000000000	1000	0000000010	0111
0100000000	1001	0000000001	1101
0010000000	1110	1010000000	0110
0001000000	1111	1000100000	1011
0000100000	0011	1000000001	0101
0000010000	1010	0100000001	0100
0000001000	0001	0010000100	1100

   (c) i. The syndrome of r_1 is $s_1 = Hr_1^T = (0011)^T$. Hence $e = (0000100000)$ and r_1 is decoded to $c_1 = (1010001010)$.
   ii. The syndrome of r_2 is $s_2 = Hr_2^T = (0010)^T$. Hence $e = (0000000100)$ and r_2 is decoded to $c_2 = (0011001000)$.

4. (a) Since the zero vector is a codeword in C and has even weight, it is also in C', and so C' is non-empty. Let $x, y \in C'$. Then $x + y \in C$ since C is closed under addition. Also, $w(x+y) = w(x) + w(y) - 2\ell$ where ℓ is the number of coordinate positions in which x and y are both 1. Since $w(x)$ and $w(y)$ are even, $w(x+y)$ is also even, and hence $x + y \in C'$. Thus, C' is closed under addition.
   If $x \in C'$ then $0 \cdot x = 0 \in C'$ and $1 \cdot x = x \in C'$. Hence, C' is closed under scalar multiplication. So, C' is a vector subspace of C.

   (b) Let O' be the vectors of odd weight in C, and let $y \in O'$. Define $f : C' \to O'$ by $f(x) = x + y$. Note that $f(x)$ is indeed in O' since $x + y \in C$ and $w(x+y) = w(x) + w(y) - 2\ell$ where ℓ is the number of coordinate positions in which x and y are both even, whence $w(x+y)$ is odd. Now, f is injective since if $f(x_1) = f(x_2)$, then $x_1 + y = x_2 + y$ whence $x_1 = x_2$. Also, f is surjective since if $z \in O'$, then $z + y \in C$ and $w(z+y)$ is even (so $z + y \in C'$) and $f(z+y) = (z+y) + y = z$. Hence, f is a bijection, so $|C'| = |O'|$. Since $|C'| + |O'| = |C|$, it follows that $C'| = \frac{1}{2}|C|$.

   (c) $n' = n$, and $k' = k - 1$ since $|C'| = \frac{1}{2}|C| = \frac{1}{2}2^k = 2^{k-1}$.

   (d) If d is even, then the nonzero codewords of weight d in C are also in C'. Hence, $w(C') = d$, so $d' = d$.
   If d is odd, then the nonzero codewords of weight d in C are not in C'. Hence, the minimum

weight of a nonzero codeword in C' is at least $d+1$. Thus, $w(C') \geq d+1$, whence $d(C')$ is an even number that is $\geq d+1$.

5. (a) $s_2 = [B|I_{12}]r_1^T = (1101\ 1001\ 0110)^T$, which has weight > 3. Since s_2 differs in positions 3 and 5 from column 5 of B, the error vector is $e_1 = (0000\ 1000\ 0000\ 0010\ 1000\ 0000)$. r_1 is decoded to $c_1 = (0011\ 0000\ 0000\ 0110\ 0100\ 1110)$.

   (b) $s_2 = [B|I_{12}]r_2^T = (1001\ 0001\ 0000)^T$. Since $w(s_2) \leq 3$, the error vector is $e_2 = (0, s_2^T)$. r_2 is decoded to $c_2 = (0000\ 0000\ 0011\ 0110\ 1100\ 1001)$.

   (c) $s_1 = [I_{12}|B]r_3^T = (0010\ 1000\ 0000)^T$. Since $w(s_1) \leq 3$, the error vector is $e_3 = (s_1^T, 0)$. r_3 is decoded to $c_3 = (1100\ 0000\ 0000\ 1001\ 0001\ 1101)$.

   (d) $s_1 = [I_{12}|B]r_4^T = (0110\ 0001\ 0110)^T$, which has weight > 3. Since s_1 differs in positions 1 and 4 from column 5 of B, the error vector is $e_4 = (1001\ 0000\ 0000\ 0000\ 1000\ 0000)$. r_4 is decoded to $c_4 = (0110\ 0000\ 0000\ 0011\ 0010\ 0111)$.