Topics in Discrete Mathematics: Error-Correcting Codes: Solutions 1.

Dan Fretwell

Spring semester 2017/18

1. (a) This code is not linear since $2(122) = 211 \notin C$. Clearly $(n, M, d) = (3, 2, 3)$.

   (b) This code is not linear since $2002 + 0220 = 2222 \notin C$. Clearly $(n, M) = (4, 9)$ and we seek the value of d. Finding all 36 possible non-zero Hamming distances gives $d = 1$. Alternatively note that $d \geq 1$ and that $d(1221, 1201) = 1$.

   (c) This code is linear, a basis is given by $\{111, 100\}$. Clearly $(n, M) = (3, 9)$ and $k = 2$. To find d we simply observe from the list that the minimum non-zero weight is 1.

   (d) This code is clearly linear and the spanning vectors are linearly independent so are a basis. Thus $(n, k) = (7, 3)$ and so $M = 3^3 = 27$. To find d we (luckily) observe that $\mathrm{wt}(0000100) = 1$ and so clearly the minimum non-zero weight of C is 1, hence $d = 1$.

   (e) This code is clearly linear but the spanning vectors are not linearly independent since $21212 = -(12121)$. However one checks that $\{21212, 11022, 00200, 10000\}$ is a linearly independent set of codewords, hence $(n, k) = (5, 4)$ and $M = 3^4 = 81$. To find d we again observe that $\mathrm{wt}(10000) = 1$ so $d = 1$.

2. (a) The generator matrix clearly has rank 2 and so the code it generates has parameters $(n, k) = (3, 2)$, so that $M = 7^2 = 49$. One might be tempted to guess that $d = 2$ since $\mathrm{wt}(042) = 2$. However $d = 1$ since $121 + 3(042) = 100 \in C$ has weight 1. This exercise shows that you have to be careful when computing the minimum non-zero weight, it is **not** enough to look only at basis codewords.

   (b) The parity check matrix clearly has rank 3 and so the code it generates has parameters $(n, k) = (7, 4)$, so that $M = 7^4 = 2401$. To find d one looks for the minimum number of linearly dependent columns. Letting h_i be the ith column of the matrix, one observes that $h_5 = 2h_3$ and so $d \leq 2$. But a single non-zero vector cannot be linearly dependent so $d = 2$.

3. (a) Recall that the ISBN code is the linear code over $\mathbb{F}_{11}$ with parity check matrix:

$$(\ 10 \quad 9 \quad 8 \quad 7 \quad 6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1 \ ).$$

This has rank 1 and so the ISBN code has parameters $(n, k) = (10, 9)$, so that $M = 11^9$. To find d note that every entry of the matrix is non-zero and $\mathbb{F}_{11}$ is a field, so that every pair of columns is linearly dependent, thus $d = 2$.

(b) Multiplying each by the parity check matrix we find that these are ISBN numbers if and only if $6a \equiv 8 \bmod 11$, $4b \equiv 0 \bmod 11$, $7c \equiv 10 \bmod 11$ and $d \equiv 10 \bmod 11$. Solving gives $(a, b, c, d) = (5, 0, 3, \mathrm{X})$.

(c) Multiplying by the parity check matrix gives $(11 - i) + a_i$. Thus letting $a_i = i$ gives a valid ISBN codeword. Since the ISBN code has dimension 9 by part (a) and $\{c_1, c_2, ..., c_9\} \subset$ ISBN are linearly independent, they must be a basis for the ISBN code. Hence a generator matrix is given by:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 9 \end{pmatrix}$$

.

(d) The first claim is that we can detect one such error. Let $\mathbf{c} = c_1...c_i...c_j...c_{10}$ be an ISBN codeword with $c_i \neq c_j$ for some $i \neq j$. We want to show that the word $\mathbf{v} = c_1...c_j...c_i...c_{10}$ is not a valid ISBN codeword. To do this, suppose $\mathbf{v}$ is a valid ISBN codeword. Then since ISBN is linear we have that $\mathbf{c}-\mathbf{v} = 0...0(c_i-c_j)0...0(c_j-c_i)0...0$ is also an ISBN codeword. Multiplying by the parity check matrix we see that:

$$(11 - i)(c_i - c_j) + (11 - j)(c_j - c_i) \equiv 0 \bmod 11.$$

Equivalently we have $(j - i)(c_i - c_j) \equiv 0 \bmod 11$. However $c_i - c_j$ is invertible mod 11 (since $c_i \neq c_j$) and so $i \equiv j \bmod 11$. This is a contradiction sincd $1 \leq i, j \leq 10$ and $i \neq j$.

The second caim is that we cannot correct a single such error. To see this we give an example. consider the word $\mathbf{v} = 0000000011$. This is not an ISBN codeword. However the two ISBN coodewords $\mathbf{c}_1 = 0100000010$ and $\mathbf{c}_2 = 1000000001$ can both be transformed into $\mathbf{v}$ by swapping a single pair of digits. Thus the single swapping error cannot be corrected.

4. This code is a simple generalisation of the repetition code. The length is $n_m = mn$ and the number of codewords is $M_m = |C| = p^k$ (since C is linear). It is clear that each code C_m is linear and so each has dimension $k_m = k$. Also, since C_m is linear we know that the minimum non-zero distance d_m is the minimum non-zero weight. This is clearly md since the minimum non-zero weight of C is d. Thus $d_m = md$.

It follows that:

$$t_m = \left\lfloor \frac{d_m - 1}{2} \right\rfloor = \left\lfloor \frac{md - 1}{2} \right\rfloor \geq \left\lfloor \frac{m(d-1)}{2} \right\rfloor \geq m \left\lfloor \frac{d-1}{2} \right\rfloor = mt.$$

When $m \geq 3$ the first inequality is strict since:

$$\frac{md-1}{2} - \frac{m(d-1)}{2} = \frac{m-1}{2} \geq 1.$$

Since $t_m > mt$ for $m \geq 3$ and $t \geq 0$ we have that $t_m \geq 1$, so that C_m corrects at least one error.

5. (a) It is clear that $n = 5$. It is also clear that C is linear, with basis $\{01011, 10101\}$, so that $k = 2$. We have that $d = 3$ since the minimum non-zero weight is visibly 3.

   (b) The rank of A is visibly 3 (look at columns $1, 2$ and 5).Thus by the rank nullity theorem it follows that $\text{nullity}(A) = 5 - \text{rank}(A) = 2$.

   (c) A simple calculation shows that $A\mathbf{c}^T = \mathbf{0}$ for each $\mathbf{c} \in C$. Thus $C \subseteq \text{NullSpace}(A)$.

   However $C = \text{NullSpace}(A)$ since both have the same dimension by (a) and (b). Hence A is a parity check matrix for C.

   (d) We calculate $\text{Null}(A)$ by explicitly solving the equations given by $A\mathbf{c}^T = \mathbf{0}$, i.e.

$$\begin{aligned} x_1 + x_3 &= 0 \\ x_2 + x_4 &= 0 \\ x_3 + x_4 + x_5 &= 0. \end{aligned}$$

   Letting $x_3 = a \in \mathbb{F}_2$ and $x_4 = b \in \mathbb{F}_2$ we see that the solution space is given by $\{(a, b, a, b, a + b) \,|\, a, b \in \mathbb{F}_2\} = \{a(1, 0, 1, 0, 1) + b(0, 1, 0, 1, 1) \,|\, a, b \in \mathbb{F}_2\} = C$.

   (e) The first method was massively quicker since we only had to compute four matrix products and then compare dimensions. Method two involved solving a system of linear equations, which becomes more cumbersome as the size of A increases.

6. Suppose such a code exists. We will see that it contradicts the sphere packing bound. The LHS of the SPB is:

$$3^8 \left(\binom{13}{0} + 2\binom{13}{1} + 4\binom{13}{2} \right) = 3^8(339) = 3^9(113) > 3^{13}.$$

   This gives the required contradiction so no such code exists.

7. We know that C is perfect so there is equality in the sphere packing bound. Since $t = 1$ this gives:

$$p^k \left(\binom{n}{0} + (p-1)\binom{n}{1} \right) = p^n,$$

i.e.

$$1 + n(p-1) = p^{n-k}.$$

Rearranging gives $n = \frac{p^{n-k}-1}{p-1}$.

If such a code is binary then $p = 2$ so that $n = 2^{n-k} - 1$. Letting $r = n - k$ we see that $n = r + k = 2^r - 1$ so that $(n, k) = (2^r - 1, 2^r - r - 1)$.

8. Recall that the binary repetition code $\mathcal{R}_n$ is a $[n, 1, n]$-linear code over $\mathbb{F}_2$. Thus $t = \left\lfloor \frac{n-1}{2} \right\rfloor$.

   Let $n = 2r + 1$ be odd, then $t = r$ and the LHS of the sphere packing bound is:

$$\begin{aligned} 2\sum_{m=0}^{r} \binom{n}{m} = \sum_{m=0}^{r} \binom{n}{m} + \sum_{m=0}^{r} \binom{n}{n-m} &= \sum_{m=0}^{r} \binom{n}{m} + \sum_{m=r+1}^{2r+1} \binom{n}{m} \\ &= \sum_{m=0}^{n} \binom{n}{m} \\ &= (1+1)^n \\ &= 2^n, \end{aligned}$$

   thus $\mathcal{R}_n$ is perfect in this case.

   Let $n = 2r$ be even, then $t = r - 1$ and the LHS of the sphere packing bound is:

$$\begin{aligned} 2\sum_{m=0}^{r-1} \binom{n}{m} = \sum_{m=0}^{r-1} \binom{n}{m} + \sum_{m=0}^{r-1} \binom{n}{n-m} &= \sum_{m=0}^{r-1} \binom{n}{m} + \sum_{m=r+1}^{2r} \binom{n}{m} \\ &= \left(\sum_{m=0}^{n} \binom{n}{m} \right) - \binom{n}{r} \\ &= (1+1)^n - \binom{n}{r} \\ &< 2^n, \end{aligned}$$

   thus $\mathcal{R}_n$ is not perfect in this case.