Hamming codes

Information Theory (APMA 1710), Fall 2011

In this assignment, you will implement and test the (n, k) Hamming codes. (Suggestion: For this assignment, Matlab is probably the easiest language to use.)

(I) Implement (n, k) Hamming

Given any desired number of parity check bits $m \geq 3$, there is a (n, k) Hamming code with codeword length $n = 2^m - 1$ and block length $k = n - m$. (Suggestion: you may want to start by implementing the $(7, 4)$ code, and then generalize to (n, k) once you have everything working.) In what follows, m, K, and α are input arguments.

(1) Write code that takes m and constructs:
   (a) the parity check matrix $H = [F \quad I_m]$, and
   (b) the generator matrix $G = [I_k \quad F^T]$.
   You may use any permutation of the columns of the F matrix that you find convenient (i.e. in the case of the $(7, 4)$ code, you don't have to use the particular one we discussed in class.) (Suggestion: In Matlab, an easy way to produce the binary vector of length m corresponding to the number j is by using `(dec2bin(j,m)=='1')`. For example, `(dec2bin(5,6)=='1')` returns `[0,0,0,1,0,1]`.)

(2) Randomly sample a binary sequence $s = s_1 s_2 \cdots s_K$, where each s_i is drawn from a Bernoulli(1/2) distribution. This s will be our source message. (For convenience, we will assume that K is a multiple of the block length k.)

(3) Encode the source sequence s using the generator matrix G, producing a "transmitted" sequence $t = t_1 t_2 \cdots t_N$ where $N = nK/k$. (Suggestion: By arranging s in a $k \times (K/k)$ matrix, the encoding can be done with a single matrix multiplication. In fact, I would encourage you to represent all of the sequences s, t, u, r, and $\hat{s}$ (defined below) in matrix form.)

(4) Randomly sample a binary sequence $u = u_1 u_2 \cdots u_N$, where each u_i is drawn from a Bernoulli(α) distribution. This u will represent the noise in the channel.

(5) Compute the binary sequence $r = r_1 r_2 \cdots r_N$ such that $r_i \equiv t_i + u_i$ (where $\equiv$ denotes congruence mod 2). This r represents the "received" sequence.

(6) Using the H matrix, perform error correction on the received sequence r, producing the decoded sequence $\hat{s} = \hat{s}_1 \hat{s}_2 \cdots \hat{s}_K$.

(II) Verify your implementation

Empirically demonstrate that your implementation is correct **in the case of** $(n, k) = (7, 4)$, by printing the following:

(1) the parity check matrix H and the generator matrix G

(2) the s, t, u, r, and $\hat{s}$ resulting from a run of your code from part (I), using $K = 32$ and $\alpha = 0.1$. (It is visually helpful here to display these in the matrix form described in (I)(3) above.)

(III) Evaluate (n, k) Hamming

(1) Write code to estimate the probability of bit error p_b, using the estimate

$$\hat{p}_b = \frac{1}{K}\sum_{i=1}^{K} I(\hat{s}_i \neq s_i).$$

In other words, $\hat{p}_b$ is the fraction of bits in which $\hat{s}$ and s disagree.

(2) Set $\alpha = 0.01$ and $K = 326040$. For each $m = 3, \ldots, 8$, run your code and print the following quantities:

$$m \quad n \quad k \quad R \quad \hat{p}_b$$

where $R = k/n$ is the rate.

(3) Repeat (2) using $\alpha = 0.001$.

(4) Describe the trends you see in R and $\hat{p}_b$.

(5) Consider the code resulting from choosing $m = 8$. Compare the values of α and $\hat{p}_b$ in the case of $\alpha = 0.01$. Does this look like a code you would want to use for this value of α? Is it any better when $\alpha = 0.001$? Try some other values of α and describe the trend you observe.

(Note: I chose the K above to be divisible by all the k's corresponding to $m = 3, \ldots, 8$.)

(Extra Credit) Analytically compute the probability of bit error p_b

(1) Recall that the definition of p_b is:

$$p_b = \frac{1}{K}\sum_{i=1}^{K} \mathbb{P}(\hat{s}_i \neq s_i).$$

Prove that for the $(7, 4)$ Hamming code, $p_b \approx 9\alpha^2$ when α is small. (More precisely, $p_b = 9\alpha^2 + g(\alpha)$, where g is some function such that $g(\alpha)/\alpha^2 \to 0$ as $\alpha \to 0$.)

(2) Can you derive a similar approximation for the general case of a (n, k) Hamming code?