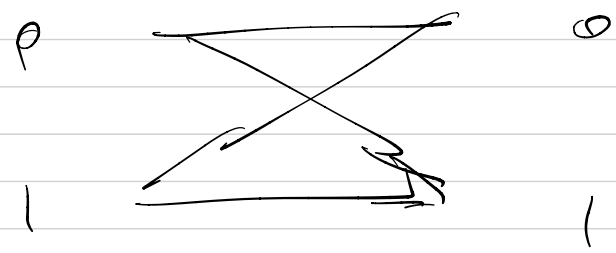
Information Theory & Coding

Nov 17th 2020

Yesterday: Coding for the Binary Symmetric Channel

$BSC(p)$ $\quad 0 \le p \le \frac{1}{2}$



$Y = X \oplus Z$ (mod 2)

$X \in \{0,1\}$ $\quad Z = \begin{cases} 0 & \text{w.p. } 1-p \\ 1 & \text{w.p. } p \end{cases}$

Z indep of X

$enc: \{1, \dots, M\} \to \{0,1\}^n$

code of rate $\frac{1}{n} \log_2 M$.

$d_{min} \triangleq \min_{\substack{m, m' \\ m \neq m'}} d_H(enc(m), enc(m'))$.

recall: $d_H(\bar{x}, \bar{x}') = \sum_{i=1}^{n} \mathbb{1}\{x_i \neq x_i'\}$.

$\bar{x}', \bar{x} \in \{0,1\}^n$

Thm: (Singleton Bound): given an encoder as above

$enc: \{1, \dots, M\} \to \{0,1\}^n$ with $M > 2^k$.
(integer)

Then $d_{min} \le (n-k)^+$.

Pf: For $k \ge n$: nothing to prove: there will be $m \neq m'$ s.t. $enc(m) = enc(m') \Rightarrow d_{min} = 0$.

For $k < n$, define the following hash function:

$\text{hash}: \{0,1\}^n \to \{0,1\}^k$

$\text{hash}(x_1 \dots x_n) = (x_1 \dots x_k)$

Consider $m \to \text{hash}(\text{enc}(m))$.

m: can take $M > 2^k$ values

$\text{hash}(\text{enc}(m))$: can take 2^k values

$\Rightarrow$ exists $m \neq m'$ s.t $\text{hash}(\underline{\text{enc}(m)}) = \text{hash}(\underline{\text{enc}(m')})$.

$d_{min} \leq d_H(\text{enc}(m), \text{enc}(m')) \leq n-k$ //.

Thm (Sphere Packing Bound): given $\text{enc}: \{1, \dots, M\} \to \{0,1\}^n$

then, $M, n, d = d_{min}$ satisfies: $r = \left\lfloor \frac{d-1}{2} \right\rfloor \leq \frac{d-1}{2}$

$$M \sum_{i=0}^{r} \binom{n}{i} \leq 2^n.$$

$\binom{n}{i}$: binomial coefficient $\frac{n!}{i!(n-i)!}$

Pf: for $x \in \{0,1\}^n$ and $r \geq 0$, let

$$B(x, r) = \{y \in \{0,1\}^n : d_H(x,y) \leq r\},$$

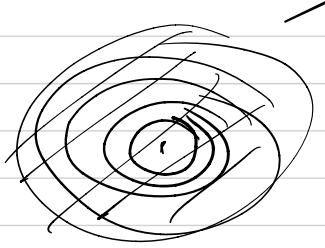
be the "Hamming ball" with center x, radius r. Observe:

$\{y : d_H(x,y) = i\}$ = all $y = (y_1 \dots y_n)$ s.t

$y_j = x_j$ for $(n-i)$ values of j and disagree

$y_j \neq x_j$ for i values of j.

$$B(x,r) = \bigcup_{i=0}^{r} \{y : d_H(x,y) = i\}$$



a $y=(y_1 \dots y_n)$ with $d_H(x,y)=i$ is specified exactly by the set $\{j : x_j \neq y_j\} = S$

$x = 01100$ $\quad\Rightarrow\quad S = \{2,3,4\}$

$y = 00010$

$$|\{y : d_H(x,y)=i\}| = |\{S : S \subset \{1..n\},\ |S|=i\}|.$$

$$\binom{n}{i}$$

$$B(x,r) = \bigcup_{i=0}^{r} \{\quad\}$$

The # of elements in $B(x,r) = \sum_{i=0}^{r} \binom{n}{i}$.

Consider now the collection

$$B(\mathrm{enc}(m), r) \qquad m = 1, \dots, M.$$

each $B(\mathrm{enc}(m), r)$ is a subset of $\{0,1\}^n$ of cardinality $*$. Furthermore for $m \neq m'$

$$B(\mathrm{enc}(m), r) \cap B(\mathrm{enc}(m'), r) = \phi.$$

[because if $y \in B(\mathrm{enc}(m), r)$ & $y \in B(\mathrm{enc}(m'), r)$

$$\Rightarrow d \overset{=}{\leq} d_H(\mathrm{enc}(m), \mathrm{enc}(m'))$$

$$\leq d_H(\mathrm{enc}(m), y) + d_H(y, \mathrm{enc}(m'))$$

$$\leq r + r = 2r \leq d-1 \quad \times]$$

So. $\bigcup_{m=1}^{M} B(\mathrm{enc}(m), r)$ has cardinality

$$M \sum_{i=0}^{r} \binom{n}{i}.$$ But this union $\subseteq \{0,1\}^n$

↑ cardinality 2^n

$$\Rightarrow M \sum_{i=0}^{r} \binom{n}{i} \leq 2^n \quad //.$$

Example: $n = 7 \qquad 2^n = 128$

$M = 16$ codewords. How large can $d = d_{min}$ be?

$\Leftrightarrow$ how large can r be.

$$\binom{n}{0} = 1, \quad \binom{n}{1} = 7, \quad \binom{n}{2} = \frac{7 \cdot 6}{1 \cdot 2} = 21, \ldots$$

$$M\left[\binom{n}{0}\right] \leq 2^n \quad \checkmark$$

$$M\left[\binom{n}{0} + \binom{n}{1}\right] \leq 2^n \quad \checkmark$$

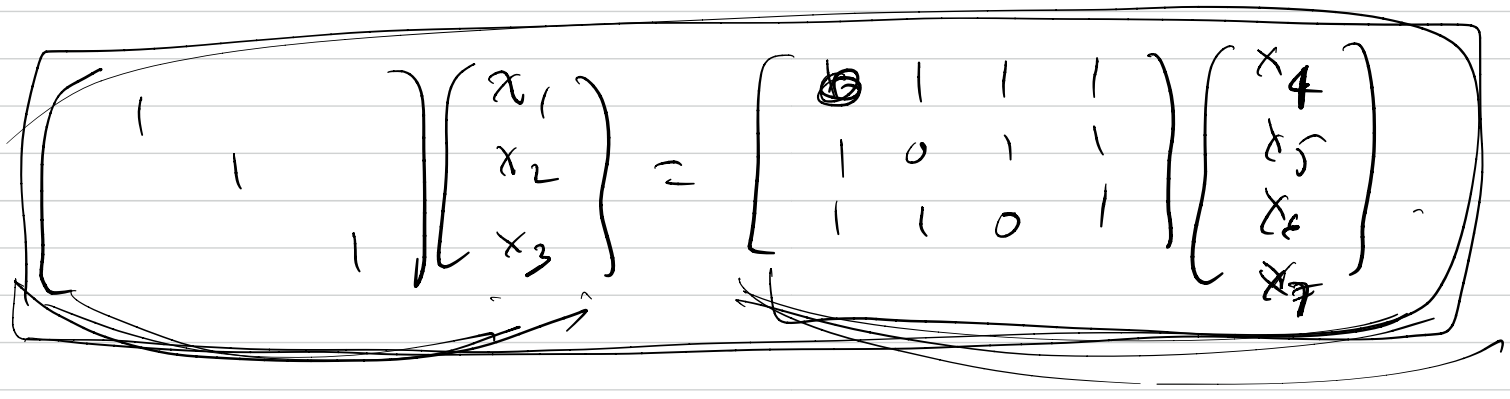
↑ ↑ ↑ =

16 ≤ 8

$$M\left[\binom{n}{0} + \binom{n}{1} + \binom{n}{2}\right) \leq 2^n \quad \times$$

$\Rightarrow r \leq 1$ ...

Consider now a code defined in the following way:

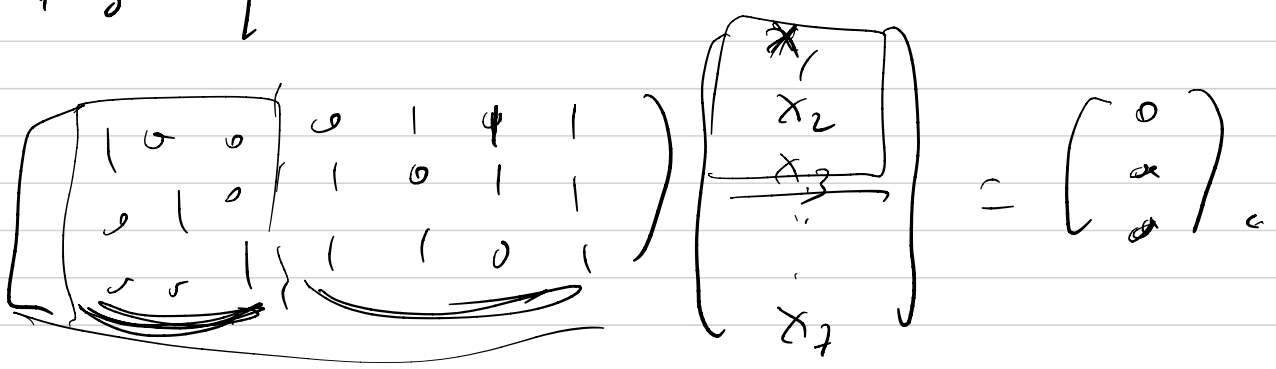


$$\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} \quad \text{mod } 2.$$

$$\text{enc}(x_4, x_5, x_6, x_7) \rightarrow (x_1\, x_2\, x_3\, x_4\, x_5\, x_6\, x_7).$$

16 possible combinations.

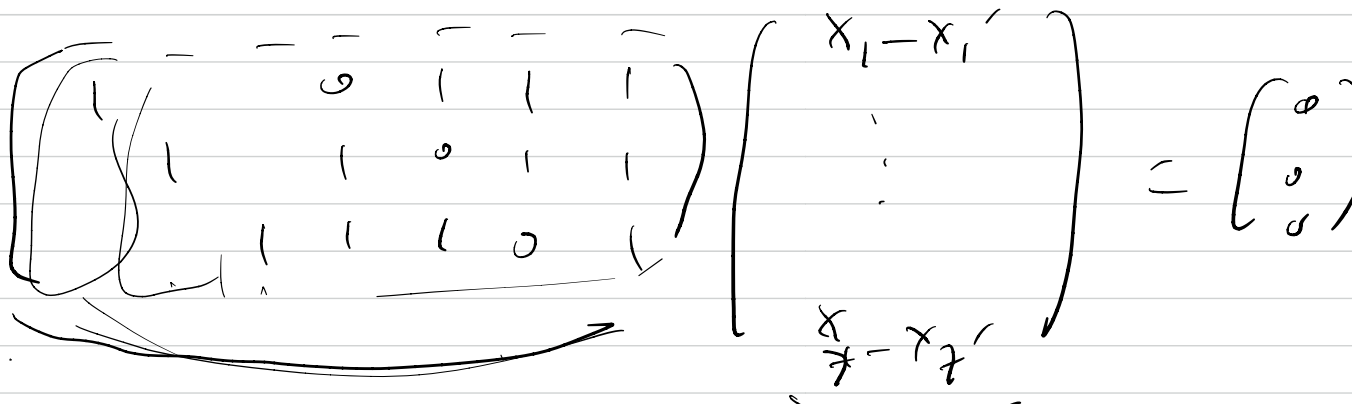
$M = 16$, $n = 7$. Q: what is the minimum distance of this code? The codewords satisfy the following equation:



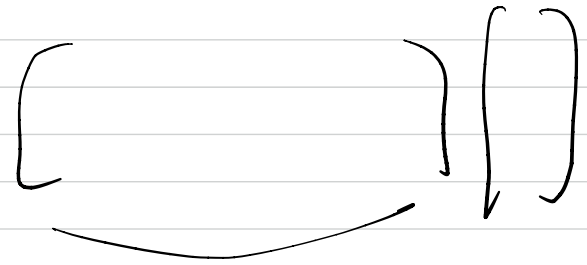
$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_7 \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

So if $(x_1 \dots x_7)$ is a codeword

$(x_1' \dots x_7')$ " " , then



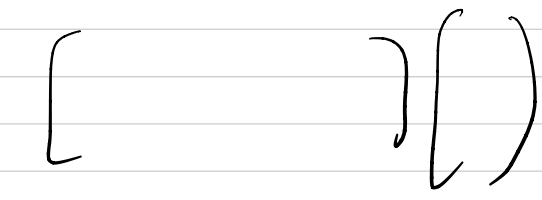
if x, x' differed in 1 position only ($d_H = 1$)

$\Rightarrow$  would be a column of the

matrix $\uparrow$. Since the matrix has no $\begin{pmatrix}0\\0\\0\end{pmatrix}$ column, the

equation cannot hold. $\Rightarrow$ $d_{min} > 1$.

if x, x' differed in exactly two positions $d_H = 2$.

then  would be the sum of two columns

of the matrix.

Since no two columns are the same, the sum of any

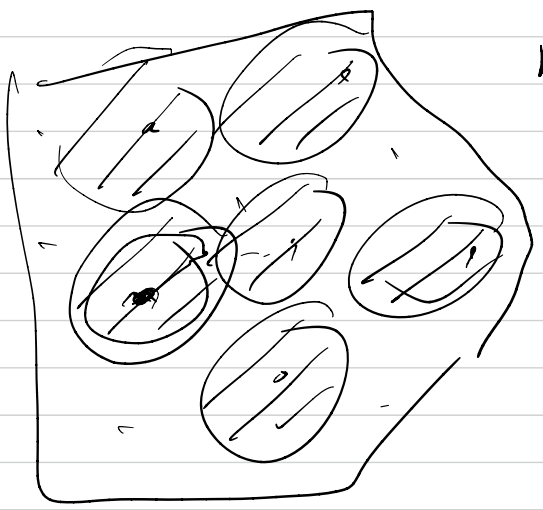
two columns $\neq \begin{pmatrix}0\\0\\0\end{pmatrix}$. $\Rightarrow$ the equation cannot hold

$\Rightarrow$ $d_{min} > 2 \Rightarrow$ $d_{min} \geq 3$.

So for this particular code with $M = 16$, $n = 7$ the

Balls of radius $r = 1$ are all disjoint and completely

fill the space $\{0,1\}^{n=7}$.



no empty space outside the shaded area.

For this code the sphere packing Bound is satisfied with equality.

In particular, we see that there is a code with

$n=7$, $d=3$, $M=16$, but no code with $n=7$, $d=3$, $M>16$.

So far we have two "negative" results (i.e. which upper bound M, d_{min}). On the "positive side" is:

Thm (Gilbert-Varshamov bound): given n, d. there is a code with M codewords,

$$M \sum_{i=0}^{d-1} \binom{n}{i} \geq 2^n, \quad \text{and } d_{min} \geq d.$$

Pf: Consider the following construction:

(1). $M=0$, and let $A = \{0,1\}^n$

(2) while $A \neq \emptyset$

$M \leftarrow M+1$

$enc(M) =$ any element of A

$A \leftarrow A - B(enc(M), d-1)$.

done.

By construction each new codeword is at distance
$\geq (d-1)$ from every other codeword picked before it.

$\Rightarrow$ we will construct a code with $d_{min} \geq d$.

Question what about M?

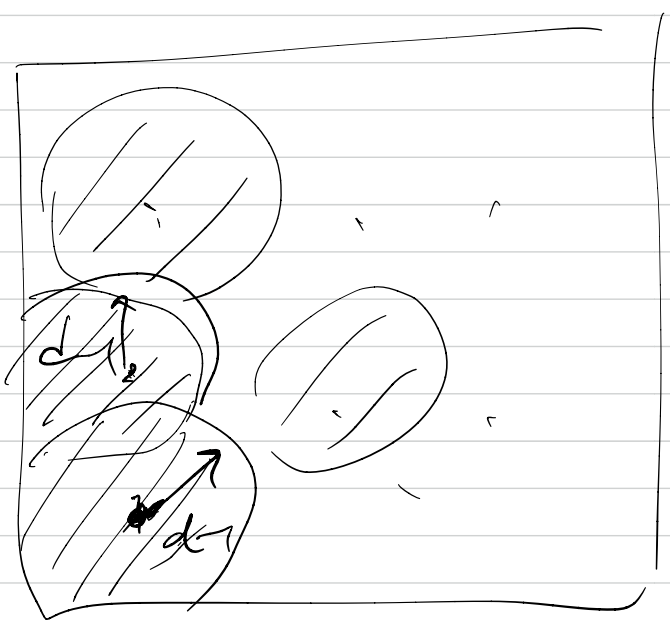
we start with $|A| = 2^n$

we end with $|A| = 0$.

at each step we eliminate $\leq |B(\cdot, d-1)|$ sequences from A, and increment the number of codewords by 1.

$\Rightarrow$ the number of steps the algorithm runs is

$$\geq 2^n / |B(\cdot, d-1)| = \frac{2^n}{\sum_{i=0}^{d-1} \binom{n}{i}}. \quad //$$



$\hookrightarrow 2^n$ binary sequences of length n, i.e. $\{0,1\}^n$.

Compared to the sphere packing bound, the distances are different by a factor 2.

Example: $M = 16$, $n = 7$.

GV will say $\exists$ code with $d_{min} \geq 2$.

SPB " " $\forall$ code $d_{min} \leq 4$.

Some reflections on coding:

Suppose we want a binary code with rate $= 1/2$.

$n = 200$. The number of codewords is

$$M = 2^{nR} = 2^{100} \approx 10^{30}.$$

$\text{enc}(m) \Longrightarrow$ 200 bit sequence

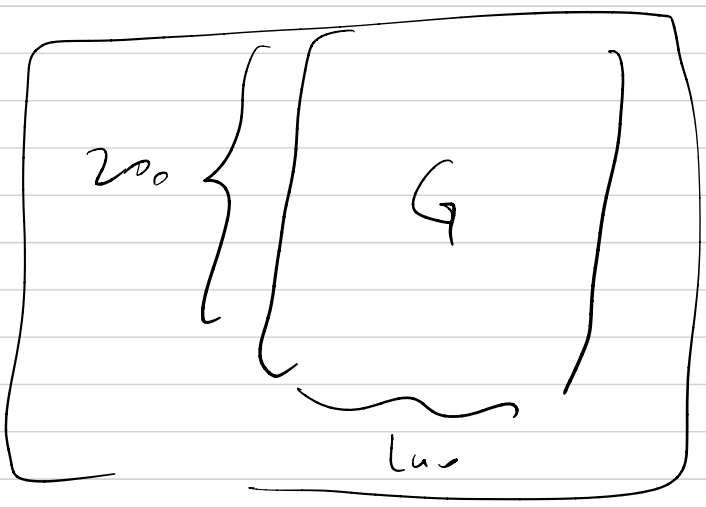
$m \in \{0,1\}^{100}$

- potentially, we may have to list 2^{100} sequences, each of length $= 200$ to describe enc.

Decoding is even worse: given $y \in \{0,1\}^{200}$, the decoder has to check each y the 2^{100} codewords to find the closest codeword to y. So unless the encoding function has some structure, there is no practical way to implement enc() or dec().

Let us try "linear" structures:

$\text{enc}: 100 \text{ bits} \longrightarrow 200 \text{ bits}$.



$$\text{enc}(\underbrace{u_1 \dots u_{100}}_{\in \{0,1\}^{100}}) = G \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_{100} \end{pmatrix}$$

The function "enc" is specified completely by the matrix G. (100×200 entries)

↑ binary entries

$$\begin{pmatrix} x_1 \\ \\ x_{200} \end{pmatrix} = \left(G \right) \begin{pmatrix} u_1 \\ \\ u_{100} \end{pmatrix}$$

$$x_i = \sum_j G_{ij} u_j$$

binary mult / addition (mod 2).

Clear: ① enc is simply described ($n(\log M)$ bits)

② enc operation is simple. ($\leq$ operations)

Under: ① decoder?

② is there a good code in this restricted class of enc's.

(perhaps all such "linear" codes have large $P(\text{Error})$?

or perhaps such " " cannot make R close to C?)

(Thm: Given $BSC(p)$, $R < C(BSC(p))$, $\varepsilon > 0$,
there is a matrix G s.t. the encoder dictated by G has rate $\geq R$ & has $P(\text{Error}) < \varepsilon$.

Pf outline: pick G randomly, (Idea: Peter Elias 1950s).
analyze its expected performance.)

Linear Codes:

recall what a vector space is:

"V is a vector space if $\forall x, y \in V$, and

(scalars a, b $\quad ax + by \in V$.)"

thus the definition needs the notion of scalars,

need " " " scalar × vector

" " " adding vectors.

We are used to vector being in $\mathbb{R}^n$ and scalars being $\mathbb{R}$.

We will need the binary equivalents, namely when vectors are in $\{0,1\}^n$ and scalars in $\{0,1\}$.