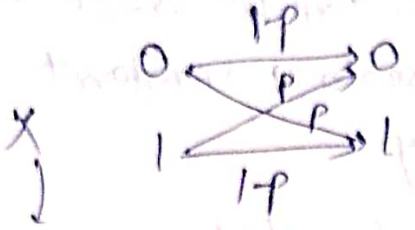
as n is increasing, such that $P_e \to 0$.

Channel Capacity.

Discrete Memoryless Channel

Examples → Binary symmetric channel.



Input r.v to the channel.

Conditional probabilities or transition probabilities.

Input to the channel → r.v → x

output to the channel → r.v → y

$P_{Y/X}(y/x)$

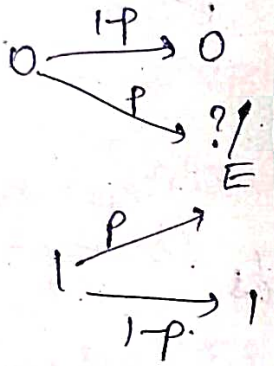
$P_{Y/X}(0/0) = 1-p = P_{Y/X}(1/1)$

$P_{Y/X}(0/1) = p = P_{Y/X}(1/0)$

$P_{X,Y}$ → joint probability is not known.

if we want to find, $P_{X,Y}$, we need to know the value of $P_X(x)$.

Binary Erasure channel



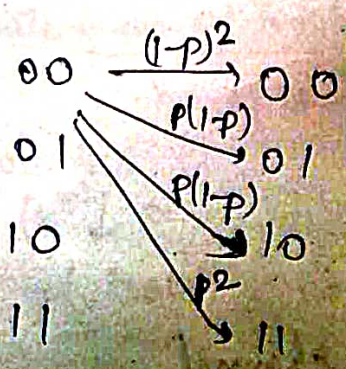
output has been erased (output) y → (we don't know what the output is)

In this channel, there is no possibility of flipping

E → something is transmitted, but it is not received.

signal

4-ary symmetric channel



$P_{Y/X}(00/00) = (1-p)^2$

Discrete memoryless Channel (DMC)

↳ Input and output are both discrete random variables.

Memoryless channel

↳ $P_{Y^n/X^n}((y_1, y_2, \cdots y_n)/(x_1, x_2, \cdots x_n))$

current output depends only on the current input.

$= P_{Y_1/X_1 \cdots X_n}(y_1/x_1, x_2, \cdots x_n) \times$

for any general channel ← $P_{Y_2/Y_1 X_1 \cdots X_n}(y_2/y_1, x_1, \cdots x_n) \times \cdots$

$(\because P_{Y_1,Y_2}(y_1, y_2) = P_{Y_1}(y_1) \times P_{Y_2/Y_1}(y_2/y_1))$

↙ unconditional.

for conditional → we have to condition it for every term.

If channel is memoryless,

$P_{Y_1/X_1 \cdots X_n}(y_1/x_1, x_2 \cdots x_n) = P_{Y_1/X_1}(y_1, x_1)$

$P_{Y_2/Y_1 X_1 X_2 \cdots X_n}(y_2/y_1 x_1 x_2 \cdots x_n) = P_{Y_2/X_2}(y_2, x_2)$

$y^n = (y_1, y_2, \cdots y_n)$

$x^n = (x_1, x_2, \cdots x_n)$

$$P_{Y^n/X^n}(y^n/x^n) = \prod_{i=1}^{n} P_{Y_i/X_i}(y_i/x_i)$$

Channel Coding Theorem

For every $R < C$, where C is the capacity of the channel (rate), there exists a sequence of codes C_n where

C_n has length n codewords & $|C_n| = 2^{\lceil nR \rceil}$ such that $P_e(C_n) \to 0$ as $n \to \infty$.

If $R = 0.5$, $C > 0.5$

$n = 2, 4, 5$

$|C_1| = 2$, $|C_2| = 4$, $|C_3| = 8$, ...

Second part of channel coding theorem says that if $R \geq C$, no matter what sequence of codes we take, $P_e > 0$.

Capacity of DMC

$$C = \max_{P_X} I(X;Y)$$

the channel is specified by only conditional probability.

Calculating Capacity of BSC

$$\max_{P_X} I(X;Y)$$

$$= \max_{\alpha} I(X;Y)$$

$$= \max_{\alpha} H(Y) - H(Y/X)$$

$$H(Y/X) = P_X(0) H(Y/X=0) + P_X(1) H(Y/X=1)$$

$P_X(0) = \alpha$,
$P_X(1) = 1-\alpha$.

$P_{Y/X}(0/0) = P_{Y/X}(1/1) = p$

$P_{Y/X}(1/0) = P_{Y/X}(0/1) = 1-p$.

$$H(Y/X=0) = -(1-p)\log(1-p) - p\log p = h_2(p)$$

(Binary entropy function)

$$H(Y/X=1) = h_2(p)$$

$$H(Y/X) = P_X(0) h_2(p) + P_X(1) h_2(p)$$

$= \alpha h_2(p) + (1-\alpha) h_2(p)$

$H(Y/X) = h_2(p)$

$\max_{\alpha} (H(Y) - h_2(p))$

↳ not dependent on α

so we want to maximize it

$H(Y) \leq \log_2 M$

where M is the number of values which Y is taking.

$\therefore H(Y) \leq 1$

$H(Y) = 1$ when $P_Y(0) = 1/2$, $P_Y(1) = 1/2$.

$P_Y(0) = 1/2 = P_X(0) P_{Y/X}(0/0) + P_X(1) P_{Y/X}(0/1)$

$= \alpha(1-p) + (1-\alpha)p = 1/2$

$P_Y(1) = 1/2 = P_X(0) P_{Y/X}(1/0) + P_X(1) P_{Y/X}(1/1)$

$= \alpha p + (1-\alpha)(1-p) = 1/2$

$\Rightarrow \alpha - \alpha p + p - \alpha p = 1/2$

$\alpha + p - 2\alpha p = 1/2 \longrightarrow (1)$

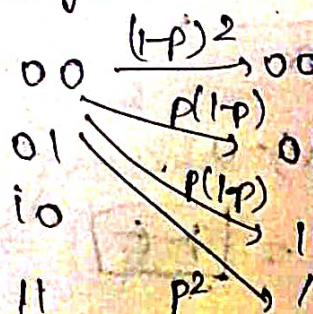
$\alpha p + 1 - \alpha - p + \alpha p = 1/2 \longrightarrow (2)$

$\Rightarrow \alpha + p - 2\alpha p = 1/2$

$1 - \alpha - p + 2\alpha p = 1/2$

So $\max_{\alpha} (H(Y) - H_1(Y/X)) = \max_{\alpha} 1 - h_2(p)$

4-ary symmetric channel.



$H(Y/X) = h_4(\,)$

$H(Y/X=00)$

$= H(Y/X=01) = H(Y/X=10)$

$= H(Y/X=11)$

$$\max_{\alpha} H(Y) - H(Y/X)$$

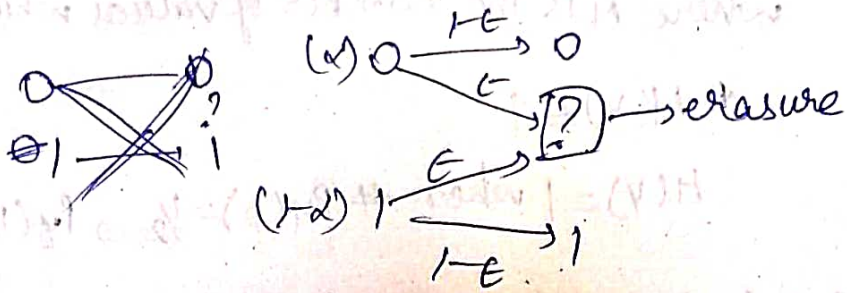
$$\Rightarrow H(Y) \leq 2 \Rightarrow \text{H.t}$$

$$= \max_{\alpha} 2 =$$

Tut →

1.) $$C = \max_{P_X} I(X;Y) = \max_{P_X} H(Y) - H(Y/X)$$

$$= 1 - \epsilon$$



$$C = \max_{P_X} I(X;Y)$$

$$= \max_{P_X} H(X) - H(X/Y) \neq \max_{P_X} H(X) -$$

$$= \max_{P_X} H(X) - \epsilon H(X/Y=?) - (1-\epsilon) H(X/Y \neq ?)$$

$\max_{P_X} H(X/Y)$

$H(X/Y=?) = P(Y=?)\, H(X/Y=?)$ (with $P(Y=?) = \epsilon$)

$+ P(Y \neq ?)\, H(X/Y \neq ?)$

$Y = X$

$H(X/X) = 0$

$$P(Y=?) = P(Y=?/X=0)P(X=0) + P(Y=?/X=1)P(X=1)$$

$$= t_1(\alpha) + t(1-\alpha) = \epsilon$$

$H(X/Y=?) \Rightarrow P_{X \neq Y=?}$ ~~$(\frac{1}{2}, \frac{1}{2})$~~ (from Baya's theorem)

$$= (P_X(0), P_X(1))$$

P_X

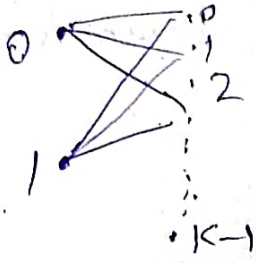
$$\Rightarrow = \max_{P_X} H(X) - \epsilon H(X) - 0$$

$$= \max_{P_X} H(X)(1-\epsilon) = \boxed{1-\epsilon}$$

capacity → max no. of bits that can be transferred reliably across the channel.

$N\epsilon$ bits are erased.
$N(1-\epsilon)$ bits are unerased.

2.)



$|y| = K < \infty$

K outputs

$Q(y/0) = Q(g(y)/1)$,

where $g: y \to y$ is one-one

g is onto → bijective

$$C = \max_{P_X} I(X;Y)$$

$$= \max_{P_X} H(Y) - H(Y/X)$$

$$Q = Q_{Y/X} = \begin{bmatrix} Q_{Y=0/X=0} & \cdots & Q_{Y=K-1/X=0} \\ Q_{Y=0/X=1} & \cdots & Q_{Y=K-1/X=1} \end{bmatrix}$$

$$P(X=0)H(Y/X=0) + P(X=1)H(Y/X=1)$$

$$Q = \begin{matrix} \\ 0 \\ 1 \end{matrix} \begin{bmatrix} 0 & 1 & 2 & \cdots & K-1 \\ & R_1 & & & \\ & R_2 & & & \end{bmatrix}$$

R_1 is a permutation of R_2

$P_{Y/X=0} = ($ ( ) $)$ — K entries, 2 permutations

P → bold font used to represent probability measure

$P_{Y/X=1} = ($ $)$

So $H(Y/X=0) = H(Y/X=1)$

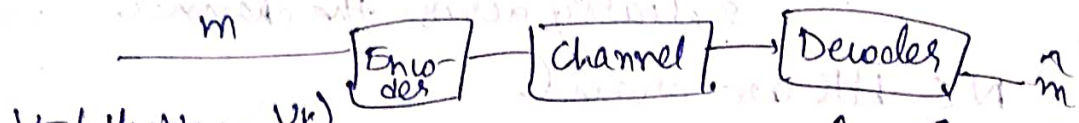
$$C = \max_{P_X} H(Y) - H(Y/X)$$

$$= \left(\max_{P_X} H(Y)\right) - H(Y/X=0) = \log K - H(Y/X=0)$$

$P_X = \left(\frac{1}{2}, \frac{1}{2}\right)$ gives rise to

$P_Y = \left(\frac{1}{K}, \frac{1}{K}, \cdots \frac{1}{K}\right)$

3.)



$V=(V_1, V_2, \ldots V_k)$ — k-tuple

$\hat{V}=(\hat{V}_1, \ldots \hat{V}_k)$

k bits form a message

$\{ P_{block} \triangleq P(V \neq \hat{V}) \}$ → the whole bit stream is different

(block error probability)

Symbol error prob / bit error prob

$\{ P_{bit} = P_{symb} \triangleq \frac{1}{k} \sum_{j=1}^{k} P(\hat{V}_j \neq V_j) \}$

* Show that: $P_{block} \geq P_{symb} \geq \frac{P_{block}}{k}$

$\{V \neq \hat{V}\} \Rightarrow \{V_j \neq \hat{V}_j\}$ for some j.

$P_{block} = P\left(\bigcup_j \{V_j \neq \hat{V}_j\}\right) \leq \sum_j P(\{V_j \neq \hat{V}_j\})$

$\{V \neq \hat{V}\} = \bigcup_j \{V_j \neq \hat{V}_j\}$

$\sum_j P(\{V_j \neq \hat{V}_j\})$ — $k \cdot P_{symb}$.

$\Rightarrow P_{symb} \geq \frac{P_{block}}{k}$

$\mathbb{1}\{\hat{V} \neq V\} \geq \mathbb{1}\{V_j \neq \hat{V}_j\}, \forall j$

↙ "indicator" RV.

$\left(\begin{array}{l} =1 \text{ if } \hat{V}=V \\ =0 \text{ if } \hat{V} \neq V \end{array}\right)$

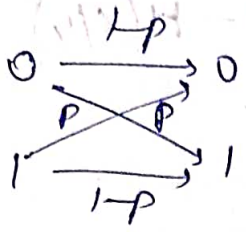
$\Rightarrow \mathbb{1}\{\hat{V} \neq V\} \geq \frac{1}{k} \sum_{i=1}^{k} \mathbb{1}\{V_j \neq \hat{V}_j\}$

$\mathbb{E}[\mathbb{1}\{\text{event}\}] = P(\text{event})$

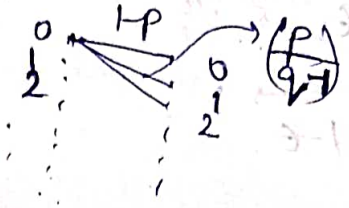
$\Rightarrow P(\hat{v} \neq v) \geq \frac{1}{k} \sum P(v_j \neq \hat{v}_j)$

$\Rightarrow P_{block} \geq P_{symb}$

4.)



BSC



q-ary symmetric channel

MAP: "maximum aposteriori"

decoder $\max_x P(x/y)$ (event which has happened before)

ML → "maximum likelihood"

$\max_x P(y/x)$ — n inputs → n outputs

likelihood $(p \leq 1 - \frac{1}{q})$

$$\max_x P(y/x) = \max_x \prod_{i=1}^{n} P(y_i/x_i)$$

$$= (1-p)^{n-d_H(x,y)} \left(\frac{p}{q-1}\right)^{d_H(x,y)}$$

$d_H(x,y) = \#\{i, x_i \neq y_i\}$

max Hamming distance

$$= (1-p)^n \left(\frac{p}{(1-p)(q-1)}\right)^{d_H(x,y)}$$

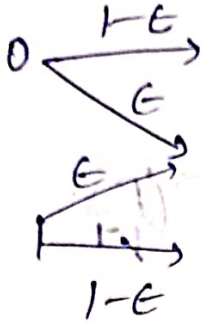
≤ 1 — minimize.

Given: $p \leq 1 - \frac{1}{q}$

$\Rightarrow \frac{p}{(1-p)(q-1)} \leq 1$

ML decoder ≡ min. d_H decoder

Channel capacity



$$C = \max_{P_x} I(X;Y)$$

$$= \max_{P_x} H(X) - H(X/Y)$$

$$H(X/Y=0) = 0 = H(X/Y=1)$$

$P_x(0) = \alpha$, $P_x(1) = (1-\alpha)$.

$$H(X/Y=E) = \alpha H(X=0/Y=E) + (1-\alpha) H(X=1/Y=E)$$

$$= \alpha\epsilon + (1-\alpha)\epsilon = \epsilon$$

$$P_{X/Y}(x/y) = \frac{P_{XY}(x,y)}{P_Y(y)} = \frac{P_{Y/X}(y/x) \cdot P_X(x)}{P_Y(y)}$$

$$P_X(0/E) = P_X(1/E) = 1/2$$

$$H(X/Y) = H(X/Y=E) \cdot P_Y(E) + H(X/Y \neq E) \cdot P_Y(Y \neq E)$$

$$\Rightarrow C = \max_{P_x} H(X) - H(X/Y)$$

$$= \max_{P_x} H(X) - \epsilon H(X) = \max_{P_x} (1-\epsilon) H(X)$$

$$= (1-\epsilon) \quad (\text{as } \max H(X) = 1)$$

Error Correcting Codes

{Information Theoretic Secrecy (cryptography)
(computationally bounded/unbounded adversary)
Secrecy capacity, one-time pad. }

$\mathbb{Z}_p$ (prime fields)

codes → Hamming code
Reed Solomon
Reed Mousely

Binary Field $(\mathbb{F}_2, +, \cdot)$

+	0	1
0	0	1
1	1	0

(XOR)

·	0	1
0	0	0
1	0	1

Prime Field

$\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ → all elements are prime numbers.

$(a+b)\%$ mod p.

$(a \cdot b)\%$ mod p

There exists a multiplicative inverse for any element in $\mathbb{Z}_p$ such that

$a \cdot b = 1 \bmod p$

as p is prime, $\gcd(a, p) = 1$

$\Rightarrow ax + py = 1$

for any arbitrary integers x, y

$\Rightarrow ax \bmod p = 1$

$(ax) \bmod p = ((a \bmod p)(x \bmod p)) \bmod p$.

Vector Space = set of vectors.

$(V, +, F_2, \cdot)$

Two operations → vector addition, scalar multiplication

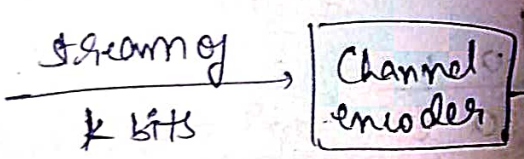
$\alpha \, v \in V$

$(F_2^n, +, F_2, \cdot)$ → Example of vector space over F_2.

$$u+v = \begin{matrix}(u_1, u_2, \ldots u_n) \\ +(v_1, v_2, \ldots v_n)\end{matrix} \} \rightarrow \text{component wise addition.}$$

$$\alpha v = (\alpha v_1, \alpha v_2, \ldots \alpha v_n)$$

1.) Linear Block Codes



Eg. if we have 100 bits, $k = 4$, then we divide 100 into 25 parts of 4 bits each and if $n = 7$, then each of the 4 bits is mapped to 7 bits.

Channel encoder = matrix multiplication

Linear Block Code over F_2 is a subspace of vector space F_2^n.

k – dimensional subspace of F_2^n.

Parameters of Linear Block Code

1.) Length of the code, n.

2.) Dimension of the code k.

3.) Rate of the code = k/n.

4.) Minimum distance of the code c,

$$d_{min} = \min_{x \neq y} d_H(x, y)$$

$x, y \in c$

$d_H(x,y)$ = no. of positions in which x and y differ.

Eg. 1 1 0 1 1 0 0 1 0 → q
1 1 1 0 0 1 0 1 1 1

Hamming distance = 5

Generator matrix of a code
(denoted by G, $(k \times n)$ matrix)

$G \longrightarrow k \times n$.

It is a matrix in which k rows are k basis elements of code C. (basis of subspace)

2) Repetition Code

$C = \{00\ldots0, 11\ldots1\}$.

length of code = n.

$k = 1$ (no. of basis elements in $C = 1$)

Rate $= \frac{1}{n}$.

Min distance = n

$G = [1\ 1\ \ldots\ 1]_{1 \times n}$.

Simple parity check code

all vectors will have even no. of one's

$C = \{(m_1, m_2, \ldots m_{n-1}, \bigoplus_{i=1}^{n-1} m_i)\}$

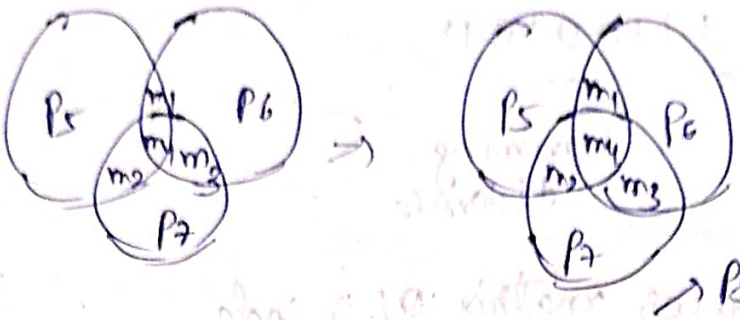
length = n $(m_1 \oplus m_2 \oplus \ldots \oplus m_{n-1})$

Dimension = n−1

Minimum distance = 2. (if one of m_i changes then m_n also changes (parity) so 2 bits are changed)

$G = \left(\begin{array}{ccccc|c} 1 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 1 & 0 & & \vdots & 1 \\ \vdots & & 1 & & & \vdots \\ & & & & & \\ 1 & & & & 1 & 1 \end{array}\right)$

3.) Hamming code (code of 3 circles)



$C = \{ (m_1, m_2, m_3, m_4, m_1+m_2+m_4, m_1+m_3+m_4, m_2+m_3+m_4) \}$.

($m_1+m_2+m_4$ → P_5; $m_1+m_3+m_4$ → P_6; $m_2+m_3+m_4$ → P_7)

$m_i \in \mathbb{F}_2$.

Length = 7.

dimension = 4

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

m_1 m_2 m_3 m_4 $m_1+m_2+m_4$ $m_1+m_3+m_4$ $m_2+m_3+m_4$

Minimum distance = 3. (atmost 3).

G is a matrix (which always contains a zero vector because it is a subspace, so maximum of minimum distance = 0)

⇒ Single error correcting code and also a double error detecting code. (Hamming code)

If a code has minimum distance d_{min}, then

No. of errors you can detect = $d_{min} - 1$

No. of errors you can correct = $\left\lfloor \frac{d_{min}-1}{2} \right\rfloor$ (floor)

d_{min} is the minimum value such that any $n - d_{min} + 1$ columns of G matrix have rank $= k$.

Theorem:

Linear Block Codes

→ subspace of $\mathbb{F}_2^n$ (or $\mathbb{F}_q^n$)

→ length n, dimension k, minimum distance d_{min}

Generator matrix G.

1.) Repetition

2.) Simple parity check.

3.) (7,4) Hamming code.

→ Generator matrix of Hamming Code

$$G = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right]$$

$v_1 + v_2 + v_3 + v_4 = 0$

So $n - d_{min} + 1 \neq 4$

$n - d_{min} + 1 = 5$

Applying Rank $= k = 4$ (dimension)

Theorem: d_{min} is the minimum +ve integer such that any $n - d_{min} + 1$ columns of generator matrix have rank k.

$$7 - d_{min} + 1 = 5 \Rightarrow \boxed{d_{min} = 3}$$

Two things to show:

1) Any $n - d_{min} + 1$ columns of G have rank k.

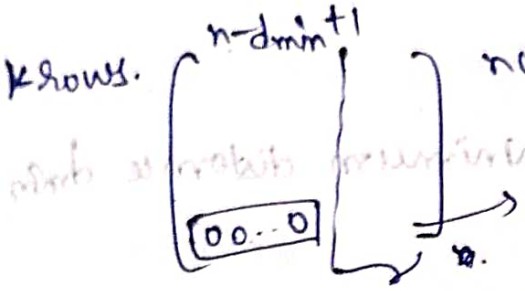
2) There exists a set of $n - d_{min}$ columns which do not have rank k.

Hamming weight → no. of non-zero elements.

Eg. 1110011
Hamming weight = 5.

Proof:

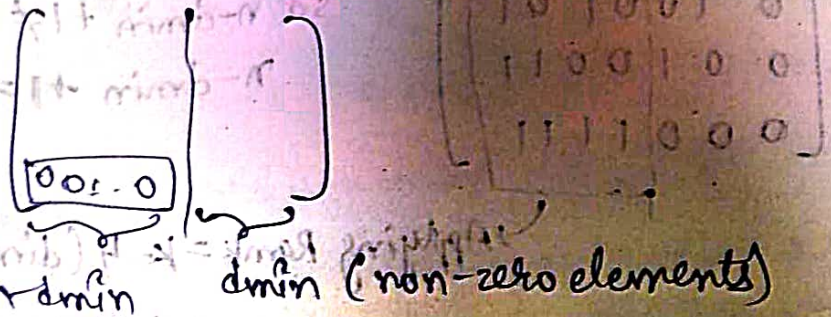
1.) if rank < k, then for another basis for the same code, the row-reduced form of the matrix will have all zeroes in the last row.



no. of non-zero elements in dmin-1 columns can be $\leq$ dmin-1

So maximum rank possible = dmin-1 which is a contradiction as dmin-1 cannot be the minimum distance, because by definition, dmin is the minimum distance.

2.)



now if dmin is the minimum distance, then Hamming weight = dmin has non-zero elements. Then the row-reduced matrix obtained from the n-dmin columns gives (0,.. 0) at the last row, because if it does not form ↗, then there is a non-zero element, which means that dmin+1 = Hamming weight which contradicts the fact that dmin = Hamming weight and hence, the rank of n-dmin columns is always less than k.

Extended Hamming Code

(parity bits)

$(m_1, m_2, m_3, m_4, P_5, P_6, P_7, \Sigma m_i + \Sigma P_i)$

Length = 8, Dimension = 4, Min distance = 4

(if any one of m_i is changed, then 2 P_i will change so in $\Sigma m_i + \Sigma P_i$, 3 bits will change, and hence this bit will also change and hence 4 bits change in total, so the min dist = 4.)

If weight$(m_1, \ldots \Sigma m_i + P_i) \geq 4$

then m_i parity bit will not change.

If weight$(m_1, \ldots \Sigma m_i + \Sigma P_i) = 3$

then parity bit also changes and hence weight becomes 4.

Read–Solomon Codes

$\mathbb{F}_p \rightarrow p$ is prime

$\mathbb{F}_p^n \rightarrow$ is vector space

length $= n$, dimension $= k$, d_{min} = min dist.

$G_{k \times n} \rightarrow$ Generator matrix.

$(m_0, m_1, \ldots m_{k-1})$, $m_i \in \mathbb{F}_p$

Message polynomial.

$$f(x) = \sum_{i=0}^{k-1} m_i x^i$$

Let $\alpha_1, \ldots \alpha_n$ be n distinct elements in $\mathbb{F}_p$.

codeword $\rightarrow C$ $\longrightarrow \mathbb{F}_p^n$

$$C = (f(\alpha_1), f(\alpha_2), \ldots f(\alpha_n))$$

Evaluating the message polynomial of n distinct points in the field.

1.) $C = mG$.

$$[C_1\ C_2 \cdots C_n] = [m_1\ m_2 \cdots m_k]\, G$$

$1 \times n$ $\quad$ $1 \times k$ $\quad$ $k \times n$

$$= [m_1\ m_2 \cdots m_k] \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_n^{k-1} \end{pmatrix}$$

$C_1 = f(\alpha_1),\ C_2 = f(\alpha_2), \ldots C_n = f(\alpha_n)$

$d_{min} = n - k + 1$.

Theorem: d_{min} is the min value such that any $n - d_{min} + 1$ columns of G have rank k.

$d_{min} = n - k + 1$.

$\Rightarrow$ $\boxed{n - d_{min} + 1 = k.}$

Any $(k-1)$ degree polynomial can have atmost $(k-1)$ roots

$(f(\alpha_1)\ f(\alpha_2) \cdots f(\alpha_n))$.

So $(k-1)$ zeroes are possible, so the weight can be $k-1$ (maximum).

So the maximum no. of $(k-1)$ zeroes for

Min distance $\geq n - k + 1$ (maximum weight)

$(k-1)$ zeroes, $(n-k+1)$ non-zeroes.

$$f(x) = \prod_{i=0}^{k-1} (x - \alpha_i)$$

$\alpha_1, \alpha_2, \ldots \alpha_n$

so for $(k-1)$ columns, it becomes zero as $(x - \alpha_i) = 0$ as α_i is a root, the rest $(n-k+1)$ columns produce non-zero answers.