Coding and Information Theory
Chapter 6:
Error-correcting Codes

Xuejun Liang

2019 Fall

Chapter 6: Error-correcting Codes

1. Introductory Concepts
2. Examples of Codes
3. Minimum Distance
4. Hamming's Sphere-packing Bound
5. The Gilbert-Varshamov Bound
6. Hadamard Matrices and Codes

The aim of this chapter

- Is to construct codes C with good transmission-rates R and low error-probabilities Pr_E, as promised by Shannon's Fundamental Theorem.
  - This part of the subject goes under the name of Coding Theory (or Error-correcting Codes), as opposed to Information Theory.
- Will concentrate on a few simple examples to illustrate some of the methods used to construct more advanced codes

Finite Field and Linear Space

- A set F is a Field
  - At least two elements $0, 1 \in F$
  - Two operations $+$ and $\times$ on F
  - Associative and commutative
  - Operation $\times$ distributes over $+$
  - 0 is the identity for $+$ and 1 for $\times$
  - Additive inverse and multiplicative inverse

Finite Fields

Goal: Given a prime p and a positive integer n, construct a field with p^n elements.

Definitions and Notations:

$Z_p[x]$: all polynomials in the indeterminate x with coefficients in Z_p.

deg(f) : the degree of f ($f \in Z_p[x]$) is the largest exponent in a term of f.

$f \mid g$: f divides g ($f, g \in Z_p[x]$), if $g = f \cdot h$ for some $h \in Z_p[x]$.

$g \equiv h \pmod{f}$: $f \mid (g - h)$ ($f, g, h \in Z_p[x]$ and def(f)$\geq$1)

$Z_p[x]/(f)$: all congruence classes modulo f in $Z_p[x]$ ($f \in Z_p[x]$).

$Z_p[x]/(f)$ is equipped with $+$, $\times$ and $|\, Z_p[x]/(f) \,| = p^n$, where n=deg(f)

Finite Fields (Cont.)

Example: $Z_3[x]/(x^2-1)$

List all the elements in forms $a_0 + a_1x$, $a_0, a_1 \in Z_3$.

List a complete multiplication table.

In general $Z_p[x]/(f)$ is a ring, not a field.

Definition: A polynomial f in $Z_p[x]$ is called irreducible, if f can not be written as $f = f_1 \cdot f_2$ where $\deg(f_1)>0$ and $\deg(f_2)>0$.

Fact: If f in $Z_p[x]$ is irreducible polynomial of degree n, then $Z_p[x]/(f)$ is a field with p^n elements.

Notation: $Z_p[x]/(f)$ is called Galois field and is denoted by $GF(p^n)$.

Linear (vector) space: Definition

A linear space V over a field F is a set whose elements are called vectors and where two operations, addition and scalar multiplication, are defined:

1. **addition**, denoted by $+$, such that to every pair $x, y \in V$ there correspond a vector $x + y \in V$, and
   - $x + y = y + x$,
   - $x + (y + z) = (x + y) + z$, $x, y, z \in V$;

   $(X, +)$ is a group, with neutral element denoted by 0 and inverse denoted by $-$, $x + (-x) = x - x = 0$.
2. **scalar multiplication** of $x \in V$ by elements $k \in F$, denoted by $kx \in V$, and
   - $k(ax) = (ka)x$,
   - $k(x + y) = kx + ky$,
   - $(k + a)x = kx + ax$, $x, y \in V$, $k, a \in F$.

   Moreover $1x = x$ for all $x \in V$, 1 being the unit in F.

- Example: V_4 of all 4-tuples over Z_2 (GF(2)).

Subspace and Linearly independent

- Subspace: $S \subseteq V$
  - addition and scalar multiplication are closed in S
- Linear combination
  - $a_1v_1+a_2v_2+ \dots + a_nv_n$
  - Linearly independent of $v_1, v_2, \dots, v_n$
    - If $a_1v_1+a_2v_2+ \dots + a_nv_n = 0$ then $a_1=0, a_2=0, \dots, a_n=0$.
  - Linearly dependent of $v_1, v_2, \dots, v_n$
    - There are $a_1, a_2, \dots, a_n$ (not all 0's) such that $a_1v_1+a_2v_2+ \dots + a_nv_n = 0$
- Example 4.11: determine if the vectors are linearly dependent or not
- Example 4.12: determine if the vectors are linearly dependent or not

Basis and Dimension

- Basis (or Base)
  - Basis: independent vectors that can span the whole vector space.
  - Any vector is a linear combination of basis vectors.
- Dimension
  - Number of vectors within the basis
  - Example: V_n is n-dimension
- Example 4.13: determine a basis and the dimension of the subspace in V_4 over Z_2 consisting of vectors:

  (0 0 0 0) (1 1 0 0) (1 0 1 0) (0 0 0 1)

  (0 1 1 0) (1 1 0 1) (1 0 1 1) (0 1 1 1)

Orthogonality and Dual Space

- Orthogonality
  - **Inner product** of $\mathbf{u} = (u_0, u_1, \ldots, u_{n-1})$ and $\mathbf{v} = (v_0, v_1, \ldots, v_{n-1})$:
    $$\mathbf{u.v} = u_0v_0 + u_1v_1 + \ldots, + u_{n-1}v_{n-1}$$
  - **u** and **v** are said orthogonal if $\mathbf{u.v} = 0$
  - Subspace S and P of V_n are said orthogonal if for any $\mathbf{u} \in S$ and any $\mathbf{v} \in P$, we have $\mathbf{u.v} = 0$
- Dual Space
  - Subspace S of V_n is the dual space (null space) of another subspace P of V_n if S and P are orthogonal and
    $$\dim(S)+\dim(P) = n$$
- Example 4.14: show S and P are dual each other
  - S = {(0 0 0 0), (1 1 0 0), (1 0 1 1), (0 1 1 1)
  - P = {(0 0 0 0), (1 1 0 1), (1 1 1 0), (0 0 1 1)

Raw space and Column Space

- Let G be a m×n matrix
  - All linear combinations of row vectors of G is a subspace of $V_{n,}$ called **row vector space** of G.
  - All linear combinations of column vectors of G is a subspace of V_m called **column vector space** of G.
  - The dimension of row vector space is called **row rank** and the dimension of column vector space is called **column rank**.
  - Row rank and column are always equal, it is called **the rank of the matrix.**
- Elementary row operations of a matrix
  - swap two rows, multiply a row with a scalar, add multiple of a row to another
- Elementary row operations do not change the row rank.

Raw space and Column Space (cont.)

- Example 4.15:
  - Determine the row space of matrix G= $\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$
- Example 4.16:
  - Consider the G in 4.15. Compute a matric G' by adding row 3 of G to row 1 of G and then interchanging rows 2 and 3 of G.
  - Show that the row space of G' is the same as that generated by G.

6.1 Introductory Concepts

- Assume channel Γ has input A and output B, and A = B = F, where F is a finite field.
- Note Z_p of integers mod (p) is a finite field, where p is a prime number.
- Theorem 6.1
  a) There is a finite field of order q if and only if $q = p^e$ for some prime p and integer $e \geq 1$.
  b) Any two finite fields of the same order are isomorphic.

Galois Field

- The essentially unique field of order q is known as the Galois field F_q or $GF(q)$.
  - When e = 1, then q = p and $F_q = F_p = Z_p$.
  - When e > 1, $F_q = Z_p[x]/f(x)$, where f(x) is an irreducible polynomial of degree e on the field Z_p.
  - When e > 1, $F_q = Z_p[\alpha]$, where α is a root of $f(x)$ which an irreducible polynomial of degree e on the field Z_p.
- Example 6.2
  - The quadratic polynomial $f(x) = x^2 + x + 1$ has no roots in the field Z_2.

$$F_4 = \{a + b\alpha \mid a, b \in \mathbf{Z}_2\} = \{0, 1, \alpha, 1 + \alpha\}$$

Linear Code

- Let F be a field, then the set $V = F^n$ of all n-tuples with coordinates in F is an n-dimensional vector space over F.
  - the operations are component wise addition and scalar multiplication
- Assume that any code-words in C are of length n
  - So C is a subset of the set $V = F^n$
- We say that C is a linear code (or a group code) if C is a non-empty linear subspace of V.
  - If $\boldsymbol{u}, \boldsymbol{v} \in C$ then $a\boldsymbol{u} + b\boldsymbol{v} \in C$ for all $a, b \in F$

The rate of a code C

- We will always denote $|C|$ by M
- When C is linear we have M = q^k, where k = dim(C) is the dimension of the subspace C.
  - We then call C a linear $[n, k]$-code.
- The rate of a code C is $$R = \frac{\log_q M}{n} \qquad (6.1)$$
  - So in the case of a linear $[n, k]$-code we have

k information digits, carrying the information
n - k check digits, confirming or protecting that information

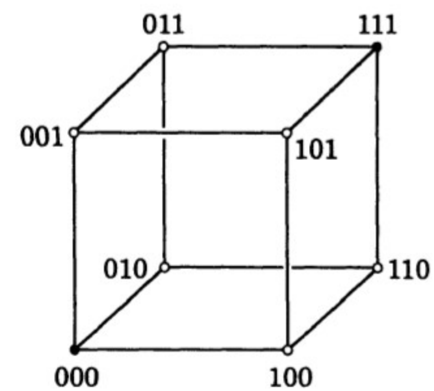
$$R = \frac{k}{n} \qquad (6.2)$$

Notes

- We will assume that all code-words in C are equiprobable, and that we use nearest neighbor decoding (with respect to the Hamming distance on V).

6.2 Examples of Codes

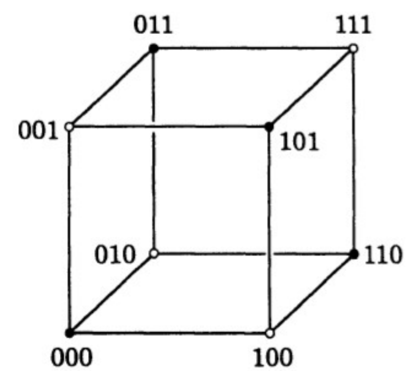
- Example 6.3: The repetition code R_n over F
  - the words $u = uu \ldots u \in V = F^n$, where $u \in F$, so M = $|F|$ = q.
  - If F is a field then R_n is a linear code of dimension k = 1, spanned by the word (or vector) 11. . . 1
  - Example:
    - Binary code R_3 = {000, 111} as a subset of $V = Z_2^3$



  - R_n corrects $\lfloor (n-1)/2 \rfloor$ errors
  - R_n has rate R = $1/n \to 0$ as $n \to \infty$,

Examples of Codes (Cont.)

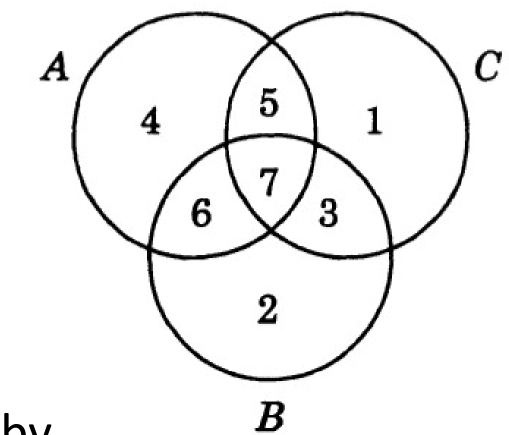
- Example 6.4: The parity-check code P_n over a field $F = F_q$
  - all vectors $u = u_1 u_2 \dots u_n \in V$ such that $\sum_i u_i = 0$.
  - if n = 3 and q = 2
    then P_3={000, 011,101, 110}.



  - $M = q^{n-1}$
  - R = (n - 1)/n, so $R \to 1$ as $n \to \infty$
  - it will detect a single error, but cannot correct it.

Hamming Code

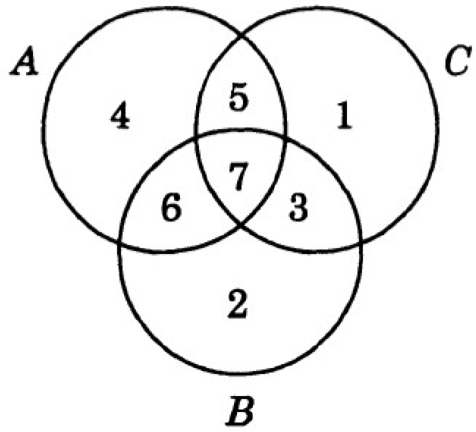
- Example 6.5
  - The binary Hamming code H_7 is a linear code of length n = 7 over F_2
    - 4 bits for data **a** = $a_1a_2a_3a_4$
    - 3 bits for checking
  - How to construct the code for **a**
    - Let the code word **u** = $u_1u_2u_3u_4u_5u_6u_7$
    - Bits $u_3 = a_1$, $u_5 = a_2$, $u_6 = a_3$, and $u_7 = a_4$
    - Bits u_1, u_2, u_4 for checking, determined by



$$u_4 + u_5 + u_6 + u_7 = 0$$
$$u_2 + u_3 + u_6 + u_7 = 0$$
$$u_1 + u_3 + u_5 + u_7 = 0$$

ABC
A=4, B=2, C=1

Hamming Code (Cont.)



- Example 6.5
  - Example: **a** = 0110

	1	2	3	4	5	6	7
	001	010	011	100	101	110	111
4 (s_1)				100	100	100	100
2 (s_2)		010	010			010	010
1 (s_3)	001		001		001		001
u	1	1	0	0	1	1	0

$$s_1 = u_4 + u_5 + u_6 + u_7$$
$$s_2 = u_2 + u_3 + u_6 + u_7$$
$$s_3 = u_1 + u_3 + u_5 + u_7$$

- The receiver will compute s_1, s_2, s_3. If they are all zero then the code is no error.
- If not, the binary number $s_1s_2s_3$ tells which bit is wrong.
- Now, assume **v** = 1110110 is received with 1-bit error in bit 3. you will get s_1= 0, s_2= 1, and s_3 = 1. So, $s_1s_2s_3$ = 011 = 3.

Hamming Code (Cont.)

- Example 6.5 (Cont.)
  - The binary Hamming code H_7 is a linear code with dimension k = 4.
    - $M = |H_7| = 16 = 2^4$
    - It can be generated by

      $\mathbf{u_1}$ = 1110000, $\mathbf{u_2}$ = 1001100, $\mathbf{u_3}$ = 0101010, $\mathbf{u_4}$ = 1101001
    - which are obtained from

      $\mathbf{e_1}$ = 1000, $\mathbf{e_2}$ = 0100, $\mathbf{e_3}$ = 0010, $\mathbf{e_4}$ = 0001
- Note:
  - Although the binary codes R_3 and H_7 both correct a single error, the rate R = 4/7 of H_7 is significantly better than the rate 1/3 of R_3.

Examples of Codes (Cont.)

- Example 6.6
  - Suppose that C is a code of length n over a field F. Then we can form a code of length n + 1 over F, called **the extended code** $\bar{C}$. by
    - adjoining an extra digit u_{n+1} to every code-word $\boldsymbol{u}$ = $u_1 u_2 \dots u_n \in C$, chosen so that $u_1 + u_2 + \cdots + u_{n+1} = 0$.
    - Clearly $|\bar{C}| = |C|$, and if C is linear then so is $\bar{C}$, with the same dimension
- Example 6.7
  - If C is a code of length n, we can form a **punctured code** C° of length n - 1 by
    - choosing a coordinate position i, and deleting the symbol u_i from each codeword $u_1 u_2 \dots u_n \in C$.

6.3 Minimum Distance

- Define the minimum distance of a code C to be

$$d = d(\mathcal{C}) = \min\{d(\mathbf{u}, \mathbf{u}') \mid \mathbf{u}, \mathbf{u}' \in \mathcal{C},\ \mathbf{u} \neq \mathbf{u}'\}, \qquad (6.3)$$

- (n, M, d)-code
  - A code of length n, with M code-words, and with minimum distance d.
- [n, k, d]-code
  - A linear (n, M, d)-code, of dimension k.
- Our aim is to choose codes C for which d is large, so that $\Pr_E$ will be small.

Minimum Distance (Cont.)

- Define the weight of any vector $v = v_1 v_2 \dots v_n \in V$ to be

$$\mathrm{wt}(\mathbf{v}) = d(\mathbf{v}, \mathbf{0}), \tag{6.4}$$

- It is easy to see that for all $u, u' \in V$, we have

$$d(\mathbf{u}, \mathbf{u}') = \mathrm{wt}(\mathbf{u} - \mathbf{u}')$$

- Lemma 6.8
  - If C is a linear code, then its minimum distance d is given by

$$d = \min\{\mathrm{wt}(\mathbf{v}) \mid \mathbf{v} \in \mathcal{C}, \mathbf{v} \neq \mathbf{0}\}.$$

Minimum Distance (Cont.)

- We say that a code C corrects t errors, or is ***t*-error-correcting**, if, whenever a code-word $u \in C$ is transmitted and is then received with errors in at most t of its symbols, the resulting received word v is decoded correctly as u.
- Equivalently, whenever $u \in C$ and v $\in V$ satisfy d(u, v) $\leq t$, the decision rule Δ gives Δ(v) = u.
- Example 6.9
  - A repetition code R_3 corrects one error, but not two.

Minimum Distance (Cont.)

- If u is sent and v is received, we call the vector e = v - u the **error pattern**.
  - A code corrects t errors if and only if it can correct all error-patterns $e \in V$ of weight wt(e) $\leq t$.
- Theorem 6.10
  - A code C of minimum distance d corrects t errors if and only if $d \geq 2t$ + 1. (Equivalently, C corrects up to $\left\lfloor \frac{d-1}{2} \right\rfloor$ errors.)
- Example 6.11
  - A repetition code R_n of length *n* has minimum distance d = *n*, since d(u, u') = n for all u $\neq$ u' in R_n. This code therefore corrects t = $\lfloor (n-1)/2 \rfloor$ errors.

Minimum Distance (Cont.)

- Example 6.12
  - Exercise 6.3 shows that the Hamming code H_7 has minimum distance d = 3, so it has t = 1 (as shown in §6.2). Similarly, $\overline{H_7}$ has d = 4 (by Exercise 6.4), so this code also has t = 1.
- Example 6.13
  - A parity-check code P_n of length n has minimum distance d = 2; for instance, the code-words u =110 ... 0 and u' = 0 = 00 . . . 0 are distance 2 apart, but no pair are distance 1 apart. It follows that the number of errors corrected by P_n is 0.

Minimum Distance (Cont.)

- C detects d - 1 errors
- Example 6.14
  - The codes R_n and P_n have $d = n$ and 2 respectively, so R_n detects n-1 errors, while P_n detects one; $\overline{H_7}$ has d = 3, so it detects two errors.

6.4 Hamming's Sphere-packing Bound

- Define Hamming's sphere to be

$$S_t(\mathbf{u}) = \{ \mathbf{v} \in \mathcal{V} \mid d(\mathbf{u}, \mathbf{v}) \leq t \} \qquad (\mathbf{u} \in \mathcal{C}) \tag{6.5}$$

- We have

$$|S_t(\mathbf{u})| = 1 + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{t}(q-1)^t \tag{6.6}$$

- Theorem 6.15
  - Let C be a q-ary t-error-correcting code of length n, with M code-words. Then

$$M\left(1 + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{t}(q-1)^t\right) \leq q^n$$

Sphere-packing Bound (Cont.)

- Example 6.16
  - If we take q = 2 and t = 1 then Theorem 6.15 gives $M \leq 2^n/(1+n)$, so $M \leq \lfloor 2^n/(1+n) \rfloor$
    since M must be an integer. Thus
    $M \leq$ 1, 1, 2, 3, 5, 9, 16, ... for n = 1, 2, 3, 4, 5, 6, 7, ...
- Corollary 6.17
  - Every t-error-correcting linear [n, k]-code C over F_q satisfies

$$\sum_{i=0}^{t} \binom{n}{i} (q-1)^i \leq q^{n-k}$$

Sphere-packing Bound (Cont.)

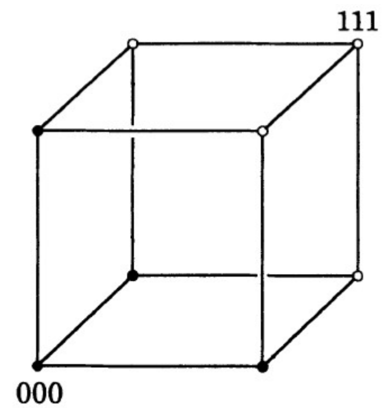
- Corollary 6.17 therefore gives us a lower bound on the number of check digits (n-k) required to correct t errors

$$n - k \geq \log_q \left(\sum_{i=0}^{t} \binom{n}{i} (q-1)^i \right)$$

- A code C is **perfect** if it attains equality in Theorem 6.15 (equivalently in Corollary 6.17, in the case of a linear code).

Sphere-packing Bound (Cont.)

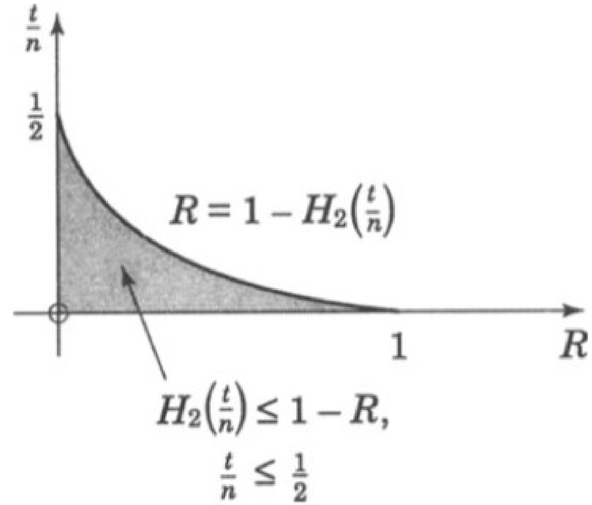
- Example 6.18
  - The binary repetition code R_n of odd length n is perfect!
  - However, when n is even or $q > 2$, R_n is not perfect.



- Example 6.19
  - The binary Hamming code H_7 is perfect.
- If C is any binary code then Theorem 6.15 gives

$$2^n \geq M\binom{n}{t} = 2^{nR}\binom{n}{t}$$

Sphere-packing Bound (Cont.)



- Thus $2^{n(1-R)} \ge \binom{n}{t}$
- So taking logarithms gives

$$1 - R \ge \frac{1}{n}\log_2\binom{n}{t}$$

- Apply Stirling's approximation $n! \sim (n/e)^n\sqrt{2\pi n}$ to the three factorials in $\binom{n}{t} = n!/t!\,(n-t)!$
- We get the Hamming's upper bound on the proportion t/n of errors corrected by binary codes of rate R, as $n \to \infty$.

$$H_2\left(\frac{t}{n}\right) \le 1 - R \qquad (6.7)$$

where H_2 is the binary entropy function.

6.5 The Gilbert-Varshamov Bound

- Let $A_q(n, d)$ denote the greatest number of code-words in any q-ary code of length n and minimum distance d, where $d \leq n$. Let $t = \lfloor (d-1)/2 \rfloor$, we have (by Theorem 6.10)

$$A_q(n,d)\left(1 + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{t}(q-1)^t\right) \leq q^n$$

- Example 6.20
  - If q = 2 and d = 3 then t = 1, so as in Example 6.16 we find that $A_2(n, 3) \leq \lfloor 2^n/(n+1) \rfloor$. Thus for n = 3, 4, 5, 6, 7, ... we have $A_2(n, 3) \leq$ 2, 3, 5, 9, 16, ...

The Gilbert-Varshamov Bound (Cont.)

- Theorem 6.21
  - If $q \geq 2$ and $n \geq d \geq 1$ then

$$A_q(n,d)\left(1+\binom{n}{1}(q-1)+\binom{n}{2}(q-1)^2+\cdots+\binom{n}{d-1}(q-1)^{d-1}\right) \geq q^n$$

- Example 6.22
  - If we take q = 2 and d = 3 again (so that t = 1), then for all n $\geq$ 3, we have

$$A_2(n,3)\left(1+n+\frac{n(n-1)}{2}\right) \geq 2^n$$

  - This gives $A_2(n,3) \geq$ 2, 2, 2, 3, 5, ... for n = 3, 4, 5, 6, 7,

The Gilbert-Varshamov Bound (Cont.)

- Two bounds on R

  $$R \geq 1 - H_2\left(\frac{d-1}{n}\right) \quad *$$

  where $d \leq \lfloor n/2 \rfloor$

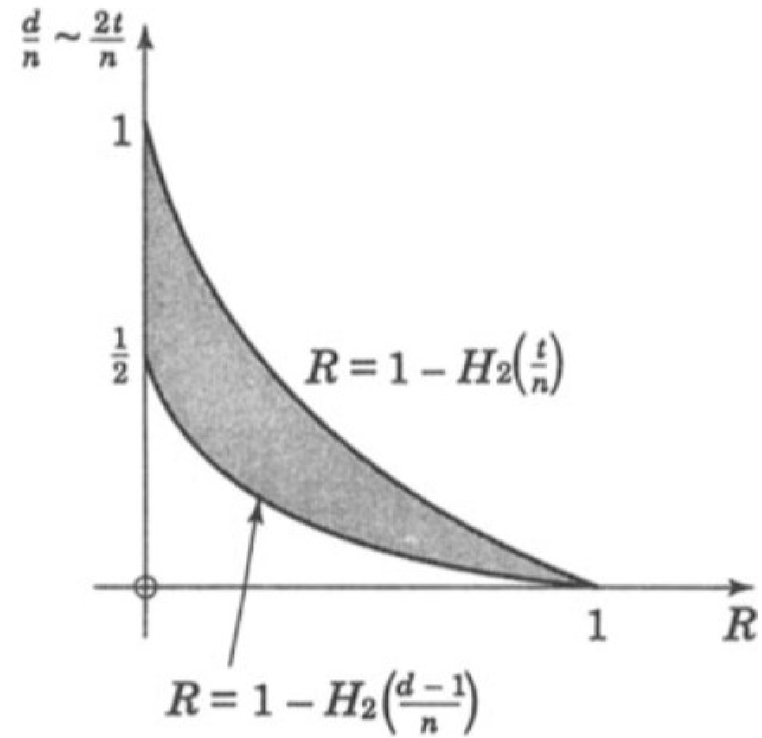
  $$R \leq 1 - H_2\left(\frac{t}{n}\right)$$ See (6.7)

  where $t = \lfloor (d-1)/2 \rfloor$

* Putting $\lambda = Q$, Exercise 5.7 gives

$$\sum_{i \leq nQ} \binom{n}{i} \leq 2^{nH(Q)}$$

for $Q < {}^1/_2$



6.6 Hadamard Matrices and Codes

- A real n x n matrix $H = \left(h_{ij}\right)$ (of order n) is called a Hadamard matrix, if it satisfies
  a) each h_{ij} = ±1, and
  b) distinct rows r_i, of H are orthogonal, that is, $r_i \cdot r_j = 0$ for all $i \neq j$.
- Note: $|\det(H)| = n^{n/2}$
- Example 6.23
  - The matrices H = (1) and $\begin{pmatrix} 1 & 1 \\ 1 & - \end{pmatrix}$ are Hadamard matrices of order 1 and 2, with $|\det H|$= 1 and 2 respectively.

Hadamard Matrices (Cont.)

- Lemma 6.24
  - Let H be a Hadamard matrix of order n, and let
$$H' = \begin{pmatrix} H & H \\ H & -H \end{pmatrix}$$
    Then H' is a Hadamard matrix of order $2n$.
- Corollary 6.25
  - There is a Hadamard matrix of order 2^m for each integer $m \geq 0$.
- Example 6.26
  - The Hadamard matrices of order 2^m obtained by this method are called Sylvester matrices. For instance, taking m = 1 or 2, ……

Hadamard Matrices and Codes

- Lemma 6.27
  - If there is a Hadamard matrix H of order n > 1, then n is even.
- Lemma 6.28
  - If there is a Hadamard matrix H of order n > 2, then n is divisible by 4.
- Theorem 6.29
  - Each Hadamard matrix H of order n gives rise to a binary code of length n, with $M = 2n$ code-words and minimum distance d = $n/2$.
- Any code C constructed as in Theorem 6.29 is called a Hadamard code of length n.

Hadamard Codes

- If n is not a power of 2 then neither is $2n$, so a Hadamard code of such a length n cannot be linear
- The transmission rate of any Hadamard code of length n is

$$R = \frac{\log_2(2n)}{n} = \frac{1 + \log_2 n}{n} \to 0 \quad \text{as} \quad n \to \infty$$

- The number of errors corrected (if $n > 2$) is

$$t = \left\lfloor \frac{d-1}{2} \right\rfloor = \left\lfloor \frac{n-2}{4} \right\rfloor = \frac{n}{4} - 1$$

- so the proportion of errors corrected is

$$\frac{t}{n} = \frac{1}{4} - \frac{1}{n} \to \frac{1}{4} \quad \text{as} \quad n \to \infty$$