



信息论与编码理论：L^AT_EX 排版

笔记：自用版

作者：xiaowen

时间：June 4, 2024



2021 级信息与计算科学专业

目录

第 1 章 引言	1
1.1 信息论的形成与发展	1
1.2 信息与编码理论的主要内容	2
1.3 本书基本内容与参考文献	4
第 2 章 信息量	6
2.1 熵	6
2.1.1 自信息量	6
2.1.2 不确定性(自信息量)的性质	7
2.1.3 自信息函数的推导	8
2.1.4 香农熵的定义 (Shannon 熵)	9
2.1.5 熵的简单性质与例子	9
2.2 联合熵与条件熵	11
2.2.1 联合熵	11
2.2.2 条件熵	12
2.2.3 联合熵和条件熵的关系	13
2.2.4 联合熵和条件熵的性质	13
2.3 熵的基本性质	14
2.3.1 对数函数的基本不等式	14
2.3.2 熵函数的性质	16
2.3.3 Fano 不等式 (不确定性的上界估计)	18
2.4 互熵与互信息	19
2.4.1 互熵	19
2.4.2 互信息	20
2.4.3 条件互信息	21
2.4.4 各种熵与互信息之间的关系	23
2.5 凸函数及其应用	23
2.5.1 凸函数的定义与判别	23
2.5.2 熵函数的凸性	24
2.6 连续型随机变量的信息量	25
2.6.1 连续型随机变量的 Shannon 熵	25
2.6.2 连续型随机变量的联合熵与条件熵	26
2.7 最大熵原理	26
2.7.1 有限区间情形的最大熵 (均匀分布时取最大熵)	26

2.7.2	半开区间情形的最大熵（指数分布时取最大熵）	27
2.7.3	全区间情形的最大熵（正态分布时取最大熵）	27
2.8	习题课	27
第3章	信源编码	34
3.1	信源编码问题	34
3.1.1	信源编码	34
3.1.2	定长编码与变长编码	35
3.1.3	信源变长码的编码问题	36
3.1.4	信源序列的定长编码问题	37
3.2	前缀码和即时码	38
3.2.1	唯一可译变长码的构造	38
3.2.2	Kraft 不等式	40
3.3	信源变长码的编码定理	41
3.3.1	最优变长码平均码长的下界估计	41
3.3.2	最优变长码平均码长的上界估计	42
3.4	Huffman 信源编码算法	43
3.4.1	Huffman 编码的实例分析	43
3.4.2	Huffman 编码的一般算法	44
3.5	Huffman 编码性能分析	46
3.5.1	Huffman 编码的前缀性	46
3.5.2	Huffman 编码的最优性	46
3.6	信源定长码的编码定理	49
3.7	习题课	50
3.7.1	基本概念	50
3.7.2	编码问题与基本结论	50
3.7.3	课后习题	51
第4章	信道编码定理	55
4.1	信道编码问题	55
4.1.1	通信系统的编码误差	55
4.1.2	信道序列的编码问题	57
4.2	离散无记忆信道	57
4.2.1	离散无记忆信道的一般定义	57
4.2.2	几种特殊的离散无记忆信道	58
4.3	无记忆信道的信道容量	61
4.3.1	信道容量的一般定义	62
4.3.2	无记忆信道序列的容量性质	65

4.4	信道容量的计算	66
4.4.1	凸函数的极大值性质	66
4.4.2	信道容量的计算	67
4.5	信道的编码和译码问题	69
4.6	信道的正编码定理和反编码定理	71
4.7	可加高斯 (Gaussian) 信道	74
4.8	习题课	76
4.8.1	基本概念	76
4.8.2	基本方法	77
4.8.3	课后习题	77
第 5 章	抽象代数的基本知识	80
5.1	群	80
5.1.1	群的概念	80
5.1.2	子群及判定	80
5.1.3	群中元素的阶	81
5.1.4	循环群	81
5.2	环与域	81
5.2.1	环的概念	81
5.2.2	域	82
5.3	理想和商环	82
5.3.1	理想	82
5.3.2	商环	83
5.3.3	环 (域) 的同构	83
5.4	域上的多项式环	83
5.4.1	域上的多项式环	83
5.4.2	带余除法	84
5.4.3	最大公因式与最小公倍式	84
5.4.4	不可约多项式	84
5.4.5	不可约多项式与有限域的构造	85
5.4.6	重因式及多项式的根	85
5.5	有限域	85
5.5.1	有限域	85
5.5.2	域的特征	86
5.5.3	素域	86
5.5.4	有限域 F_q 的性质	87
5.5.5	极小多项式与本原多项式	88

5.6	域上的线性代数	89
5.6.1	域上的向量空间	89
5.6.2	极大线性无关组	91
5.6.3	域 F 上的 $m \times n$ 矩阵	91
第 6 章	编码理论的基本知识	92
6.1	码的基本概念	92
6.1.1	码的定义	92
6.1.2	Hamming 距离和 Hamming 重量	92
6.1.3	Hamming 距离 (重量) 与译码	93
6.1.4	系统码	94
6.2	码的检错和纠错能力	95
6.2.1	最小距离	95
6.2.2	码的检错和纠错能力	95
6.3	编码理论的基本问题	97
6.3.1	码的等价变换	98
6.3.2	$A_q(n, d)$ 的性质	100
6.3.3	$A_q(n, d)$ 的界	103
6.4	习题课	104
6.4.1	基本概念	104
6.4.2	基本结论	105
6.4.3	课后习题	105
第 7 章	线性码	109
7.1	线性码的定义	109
7.1.1	线性码的概念及性质	109
7.1.2	线性码的表示方法	109
7.2	线性码的对偶码	111
7.2.1	对偶码的定义	111
7.2.2	对偶码的性质	111
7.2.3	线性码的校验矩阵	112
7.2.4	线性码的界	114
7.3	线性码的译码方法	114
7.4	线性码的重量分布	117
7.5	习题课	120
7.5.1	本章小结	120
7.5.2	课后习题	120

第 8 章 Hamming 码	129
8.1 Hamming 码的定义	129
8.1.1 Hamming 码的构造	129
8.2 Hamming 码的性质	129
8.3 Hamming 码的译码方法	131
8.4 二元 Hamming 码的对偶码	133
8.5 习题课	134
8.5.1 基本概念及方法	134
8.5.2 课后习题	135
第 9 章 循环码	138
9.1 循环码的定义	138
9.2 循环码的性质	139
9.3 循环码的校验矩阵及其对偶码	142
9.4 循环码的编码方法	145
9.5 循环码的检错性能	146
9.6 习题课	147
9.6.1 基本概念	147
9.6.2 基本性质与结论	147
9.6.3 课后习题	148

第1章 引言

在本章中，我们介绍信息论与编码理论的基本情况，使大家对信息理论和编码理论的全貌有一个大致的了解。

内容提要

- 信息论的早期酝酿
- Shannon 信息论的建立与发展
- 信息论的近期发展
- 信息的度量问题
- 通信系统的基本要素与模型
- 通信系统的概率统计模型

1.1 信息论的形成与发展

一、信息论的早期酝酿

1. 早期编码问题

在有线和无线电通信产生的同时，编码技术随之产生，早期的编码有 Morse 码和 Bodo 码。他们将文字通过点、划、空等信号给以表达，这些码虽然很原始，但它们实现了从文字到通讯信号的转变。中文通信等用的是电报码的方式，先将汉字转成数字，再用电码发射。

2. 通信的有效性和可靠性

随着通信距离的加大，出现了信号强度衰减与噪声干扰的问题，因此，如何克服噪声干扰问题就成为通信技术中的一个迫切问题。为解决这些问题，人们对通信中的各种因素进行分析，发现在通信技术中，通信的数量与质量存在相互制约的关系。如果牺牲通信的数量，则可以达到提高质量的目的，但二者之间究竟有何定量关系，却无法说明。到了 20 世纪 20 年代，奈奎斯特等人对上述问题进行了一系列讨论，说明了信息传递的速率与带宽成正比，信息的度量与信号的概率分布、对数函数有关等等。

二、Shannon 信息论的建立与发展

信息论的产生以 1948 年 C.E.Shannon 发表的“通信的数学理论”这一奠基性论文为起始，作为信息理论的奠基人，对其主要生平作一个简单的介绍。

Claude Edwoods Shannon(香农、仙农)

C.E.Shannon(1916-2001) 数学家，工程学家。信息论创始人，奠基人，电子计算机理论的重要奠基人之一。代表性著作：

- (1)1938 年 (22 岁)，发表著名论文《继电器和开关电路的符号分析》文中首次使用了比特 (bit) 的概念。
- (2) 1948 年，《通信的数字理论》第一次提出信息量的概念，并且应用数理统计的方法来研究通信系统，从而创立了影响深远的信息论。
- (3)1949 年，《噪声下的通信》经典的阐明了通信的基本问题，提出了通信系统的模型，给出了信息量的数字表达式，解决了通道容量，信源统计特性，信源编码，信道编码等有关精确地传递

通信符号的基本技术问题.

(4)1956 年,《噪声信道的零差错容量》开拓了零差错容量的研究领域.

(5)1959 年,《在保真度准则下的离散信源编码定理》推动了信息率失真理论研究等.

Shannon 提出完善信息理论离不开对编码问题的研究,因此信息论与编码理论密不可分.

1.Shannon 信息论的确立期(1948 年-20 世纪 60 年代)此阶段主要是对 Shannon 理论进行研究与说明,包括对通信系统的数学模型和基本问题的说明与讨论.

2.Shannon 信息论的发展期(20 世纪 70-80 年代)这一时期主要内容在“率失真理论”与“多用户信息论”方面,后发展为数据压缩理论.

三、信息论的近期发展

信息论近期发展的主要特点是向多学科结合方向发展.

1. 信息论与密码学(通信编码问题的一种表现形式)

2. 算法信息论与分形数学(信息论,计算机科学,分形理论的共同本质,决定如何互等价转化).

3. 信息论在统计与智能计算中的应用

(1) 智能计算中的信息统计问题.

(2) 信息计算与组合投资决策关系密切.

(3) 编码理论在与试验设计,假设检验理论的结合中发挥了重要作用.

4. 信息论在工程领域中的应用

主要是编码理论在工程领域有广泛应用.

1.2 信息与编码理论的主要内容

一、信息的度量问题

Shannon 熵是 Shannon 信息论中信息的度量基础,它与概率分布有关,以不肯定性(不确定性)作为度量信息的基础,即信息量是描述消息中不确定性的概念.例如:一个消息“太阳从东边升起,西边落下”,这一消息中没有任何信息,即含有的信息量为 0.这是一确定的消息,不含有任何的不肯定性(不确定性).

二、通信系统的基本要素与模型

1. 一个系统由以下基本要素与模型

(1) 信源:产生信息的来源,信源产生的信息称为消息.

(2) 信道:传递信息的通道,消息通过信道以信号的形式传播.

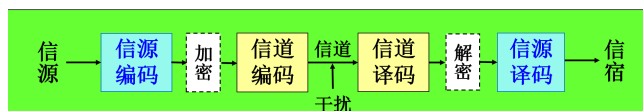
(3) 编码:由消息产生信号的运算.

(4) 译码:由信号还原为消息的运算称为译码.

(5) 通信系统:由信源、信道、编码和译码四个要素组成.

(6) 用户:通信系统的使用者,分为发送者和接收者.

2. 通信系统的基本模型 由于干扰的存在,信道的输出信号可能与输入信号不同,从而使得还原消息与原始消息可能不同,这种现象称为通信误差,是通信系统中需要克服的现象.



克服通信误差 $\left\{ \begin{array}{l} \text{硬件途径: 元器件改进, 降低噪声干扰} \\ \text{软件途径: 用编码方式克服} \rightarrow \text{纠错和检错码} \end{array} \right.$

三、通信系统的概率统计模型

1. 信源的概率统计模型

定义 1.2.1

信源是消息的来源, 用 $\mathcal{S} = [\mathcal{X}, p(x)]$ 来表示, 其中 \mathcal{X} 为信源字母表, 是消息中可能使用的全体符号, 它的元素 $x \in \mathcal{X}$ 是信源字母表中的字母, $p(x)$ 是字母 x 的使用概率. 由概率分布的性质知, 对 $\forall x \in \mathcal{X}$, 有 $p(x) \geq 0$, 且 $\sum_{x \in \mathcal{X}} p(x) = 1$.



2. 信道的概率统计模型

信道由以下因素组成: 输入信号字母集, 输出信号字母集与转移概率分布, 分别记为 $\{\mathcal{U}, \mathcal{V}, p(v|u), u \in \mathcal{U}, v \in \mathcal{V}\}$, 其中 \mathcal{U} 为全体能使用的输入信号字母集, 简称为输入信号字母表, \mathcal{V} 是可能输出的信号字母集, 简称为输出信号字母表. 若 $u \in \mathcal{U}, v \in \mathcal{V}$, 表示输入和输出信号字母, 而 $p(v|u)$ 是输入、输出信号字母的转移概率, 即当输入信号为 u , 输出信号为 v 的概率, 有 $p(v|u) \geq 0, \sum_{v \in \mathcal{V}} p(v|u) = 1$.

定义 1.2.2

信道的概率统计模型为 $\mathcal{C} = (\mathcal{U}, p(v|u), \mathcal{V})$, 其中 \mathcal{U}, \mathcal{V} 分别是输入和输出信号字母表, $p(v|u)$ 是输入和输出信号字母的转移概率.



3. 编码与译码的函数表示

定义 1.2.3

由消息变为信道输入信号的运算称为编码, 由信道输出信号变为还原消息的运算称为译码. 记 \mathcal{Y} 为还原消息字母表. 则 $f: \mathcal{X} \rightarrow \mathcal{U}, g: \mathcal{V} \rightarrow \mathcal{Y}$ 分别是编码和译码运算, 在通信系统中, 一般 \mathcal{U} 与 \mathcal{V} 相同, 但也可以不同.



4. 通信系统的数学模型

定义 1.2.4

称信源与信道的组合为通信系统, 记为

$$\mathcal{E} = \{\mathcal{S}, \mathcal{C}\} = \{\mathcal{X}, p(x), \mathcal{U}, p(v|u), \mathcal{V}\}$$

对于一个通信系统 \mathcal{E} , 如果编码和译码函数 f, g 给定, 那么我们称这个通信系统是有编码

的通信系统，记为

$$\mathcal{E}(f, g) = \{\mathcal{S}, \mathcal{C}, (f, g), \mathcal{Y}\} = \{\mathcal{X}, p(x), \mathcal{U}, p(v | u), \mathcal{V}, (f, g), \mathcal{Y}\}$$



5. 由通信系统决定的随机变量

因为通信系统由一系列概率分布组成，因此我们可以用随机变量表示，记：

$\bar{\xi}$ 由信源 $\mathcal{S} = \{\mathcal{X}, p(x)\}$ 决定的随机变量。

ξ 由信道输入信号决定的随机变量。

η 由信道输出信号决定的随机变量。

$\bar{\eta}$ 由还原消息决定的随机变量。

$\bar{\xi}$ 在 \mathcal{X} 上取值，其概率分布为 $p(x)$ 。

定理 1.2.1

如果有编码的通信系统 $\mathcal{E}(f, g)$ 给定，则这个通信系统的随机变量 $(\bar{\xi}, \xi, \eta, \bar{\eta})$ 的联合概率分布确定，它们的联合概率分布 $p(x, u, v, y) = p(x)f(u | x)p(v | u)g(y | v)$ ，其中 $p(x), p(v | u)$ 分别由信源 S 和信道 C 给定。

$$f(u | x) = \begin{cases} 1, & u = f(x) \\ 0, & u \neq f(x) \end{cases}$$

$$g(y | v) = \begin{cases} 1, & y = g(v) \\ 0, & y \neq g(v) \end{cases}$$

这时， $(\bar{\xi}, \xi, \eta, \bar{\eta})$ 构成一个马尔可夫 (Markov) 链，因为当 ξ 固定时， $\bar{\xi}$ 与相互独立，而当 η 固定时， $(\bar{\xi}, \xi)$ 与 $\bar{\eta}$ 相互独立。



有关概率记号： $p(x, u, v, y) = P_r\{(\bar{\xi}, \xi, \eta, \bar{\eta}) = (x, u, v, y)\}$

$$p(x) = \Pr\{\bar{\xi} = x\}$$

$$p(v | u) = \Pr\{\eta = v | \xi = u\}$$

其中， $P_r(A)$ 表示事件 A 发生的概率， $P_r(A | B)$ 表示事件 A 关于事件 B 的条件概率。

$$P(A) = P_r(A) \quad P(A | B) = P_r(A | B)$$

定义 1.2.5

如果随机变量 $(\bar{\xi}, \xi, \eta, \bar{\eta})$ 的概率分布为 $p(x, u, v, y) = p(x)f(u | x)p(v | u)g(y | v)$ ，则称 $(\bar{\xi}, \xi, \eta, \bar{\eta})$ 是由通信系统 $\mathcal{E}(f, g)$ 决定的随机变量，如果 $(\bar{\xi}, \xi, \eta, \bar{\eta})$ 构成一个马尔可夫链，那么记之为：

$$\bar{\xi} \longrightarrow \xi \longrightarrow \eta \longrightarrow \bar{\eta}$$



1.3 本书基本内容与参考文献

一、基本内容

1. 信息论部分：
 - (a). 信息的度量和性质
 - (b). 信源和信道的编码问题及编码定理
2. 编码理论部分
 - (a). 代数码的概念和性质
 - (b). 几种典型的代数码 (Hamming 码, 循环码)

二、前续知识:

- 概率论
- 抽象代数 (近世代数)

三、参考文献

- Coding and Information Theory.Steven Roman.Springer.
- 万哲先. 代数和编码 (第三版). 北京. 高等教育出版社.
- 叶中行. 信息论基础. 高等教育出版社.
- 曹雪虹, 张宗橙. 信息论与编码, 北京邮电大学出版社.

第2章 信息量

2.1 熵

设信源 $\mathcal{S} = \{\mathcal{X}, p(x)\}$ 为离散信息, $\bar{\xi}$ 为一随机变量, 取自 \mathcal{S} . 如: 随机事件掷骰子. 信源集 $\mathcal{X} = \{1, 2, 3, 4, 5, 6\}$ 的概率分布为 $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$, $\bar{\xi}$ 取某一可能, 如 $\bar{\xi} = 4$, 则 $\bar{\xi}$ 带有多少信息量与 4 的概率有关, 即与不确定性有关. 此时, $\bar{\xi} = 4$ 提供了多少信息.

一般地 $\mathcal{X} = \{x_1, x_2, \dots, x_a\}$, 其中 a 为某个正整数, 对 $\forall x_i \in \mathcal{X}$, $\bar{\xi}$ 取值为 x_i 的概率为

$$p_i = p(\xi_i) = P_r(\xi = x_i), i = 1, 2, \dots, a.$$

$\bar{\xi}$ 和它的概率分布表示为 $\begin{pmatrix} x_1 & x_2 & \cdots & x_a \\ p_1 & p_2 & \cdots & p_a \end{pmatrix}$, 则有 $p_i \geq 0, i = 1, 2, \dots, a, \sum_{i=1}^a p_i = 1$.

$\bar{\xi}$ 取某一信号 $x_i, \bar{\xi} = x_i$, 带有多少信息量, 与 x_i 的概率有关, 即与不确定性有关.

2.1.1 自信息量

定义 2.1.1

一个随机事件发生某一结果后所带来的信息量称为自信息量. 定义为其发生概率的对数的负值. 即若随机事件 x_i 发生的概率为 $p(x_i)$, 那么它的自信息量为


$$I(x_i) = -\log p(x_i)$$



由于 x_i 是随机出现的, 它是 \mathcal{X} 的一个样值, 所以是一个随机量. 而 $I(x_i)$ 是 x_i 的函数, 它必须也是一个随机量.

自信息量的单位与所用的对数底有关. 在信息论中常用的对数底是 2, 信息量的单位为比特 (bit); 若取自然对数, 则信息量的单位为奈特 (nat); 若以 10 为对数底, 则信息量的单位为笛特 (det). 这三个信息量单位之间的转换关系如下:

$$1 \text{ nat} = \log_2 e \approx 1.433 \text{ bit}, \quad 1 \text{ det} = \log_2 10 \approx 3.322 \text{ bit}$$

 **笔记** 信息量是纯数, 信息量单位只是为了标示不同底数的对数值, 并无量纲的含义.

自信息量的特点是, 当事件的概率越小, 其自信息量越大, 表示提供的信息越多; 而当事件的概率越大, 其自信息量越小, 表示提供的信息越少. 当事件的概率越小, 其自信息量越大的原因可以通过信息的意外性来解释. 假设一个事件的概率非常小, 意味着这个事件发生的可能性非常低, 它是一种罕见或异常的情况. 当这个罕见事件发生时, 它提供了大量的信息, 因为它与我们的预期或常见情况相悖, 具有较高的意外性.

这里要引入随机事件的不确定度概念. 根据日常知识, 各个出现概率不同的随机事件所包含的不确定度是有差别的. 一个出现概率接近于 1 的随机事件, 发生的可能性很大, 所以它包

含的不确定度就很小. 反之, 一个出现概率很小的随机事件, 很难猜测在某个时刻它能否发生, 所以它包含的不确定度就很大. 若是确定性事件, 出现概率为 1, 则它包含的不确定度为 0.

注: 随机事件的不确定度在数量上等于它的自信息量, 两者的单位相同, 但含义却不相同. 具有某种概率分布的随机事件不管发生与否, 都存在不确定度, 不确定度表征了该事件的特性, 而自信息量是在该事件发生后给予观察者的信息量.

2.1.2 不确定性(自信息量)的性质


上述自信息函数 $I(x_i)$ 应满足以下性质:

- (1) $p(x_i)$ 越大, $I(x_i)$ 越小, 即 $I(x)$ 是关于 $P(x)$ 单调递减的;
- (2) x, y 是相互独立的两个信号, 即联合分布 $p(x, y) = p(x)p(y)$ 时, $I(x, y) = I(x) + I(y)$. 于是, 可推出自信息函数 $I(x)$ 应满足以下公理:

公理 2.1

- (i) $I(x) \geq 0$;
- (ii) 当 $p(x) = 1$ 时, $I(x) = 0$;
- (iii) 当 $p(x) = 0$ 时, $I(x) = \infty$;
- (iv) 若 $p(x) > p(y)$, 则 $I(x) < I(y)$;
- (v) 若 $p(x, y) = p(x)p(y)$, 则 $I(x, y) = I(x) + I(y)$.



 **笔记** (1) 信息量非负说明随机事件发生后总能提供一些信息量, 最差情况是零, 即什么信息也没提供, 但不会因事件发生使得不确定性更大. P_i 表示随机事件发生的概率, 在闭区间 $[0, 1]$ 上取值, 根据对数的性质也可知 $I(x) \geq 0$.

(2) 当 $p(x) = 1$ 时, 说明该事件是必然事件. 必然事件不含有任何不确定性, 所以不含任何信息量.

(3) 当一个事件的概率接近于 1 时, 它的信息量趋近于 0, 表示这个事件是高度可预测的, 提供的信息量很少. 相反, 当一个事件的概率接近于 0 时, 它的信息量趋近于正无穷大, 表示这个事件是高度意外的, 提供的信息量很大.

(4) 概率越大的事件, 不确定性越小, 发生后提供的信息量就越小. 即 $I(x)$ 是 p_i 的单调递减函数.

(5) 首先, 根据独立事件的定义, 事件 x 和 y 的联合概率等于它们的边缘概率的乘积, 即 $p(x, y) = p(x) \cdot p(y)$. 然后, 根据自信息的定义, $I(x) = -\log p(x)$ 和 $I(y) = -\log p(y)$. 将上述两个式子代入 $I(x, y)$ 的定义中, 有: $I(x, y) = -\log p(x, y) = -\log(p(x) \cdot p(y))$. 根据对数运算的性质, 上式可以改写为: $I(x, y) = -\log p(x) - \log p(y) = I(x) + I(y)$. 因此, 当事件 x 和 y 是独立事件且满足 $p(x, y) = p(x) \cdot p(y)$ 时, 可以得出 $I(x, y) = I(x) + I(y)$ 的结论. 这意味着两个独立事件的联合自信息等于它们各自的自信息之和.

于是, 要对信息量进行度量(定量表示), 需要寻求一个满足上述条件的自信息函数.

2.1.3 自信息函数的推导

引理 2.1.1

若实函数 $f(x)$ ($1 \leq x \leq \infty$) 满足以下条件:

(i) $f(x) \geq 0$;

(ii) $f(x)$ 是严格单调增加的, 即 $x < y \Rightarrow f(x) < f(y)$;

(iii) $f(x \cdot y) = f(x) + f(y)$;

则有 $f(x) = c \log x$, 其中 c 为常数.



证明 反复使用 (iii), 对任意正整数 k , 我们有

$$f(x^k) = f(x \cdot x^{k-1}) = f(x) + f(x^{k-1}) = \cdots = kf(x)$$

从而 $f(1) = 0$. 进而由于 (i) 和 (ii), 对于任意 $x > 1$, $f(x) > 0$, 对于任意大于 1 的 x, y 与任意正整数 k , 总可以找到非负整数 n , 使

$$y^n \leq x^k < y^{n+1}$$

(事实上, 由 $y > 1$, 则区间 $(1, +\infty)$ 可由 y 分为 $(1, y], (y, y^2], [y^2, y^3], \cdots$, x 取定, x^k 是固定的正函数, 必落在某个区间内.)

取对数并除以 $k \log_a y$ 得

$$\frac{n}{k} \leq \frac{\log_a x}{\log_a y} < \frac{n+1}{k} \quad (2.1.1)$$

另一方面, 由 $f(x^k) = kf(x)$ 结合条件 (ii) 可得

$$nf(y) \leq kf(x) < (n+1)f(x)$$

两边同时除以 $kf(y)$ 得:

$$\frac{n}{k} \leq \frac{f(x)}{f(y)} < \frac{n+1}{k}$$

两边同时乘以 -1 得:

$$-\frac{n+1}{k} \leq -\frac{f(x)}{f(y)} \leq -\frac{n}{k} \quad (2.1.2)$$

于是联立(2.1.1)和(2.1.2)得:

$$-\frac{1}{k} \leq \frac{\log_c x}{\log_c y} - \frac{f(x)}{f(y)} \leq \frac{1}{k}$$

我们有

$$\left| \frac{f(x)}{f(y)} - \frac{\log_a x}{\log_a y} \right| \leq \frac{1}{k}$$

当 $k \rightarrow \infty$ 时,

$$\frac{f(x)}{f(y)} = \frac{\log_a x}{\log_a y}$$

因此,

$$\frac{f(x)}{\log_a x} = \frac{f(y)}{\log_a y} = c$$

或

$$f(x) = c \log_a x$$

注: 对于自信息函数 $I(x) = I(p(x))$, 令 $t = \frac{1}{p(x)}$, 则 $I(t) = I\left(\frac{1}{p(x)}\right)$, $I(x)$ 关于 $p(x)$ 递减, 又 $I(t(x), t(y)) = I(t(x)) + I(t(y))$ 成立, 由引理知: $I(t) = c \log_a t$, 从而 $I(x) = c \log_a \frac{1}{p(x)}$.

定理 2.1.1

若自信息 $I(x)$ 满足 5 条公理, 则 $I(x) = c \log_a \frac{1}{p(x)}$, 其中 c 为常数, 此时称 $I(x)$ 为自信息函.



设 $x \in \mathcal{X}$, x 具有概率分布 $p(x)$, 则 x 的自信息定义为

$$I(x) = \log_a \frac{1}{p(x)}$$

2.1.4 香农熵的定义 (Shannon 熵)

信源 $\mathcal{S} = \{\mathcal{X}, p(x)\}$, ξ 是随机变量 (考虑随机信号的信息量).

前面介绍 $x \in \mathcal{X}$ 的自信息是在 X 确定的情况下, 现在 ξ 是随机变量, 发送哪一个信号是不确定的, 因此需考虑在信号随机选取时携带的信息量如何刻画, 即对整个信源, 每个信号的平均信息量为多少?

定义 2.1.2

如果一个离散随机变量 ξ 的概率分布为 $\bar{p} = (p_1, p_2, \dots, p_a)$, 则它的熵 $H(\xi)$ 或 $H(\bar{p})$ 定义为

$$\begin{aligned} H(\xi) &= H(p_1, p_2, \dots, p_a) = - \sum_{i=1}^a p_i \log_c p_i \\ &= \sum_{i=1}^a p_i \log_c \frac{1}{p_i} = E \left(\log_c \frac{1}{p} \right) \end{aligned}$$



注:

(1) $H(\xi)$ 是 $\log_c \frac{1}{p(\xi)}$ 的期望值

(2) 底数 $c = 2$ 时, $H(\xi)$ 的单位规定为比特;

底数 $c = e$ 时, $H(\xi)$ 的单位规定为奈特;

底数 $c = 3$ 时, $H(\xi)$ 的单位规定为铁特;

各个单位之间可根据对数的换底公式进行换算; 如: $\log_2 \frac{1}{p_i} = \frac{\log_e \frac{1}{p_i}}{\log_e 2}$

(3) 除特别说明外, 取 $c = 2$.

2.1.5 熵的简单性质与例子

$H(p_1, p_2, \dots, p_a)$ 为上述定义的 Shannon 熵, 则其满足以下性质:

(1) 对称性: $H(p_1, p_2, \dots, p_a) = H(p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(a)})$ 其中 σ 为有限集 $\{1, 2, \dots, a\}$ 的一个置换.

(2) 非负性: 对任意的概率分布 $\bar{p} = \{p_1, p_2, \dots, p_a\}$, 有 $H(p_1, p_2, \dots, p_a) \geq 0$, 且等号成立当且仅当 \bar{p} 是一个确定性分布 (其中一个取 1, 其它的取 0).

证明: (1) 显然当概率的顺序发生置换后, 只是求和顺序不同, 并不影响求和结果. (平均值, 数学期望).

(2) \Leftarrow : 显然

\Rightarrow : $H(p_1, p_2, \dots, p_a) = 0$, 由 $H(p_1, p_2, \dots, p_a)$ 的

定义可知, $\forall i, p_i \log_c p_i = 0$, 或者 $p_i = 0$, 或者 $\log_c p_i = 0$, 由于 $\sum_{i=1}^a p_i = 1, p_i \geq 0$, 存在 i 使得 $p_i = 1$, 而其它 $p_j = 0$, 因此 \bar{p} 必为确定型分布.

例题 2.1.1 以等概率 $p_i = \frac{1}{3}$ 从集合 $\mathcal{X} = \{x_1, x_2, x_3\}$ 中抽样, 以 $p_1 = p_2 = \frac{1}{4}, p_3 = \frac{1}{2}$ 从集合 $\mathcal{X} = \{x_1, x_2, x_3\}$ 中抽样, 以上两个随机事件那个信息量大? (不确定性大, 信息量越大)

解: 设两个随机事件的随机变量分别为 ξ_1, ξ_2 , 则有

$$\begin{aligned} H(\xi_1) &= \sum_{i=1}^3 p_i \log_2 \frac{1}{p_i} \\ &= \left(\frac{1}{3} \log_2 3 + \frac{1}{3} \log_2 3 + \frac{1}{3} \log_2 3 \right) \\ &= \log_2 3 \approx 1.585 \\ H(\xi_2) &= \sum_{i=1}^3 p_i \log_2 \frac{1}{p_i} \\ &= \left(\frac{1}{4} \log_2 4 + \frac{1}{4} \log_2 4 + \frac{1}{2} \log_2 2 \right) \\ &= \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = 1.5 \end{aligned}$$

第一个随机事件信息量大.

例题 2.1.2 设 ξ 是一个二元随机变量, 即 $\mathcal{X} = \{0, 1\}$, 令

$$\begin{aligned} p(\xi = 1) &= p, \text{ 则 } p(\xi = 0) = 1 - p. \text{ 则有} \\ H(\xi) &= p \log_2 \frac{1}{p} + (1 - p) \log_2 \frac{1}{1-p} \triangleq H(p) \end{aligned}$$

$H(p)$ 被称为熵函数.

例题 2.1.3 信源集 $\mathcal{X} = \{x_1, x_2, \dots, x_a\}$, 以等概率从 \mathcal{X} 中抽样, 即 ξ 服从分布 $\begin{pmatrix} x_1 & x_2 & \cdots & x_a \\ p_1 & p_2 & \cdots & p_a \end{pmatrix}$, 则该随机事件的信息量为多少?

解:

$$H(\xi) = \sum_{i=1}^a p_i \log_2 \frac{1}{p_i} = \sum_{i=1}^a \frac{1}{a} \log_2 a = \log_2 a$$

可知 a 越大, $H(\xi)$ 越大 (信号越多, 不确定性越大).

例题 2.1.4 令黑白电视机的分辨率为 500×600 , 且灰度为 10, 令文章中的字可以从一万个字中

任意挑选, 分别求出一副电视画面和一篇千字文章中所含的信息量并比较大小.

解: 设两个随机事件的随机变量分别为 ξ_1, ξ_2 . 对于随机事件一, 一幅电视画面, 分辨率为 500×600 , 则像素数为 $500 \times 600 = 300000 = 3 \times 10^5 = N_1$

对每个像素我们计算它携带的信息量, 由于灰度为 10 (每个灰度值均一样), 则

$$\xi_1 \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_{10} \\ \frac{1}{10} & \frac{1}{10} & \cdots & \frac{1}{10} \end{pmatrix}$$

$$H(\xi_1) = \sum_{i=1}^{10} \frac{1}{10} \log_2 10 = \log_2 10 \approx 3.32 \text{ 比特/像素}$$

则一副电视画面携带的信息量为

$$N_1 H(\xi_1) = 3 \times 10^5 \times 3.32 = 9.96 \times 10^5 \text{ 比特.}$$

对于随机事件二, 一篇千字文章, 计算每个文字所携带的信息量 $N_2 = 1000$,

$$\xi_2 \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_{10^4} \\ \frac{1}{10^4} & \frac{1}{10^4} & \cdots & \frac{1}{10^4} \end{pmatrix}$$

$$H(\xi_2) = \sum_{i=1}^{10^4} \frac{1}{10^4} \log_2 10^4 = 4 \log_2 10 = 13.29 \text{ 比特/字}$$

则一篇千字文章所携带的信息量为 $N_2 H(\xi_2) = 1000 \times 13.29 = 1.329 \times 10^4$ 比特. 故一副电视画面携带的信息量大.

2.2 联合熵与条件熵

2.2.1 联合熵

定义 2.2.1

设一维随机变量 (ξ, η) 的联合分布为

$$p(x, y) = P_r\{\xi = x, \eta = y\}, x \in \mathcal{X}, y \in \mathcal{Y}$$

对于 $x \in \mathcal{X}, y \in \mathcal{Y}$, 二维随机变量 ξ, η 的联合熵定义为:

$$H(\xi, \eta) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log_2 p(x, y)$$

或者写成数学期望的形式:

$$H(\xi, \eta) = -E(\log_2 p(x, y))$$



注: 若 ξ, η 是相互独立的, 即 $p(x, y) = p(x)p(y)$, 则二维随机变量 ξ, η 的联合熵满足 $H(\xi, \eta) = H(\xi) + H(\eta)$

证明

$$\begin{aligned}
H(\xi, \eta) &= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x, y) \\
&= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) [\log p(x) + \log p(y)] \\
&= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x) - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(y) \\
&= - \sum_{x \in \mathcal{X}} \left(\sum_{y \in \mathcal{Y}} p(x, y) \right) \log p(x) - \sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} p(x, y) \right) \log p(y) \\
&= - \sum_{x \in \mathcal{X}} p(x) \log p(x) - \sum_{y \in \mathcal{Y}} p(y) \log p(y) \\
&= H(\xi) + H(\eta)
\end{aligned}$$

注: $-\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x)$, x 取定, $y \in \mathcal{Y}$, 求和与 x 无关, 故

$$-\sum_{x \in \mathcal{X}} \left(\sum_{y \in \mathcal{Y}} p(x, y) \right) \log p(x) = - \sum_{x \in \mathcal{X}} p(x) \log p(x)$$

2.2.2 条件熵

设随机变量对 (ξ, η) 有联合分布 $p(x, y)$, 用

$$p(y | x) = P_r\{\eta = y | \xi = x\}, x \in \mathcal{X}, y \in \mathcal{Y}$$

表示条件概率分布, 则给定 $\xi = x$ 条件下 η 的熵定义为

$$H(\eta | \xi = x) = - \sum_{y \in \mathcal{Y}} p(y | x) \log p(y | x)$$

而给定随机变量 ξ 条件下 η 的熵记为 $H(\eta | \xi)$, 它是 $H(\eta | \xi = x)$ 关于 ξ 的平均值, 即

定义 2.2.2

如果 $(\xi, \eta) \sim p(x, y)$, 那么 η 关于 ξ 的条件熵定义为:

$$\begin{aligned}
H(\eta | \xi) &= \sum_{x \in \mathcal{X}} p(x) H(\eta | \xi = x) \\
&= - \sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y | x) \log p(y | x) \\
&= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(y | x) \\
&= -E(\log p(\eta | \xi))
\end{aligned}$$

同样地, 有 $H(\xi | \eta) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x | y)$



2.2.3 联合熵和条件熵的关系

定理 2.2.1

联合熵与条件熵的关系为:

$$H(\xi, \eta) = H(\xi) + H(\eta | \xi) = H(\eta) + H(\xi | \eta) = H(\eta, \xi)$$



证明

$$\begin{aligned}
 H(\xi, \eta) &= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x, y) \\
 &= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x) p(y | x) \\
 &= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) [\log p(x) + \log p(y | x)] \\
 &= - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x) - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(y | x) \\
 &= - \sum_{x \in \mathcal{X}} \left(\sum_{y \in \mathcal{Y}} p(x, y) \right) \log p(x) - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(y | x) \\
 &= - \sum_{x \in \mathcal{X}} p(x) \log p(x) - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(y | x) \\
 &= H(\xi) + H(\eta | \xi)
 \end{aligned}$$

同理可证明 $H(\xi, \eta) = H(\eta) + H(\xi | \eta)$. 根据此关系式可由联合熵和随机变量的熵求出条件熵.

2.2.4 联合熵和条件熵的性质

性质 1: $H(\eta | \xi) \geq 0$, 等号成立的充要条件为 η 是由 ξ 决定的随机变量.

注: 若 $p(y | x) = 1$, 则在 ξ 的条件下, η 一定发生.

性质 2: $H(\xi, \eta) \geq H(\xi)$ (或者 $H(\xi, \eta) \geq H(\eta)$), 等号成立的充要条件为 η 是由 ξ 决定的随机变量 (或者 ξ 是由 η 决定的随机变量).

性质 3: $H(\xi | \eta) = H(\xi)$ 的充要条件为 ξ, η 是相互独立的.

$$\begin{aligned}
 p(x, y) &= p(x)p(y), p(x | y) = \frac{p(x)p(y)}{p(y)} = p(x) \\
 (p(x | y) &= p(x), \quad y \text{ 不影响 } x)
 \end{aligned}$$

例题 2.2.1 令 (ξ, η) 具有如下联合分布

ξ, η 的边际分布分别为:

$$\xi \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{pmatrix} \quad \eta \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

对 $\xi: p(x) = \sum_{y \in \mathcal{Y}} p(x, y)$ (x 定).

$$\text{故 } p(1) = p(1, 1) + p(1, 2) + p(1, 3) + p(1, 4) = \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + \frac{1}{4} = \frac{1}{2},$$

$\eta \backslash \xi$	1	2	3	4	Σ
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{4}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{4}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
4	$\frac{1}{4}$	0	0	0	$\frac{1}{4}$
Σ	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	1

$$p(2) = p(2, 1) + p(2, 2) + p(2, 3) + p(2, 4) = \frac{1}{16} + \frac{1}{8} + \frac{1}{16} + 0 = \frac{1}{4}$$

于是, 可求得

$$H(\xi) = \sum_{i=1}^4 p_i \log \frac{1}{p_i} = \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + 2 \cdot \frac{1}{8} \log_2 8 = \frac{1}{2} + \frac{2}{4} + \frac{3}{4} = \frac{7}{4}$$

$$H(\eta) = \sum_{i=1}^4 p_i \log \frac{1}{p_i} = 4 \cdot \frac{1}{4} \log_2 4 = 4 \cdot \frac{1}{2} = 2$$

$$\begin{aligned} H(\xi, \eta) &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{1}{p(x, y)} \\ &= 2 \cdot \frac{1}{8} \log 8 + 6 \cdot \frac{1}{16} \log 16 + 4 \cdot \frac{1}{32} \log 32 + \frac{1}{4} \log 4 \\ &= \frac{3}{4} + \frac{6}{4} + \frac{5}{8} + \frac{1}{2} = \frac{27}{8} \end{aligned}$$

$$H(\xi | \eta) = H(\xi, \eta) - H(\eta) = \frac{27}{8} - 2 = \frac{11}{8}$$

$$H(\eta | \xi) = H(\xi, \eta) - H(\xi) = \frac{27}{8} - \frac{7}{4} = \frac{13}{8}$$

2.3 熵的基本性质

2.3.1 对数函数的基本不等式

引理 2.3.1

对任意实数 $x > 0$, 有 $1 - \frac{1}{x} \leq \ln x \leq x - 1$, 等号成立的条件为 $x = 1$.



引理 2.3.2

对任意两组实数 $p_i, q_i, i = 1, 2, \dots, a$, 若满足:

(1) $p_i \geq 0, q_i \geq 0, i = 1, 2, \dots, a$;

(2) $\sum_{i=1}^a p_i = 1, \sum_{i=1}^a q_i = 1$, 则有:

$$-\sum_{i=1}^a p_i \log p_i \leq -\sum_{i=1}^a p_i \log q_i \text{ 或 } \sum_{i=1}^a p_i \log \frac{p_i}{q_i} \geq 0$$

其中等号成立的条件 (充要) 为 $\frac{p_i}{q_i} = 1, i = 1, 2, \dots, a$.



证明

$$\ln \frac{q_i}{p_i} = \frac{\log \frac{q_i}{p_i}}{\log e} \Rightarrow \log \frac{q_i}{p_i} = (\log e) \left(\ln \frac{q_i}{p_i} \right)$$

由引理2.3.1可知: $\log \frac{q_i}{p_i} \leq (\log e) \left(\frac{q_i}{p_i} - 1 \right)$ 两边同时乘以 p_i 得:

$$p_i \log \frac{q_i}{p_i} \leq (\log e) p_i \left(\frac{q_i}{p_i} - 1 \right) = \log e (q_i - p_i)$$

两边求和得:

$$\sum_{i=1}^a p_i \log \frac{q_i}{p_i} \leq \log e \sum_{i=1}^a (q_i - p_i) = 0$$

$$\left(\sum_{i=1}^a q_i = 1, \sum_{i=1}^a p_i = 1 \right)$$

于是有: $\sum_{i=1}^a p_i \log \frac{q_i}{p_i} \leq 0 \Rightarrow \sum_{i=1}^a p_i \log \frac{p_i}{q_i} \geq 0$ 由引理2.3.1知等号成立的条件为 $x = 1$. 故上述不等式成立的条件为:

$$\frac{p_i}{q_i} = 1 \Rightarrow p_i = q_i, \quad i = 1, 2, \dots, a.$$

引理 2.3.3

对任意两组实数 $u_i > 0, v_i > 0, i = 1, 2, \dots, a$, 有:

$$\sum_{i=1}^a u_i \log \frac{u_i}{v_i} \geq \left(\sum_{i=1}^a u_i \right) \log \frac{\sum_{i=1}^a u_i}{\sum_{i=1}^a v_i}$$

其中等号成立的充要条件为: $\frac{u_k}{v_k} = \frac{\sum_{i=1}^a u_i}{\sum_{i=1}^a v_i}, i = 1, 2, \dots, a$.



证明 令 $p_k = \frac{u_k}{\sum_{i=1}^a u_i}, q_k = \frac{v_k}{\sum_{i=1}^a v_i}$;

$$\sum_{k=1}^a p_k = \sum_{k=1}^a \frac{u_k}{\sum_{i=1}^a u_i} = \frac{\sum_{k=1}^a u_k}{\sum_{i=1}^a u_i} = 1;$$

$$\sum_{k=1}^a q_k = 1, p_k \geq 0, q_k \geq 0$$

由引理2.3.2有: $\sum_{k=1}^a p_k \log \frac{p_k}{q_k} \geq 0$, 即:

$$\sum_{k=1}^a \left(\frac{u_k}{\sum_{i=1}^a u_i} \log \frac{\frac{u_k}{\sum_{i=1}^a u_i}}{\frac{v_k}{\sum_{i=1}^a v_i}} \right) \geq 0 \Rightarrow \sum_{k=1}^a \frac{u_k}{\sum_{i=1}^a u_i} \log \frac{u_k}{v_k} \cdot \frac{\sum_{i=1}^a v_i}{\sum_{i=1}^a u_i} \geq 0$$

$$\begin{aligned} \text{由 } \sum_{i=1}^a u_i \geq 0 \text{ 得: } \sum_{k=1}^a u_k \log \frac{u_k}{v_k} \cdot \frac{\sum_{i=1}^a v_i}{\sum_{i=1}^a u_i} &\geq 0 \\ \Rightarrow \sum_{k=1}^a u_k \left[\log \frac{u_k}{v_k} - \log \frac{\sum_{i=1}^a u_i}{\sum_{i=1}^a v_i} \right] &\geq 0 \end{aligned}$$

$$\text{即: } \sum_{k=1}^a u_k \log \frac{u_k}{v_k} \geq \sum_{k=1}^a u_k \log \frac{\sum_{i=1}^a u_i}{\sum_{i=1}^a v_i}$$

$$\text{亦即: } \sum_{k=1}^a u_k \log \frac{u_k}{v_k} \geq \left(\sum_{k=1}^a u_k \right) \log \frac{\sum_{k=1}^a u_k}{\sum_{k=1}^a v_k}. \text{ 等号成立的条件为 } p_k = q_k, \text{ 即:}$$

$$\frac{u_k}{\sum_{i=1}^a u_i} = \frac{v_k}{\sum_{i=1}^a v_i} \Rightarrow \frac{u_k}{v_k} = \frac{\sum_{i=1}^a u_i}{\sum_{i=1}^a v_i}$$

2.3.2 熵函数的性质

定理 2.3.1 (熵函数的最大值)

令 ξ 是一个离散随机变量, 它在 $\mathcal{X} = \{x_1, x_2, \dots, x_a\}$ 中取值, 那么有: $H(\xi) \leq \log a$. 其中等号成立的充要条件为对所有的 i 都有 $p(x_i) = \frac{1}{a}, i = 1, \dots, a$



证明 由引理 2.3.3, 取 $(u_1, \dots, u_a) = (p_1, \dots, p_a)$ 为 ξ 的概率分布, 取 $v_i = 1, i = 1, 2, \dots, a$ 可知:

$$\begin{aligned} H(\xi) &= - \sum_{i=1}^a p_i \log p_i \quad (u_i = p_i, v_i = 1) \\ &= - \sum_{i=1}^a u_i \log \frac{u_i}{v_i} \\ &\leq - \left(\sum_{i=1}^a u_i \right) \log \frac{\sum_{i=1}^a u_i}{\sum_{i=1}^a v_i} \\ &= - \left(\sum_{i=1}^a p_i \right) \log \frac{\sum_{i=1}^a p_i}{a} \quad \left(\sum_{i=1}^a p_i = 1 \right) \\ &= \log a \end{aligned}$$

定理 2.3.2 (熵函数的可加性)

如果 $q_{ij}, j = 1, 2, \dots, k_i, i = 1, 2, \dots, a$ 是一组非负数, 满足:

$$q_{ij} \geq 0, p_i = \sum_{j=1}^{k_i} q_{ij}, \sum_{i=1}^a p_i = 1$$

对任何 $j = 1, 2, \dots, k_i, i = 1, 2, \dots, a$ 成立, 那么有:

$$\begin{aligned} & H(q_{11}, q_{12}, \dots, q_{1k_1}, q_{21}, q_{22}, \dots, q_{2k_2}, \dots, q_{a1}, q_{a2}, \dots, q_{ak_a}) \\ &= H(p_1, p_2, \dots, p_a) + \sum_{i=1}^a p_i H\left(\frac{q_{i1}}{p_i}, \dots, \frac{q_{ik_i}}{p_i}\right). \end{aligned}$$



证明 由熵的定义可知:

$$\begin{aligned} & H(q_{11}, q_{12}, \dots, q_{1k_1}, q_{21}, q_{22}, \dots, q_{2k_2}, \dots, q_{a1}, q_{a2}, \dots, q_{ak_a}) \\ &= - \sum_{i=1}^a \sum_{j=1}^{k_i} q_{ij} \log q_{ij} = - \sum_{i=1}^a \sum_{j=1}^{k_i} q_{ij} \log \frac{q_{ij}}{p_i} \cdot p_i \\ &= - \sum_{i=1}^a \sum_{j=1}^{k_i} q_{ij} \left(\log \left(\frac{q_{ij}}{p_i} \right) + \log p_i \right) \\ &= - \sum_{i=1}^a \sum_{j=1}^{k_i} q_{ij} \log p_i - \sum_{i=1}^a \sum_{j=1}^{k_i} q_{ij} \log \frac{q_{ij}}{p_i} \\ &= - \sum_{i=1}^a \log p_i \sum_{j=1}^{k_i} q_{ij} - \sum_{i=1}^a p_i \sum_{j=1}^{k_i} \frac{q_{ij}}{p_i} \log \frac{q_{ij}}{p_i} \\ &= - \sum_{i=1}^a p_i \log p_i - \sum_{i=1}^a p_i \sum_{j=1}^{k_i} \frac{q_{ij}}{p_i} \log \frac{q_{ij}}{p_i} \\ &= H(p_1, p_2, \dots, p_a) + \sum_{i=1}^a p_i H\left(\frac{q_{i1}}{p_i}, \dots, \frac{q_{ik_i}}{p_i}\right) \end{aligned}$$

注: 若集合 \mathcal{X} 被划分为 a 个子集 $\mathcal{S}_i (i = 1, 2, \dots, a)$, 每个子集的概率为 $p_i (i = 1, 2, \dots, a)$. 其熵为 $H(p_1, p_2, \dots, p_a)$, 对于这 a 个子集, 我们把每一个子集 \mathcal{S}_i 又分成 k_i 个小子集 ($i = 1, 2, \dots, a$), 每个小子集的概率为 q_{ij} . 即有 $\sum_{j=1}^{k_i} q_{ij} = p_i$

判断事件具体属于哪个子集的不确定性 (在哪个子集中选取的不确定性), 等于大子集的不确定性 $H(p_1, \dots, p_a)$ 与小子集不确定性的概率加权统计平均值之和.

例题 2.3.1 一个班的同学的集合设为 X , 把 X 分为 4 小组, 每个小组有若干排, 某一个同学在哪个座位上的不确定性为多少? 我们可以分两种不同方法考虑这个问题:

一种方法, 从第一组开始, 沿着每一个座位查找, 直到找到有他名字的座位为止;

第二种方法, 我先考虑它可能在哪个小组中, 然后再在这个小组中, 考虑他在哪一排哪个座位上, 去找他的座位.

例题 2.3.2 电脑中文件的查找 $\left\{ \begin{array}{l} \text{所有文件均列出来一一查找} \\ \text{(提供的不确定性相同, 通常采用这种方法)} \\ \text{找到文件所在的文件夹再查找} \end{array} \right.$

定理 2.3.3

如果 f 是从 \mathcal{X} 到 \mathcal{Z} 的任意映射, 那么必有 $H(\xi) \geq H(f(\xi))$ 成立, 等号成立的充要条件为 f 是一个 1-1 变换, 也就是对任何 $x, x' \in \mathcal{X}, x \neq x', p(x) \neq 0, p(x') \neq 0$, 必有 $f(x) \neq f(x')$, 即单射.



证明 记 $\mathcal{Z}_0 = \{f(x) \mid x \in \mathcal{X}\}$ (f 的像集), 令 $\mathcal{Z}_0 = \{z_1, z_2, \dots, z_b\}$ 记 $A_i = \{x \mid f(x) = z_i\}$ ($i = 1, 2, \dots, b$), 则 A 是 z_i 的原像的集合. 记 A_i 中的元为 $A_i = \{x_{i1}, x_{i2}, \dots, x_{ik_i}\} \subseteq \mathcal{X}$, (z_i 的原像有 k_i 个) 且记 $q_{ij} = p(x_{ij}); j = 1, 2, \dots, k_i, i = 1, 2, \dots, b, p_i = \sum_{j=1}^{k_i} q_{ij}; i = 1, 2, \dots, b$, (p_1, \dots, p_b 是 \mathcal{Z}_0 的概率分布), 那么 q_{ij}, p_i 满足定理 2.3.2 的条件, 而且:

$$H(\xi) = H(q_{11}, \dots, q_{1k_1}, \dots, q_{b1}, \dots, q_{bk_b}) \left(\bigcup_{i=1}^b A_i = \mathcal{X} \right)$$

$$H(f(\xi)) = H(p_1, p_2, \dots, p_b)$$

由定理 2.3.2 知: $H(\xi) \geq H(f(\xi))$, 且等号成立的条件为:

$$\sum_{i=1}^b p_i H\left(\frac{q_{i1}}{p_i}, \frac{q_{i2}}{p_i}, \dots, \frac{q_{ik_i}}{p_i}\right) = 0$$

这时必须有 $p_i = 0$ 或 $\frac{q_{ij}}{p_i} = 1$ 或 0. 因为 $\sum_{i=1}^{k_i} q_{ij} = p_i$, 所以有且只有一个 j 使得 $\frac{q_{ij}}{p_i} = 1$ 即 f 是一个 1-1 映射.

2.3.3 Fano 不等式 (不确定性的上界估计)

Fano 不等式给出了两个随机变量的条件熵与误差之间的关系.

定理 2.3.4 (Fano 不等式)

设 ξ 与 η 为两个离散随机变量, 它们有相同的值域 \mathcal{X} , 且联合分布是 $p(x, y)$, 令 $p_e = P_r\{\xi \neq \eta\}$, 则:

$$H(\xi \mid \eta) \leq H(p_e) + p_e \log(|\mathcal{X}| - 1)$$

其中 $|\mathcal{X}|$ 表示 \mathcal{X} 的元素个数.



证明 因为 $p_e = P_r\{\xi \neq \eta\} = \sum_{x \neq y} p(x, y)$, 所以由条件熵的定义有:

$$\begin{aligned}
 H(\xi | \eta) &= - \sum_{x \neq y} p(x, y) \log p(x | y) - \sum_{x=y} p(x, y) \log p(x | y) \\
 &= - \sum_{x \neq y} p(x, y) \log \frac{p(x, y)}{p(y)} - \sum_{x=y} p(x, y) \log \frac{p(x, y)}{p(y)} \\
 &\leq \left(\sum_{x \neq y} p(x, y) \right) \log \frac{\sum_{x \neq y} p(y)}{\sum_{x \neq y} p(x, y)} + \left(\sum_{x=y} p(x, y) \right) \log \frac{\sum_{x=y} p(y)}{\sum_{x=y} p(x, y)} \\
 &= \left(\sum_{x \neq y} p(x, y) \right) \log \frac{(|\mathcal{X}| - 1)}{\sum_{x \neq y} p(x, y)} + \left(\sum_{x=y} p(x, y) \right) \log \frac{1}{\sum_{x=y} p(x, y)} \\
 &= p_e \log \frac{(|\mathcal{X}| - 1)}{p_e} + (1 - p_e) \log \frac{1}{1 - p_e} \\
 &= p_e \log \frac{1}{p_e} + (1 - p_e) \log \frac{1}{1 - p_e} + p_e \log(|\mathcal{X}| - 1) \\
 &= H(p_e) + p_e \log(|\mathcal{X}| - 1),
 \end{aligned}$$

其中的不等号根据引理2.3.3得来.

2.4 互熵与互信息


前面我们给出了熵、联合熵、条件熵的定义与性质, 这些熵都是概率分布不肯定性的度量. 这一节我们给出两个概率分布“差异性”的度量值, 且把这种“差异性”的度量也看成一种信息量.

2.4.1 互熵

定义 2.4.1

设 $p(x), q(x)$ 是 \mathcal{X} 上的两个概率分布, 它们的互熵定义为

$$H(p||q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}$$

规定 $0 \log \frac{0}{0} = 0$, 另外, 如有一个 $x \in \mathcal{X}$, 使 $p(x) = 0$, 而 $q(x) > 0$, 那么规定 $H(p||q) = \infty$. 

注: (1) 互熵是两个概率分布“差异性”的度量;

(2) 由上一节引理2.3.2知 $H(p||q) \geq 0$;

(3) $H(p||q) = 0 \Leftrightarrow$ 对于 $p(x) \neq 0$ 的 x 均有 $p(x) = q(x)$.

例题 2.4.1 设 $\mathcal{X} = \{x_1, x_2, x_3\}$, 以等概率 $p_i = \frac{1}{3}$ 和以概率 $q_1 = q_2 = \frac{1}{4}, q_3 = \frac{1}{2}$ 从 $\{x_1, x_2, x_3\}$ 中

抽样, 这两个概率分布的互熵是多少?

$$\begin{aligned}
 H(p||q) &= \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)} = \sum_{i=1}^3 p_i \log \frac{p_i}{q_i} \\
 &= \frac{1}{3} \log \frac{1/3}{1/4} + \frac{1}{3} \log \frac{1/3}{1/4} + \frac{1}{3} \log \frac{1/3}{1/2} \\
 &= \frac{1}{3} \log \frac{4}{3} \times \frac{4}{3} \times \frac{2}{3} = \frac{1}{3} \log \frac{32}{27} \\
 &= \frac{5}{3} - \frac{1}{3} \log 27 = \frac{5}{3} - \frac{1}{3} \log 3^3 \\
 &= \frac{5}{3} - \log 3 \approx \frac{5}{3} - 1.585 = 0.082
 \end{aligned}$$

2.4.2 互信息

定义 2.4.2

对于两个随机变量 ξ 与 η , 它的联合分布为 $p(x, y)$, 边际分布为 $p(x)$ 与 $q(y)$, 则 ξ 与 η 的互信息 $I(\xi; \eta)$ 定义为:

$$\begin{aligned}
 I(\xi; \eta) &= H(p(x, y) || p(x)q(y)) \\
 &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)q(y)}
 \end{aligned}$$



定理 2.4.1

由互信息 $I(\xi; \eta)$ 的定义可知如下性质成立:

(1) 对称性: $I(\xi; \eta) = I(\eta; \xi)$;

(2) 互信息与联合熵及条件熵的关系为

$$I(\xi; \eta) = H(\xi) + H(\eta) - H(\xi, \eta);$$

$$I(\xi; \eta) = H(\xi) - H(\xi | \eta);$$

$$I(\xi; \eta) = H(\eta) - H(\eta | \xi);$$

(3) 非负性: 对任何随机变量 ξ, η , 总有 $I(\xi; \eta) \geq 0$, 等号成立的充要条件为 ξ 与 η 是相互独立的随机变量;

(4) 如果 f 是从 \mathcal{Y} 到 \mathcal{Z} 的任意映射, 那么必有 $I(\xi; \eta) \geq I(\xi; f(\eta))$, 等号成立的充要条件为 f 是一个 1-1 变换;

(5) $I(\xi, \xi) = H(\xi)$.



证明 (1) 显然;

(2)

$$\begin{aligned}
I(\xi; \eta) &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)q(y)} \\
&= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) [\log p(x, y) - \log p(x)q(y)] \\
&= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) [\log p(x, y) - \log p(x) - \log q(y)] \\
&= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x, y) - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x) - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log q(y) \\
&= -H(\xi, \eta) + H(\xi) + H(\eta) = H(\xi) + H(\eta) - H(\xi, \eta)
\end{aligned}$$

而

$$\begin{aligned}
H(\xi, \eta) &= H(\eta) + H(\xi | \eta) \\
H(\xi, \eta) &= H(\xi) + H(\eta | \xi)
\end{aligned}$$

于是有

$$I(\xi; \eta) = H(\xi) + H(\eta) - [H(\eta) + H(\xi | \eta)] = H(\xi) - H(\xi | \eta)$$

和

$$I(\xi; \eta) = H(\xi) + H(\eta) - [H(\xi) + H(\eta | \xi)] = H(\eta) - H(\eta | \xi)$$

(3) 显然; (4) 类似于定理2.3.3可证;

(5)

$$\begin{aligned}
I(\xi, \xi) &= \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{p(x)p(x)} \\
&= - \sum_{x \in \mathcal{X}} p(x) \log p(x) = H(\xi)
\end{aligned}$$

2.4.3 条件互信息

定义 2.4.3

随机变量 ξ 与 η 在给定随机变量 ζ 时的条件互信息定义为

$$\begin{aligned}
I(\xi; \eta | \zeta) &= H(p(x, y | \zeta) || p(x | \zeta) \cdot p(y | \zeta)) \\
&= \sum_{(x, y, \zeta) \in \mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}} p(x, y | \zeta) \log \frac{p(x, y | \zeta)}{p(x | \zeta)p(y | \zeta)}
\end{aligned}$$



条件互信息与条件熵的关系为

$$\begin{aligned}
I(\xi; \eta | \zeta) &= H(\xi | \zeta) + H(\eta | \zeta) - H(\xi, \eta | \zeta) \\
&= H(\xi | \zeta) - H(\xi | \eta, \zeta) \\
&= H(\eta | \zeta) - H(\eta | \xi, \zeta)
\end{aligned}$$

例题 2.4.2 设有一信源输出 $\mathcal{X} = \{0, 1, 2\}$, 其概率为 $p_0 = \frac{1}{4}, p_1 = \frac{1}{4}, p_2 = \frac{1}{2}$, 设计一个实验去观察, 其结果为 $\mathcal{Y} \in \{0, 1\}$, 已知条件概率为求 $I(\mathcal{X}; \mathcal{Y})$.

$p(y x)$	0	1
0	1	0
1	0	1
2	$\frac{1}{2}$	$\frac{1}{2}$

解: 对于 \mathcal{Y}

$$\begin{aligned}
 p(y=0) &= \sum_{i=0}^2 p(x_i, y=0) \\
 &= \sum_{i=0}^2 p(x_i) p(y=0 | x_i) \\
 &= p(0)p(y=0 | 0) + p(1)p(y=0 | 1) + p(2)p(y=0 | 2) \\
 &= \frac{1}{4} \times 1 + \frac{1}{4} \times 0 + \frac{1}{2} \times \frac{1}{2} \\
 &= \frac{1}{2} \\
 p(y=1) &= \sum_{i=0}^2 p(x_i, y=1) \\
 &= \sum_{i=0}^2 p(x_i) p(y=1 | x_i) \\
 &= p(0)p(y=1 | 0) + p(1)p(y=1 | 1) + p(2)p(y=1 | 2) \\
 &= \frac{1}{4} \times 0 + \frac{1}{4} \times 1 + \frac{1}{2} \times \frac{1}{2} \\
 &= \frac{1}{2} \\
 I(\mathcal{X}; \mathcal{Y}) &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \\
 &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x)p(y | x) \log \frac{p(y | x)}{p(y)} \\
 &= \frac{1}{4} \times 1 \times \log 2 + \frac{1}{4} \log 2 + \frac{1}{2} \log 1 + \frac{1}{2} \log 1 \\
 &= \frac{1}{4} + \frac{1}{4} \\
 &= 0.5 \text{ (比特/符号)}
 \end{aligned}$$

或者利用 $I(\mathcal{X}; \mathcal{Y}) = H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X})$, 我们计算知 $H(\mathcal{Y}) = H(\frac{1}{2}) = 1, H(\mathcal{Y}|\mathcal{X}) = \frac{1}{4} \times 0 + \frac{1}{4} \times 0 + \frac{1}{2} \times (\frac{1}{2} \log 2 + \frac{1}{2} \log 2) = \frac{1}{2}$, 于是 $I(\mathcal{X}; \mathcal{Y}) = 1 - \frac{1}{2} = \frac{1}{2}$

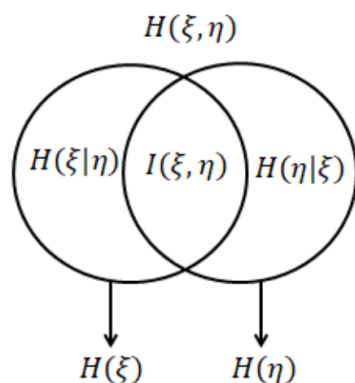
2.4.4 各种熵与互信息之间的关系

$$I(\xi, \eta) = H(\xi) + H(\eta) - H(\xi, \eta)$$

$$H(\xi | \eta) = H(\xi) - I(\xi, \eta)$$

$$H(\xi, \eta) = H(\xi) + H(\eta | \xi)$$

$$I(\xi, \eta) = H(\eta) - H(\eta | \xi)$$



2.5 凸函数及其应用

2.5.1 凸函数的定义与判别

定义 2.5.1

设 $g(x)$ 是定义在区间 (a, b) 上的函数, 如果对任意的 $x_1, x_2 \in (a, b)$ 和任意的 $0 \leq \lambda \leq 1$, 都有

$$g(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda g(x_1) + (1 - \lambda)g(x_2)$$

则称 $g(x)$ 是 (a, b) 上的上凸函数, 如果等号只有当 $\lambda = 0$ 或 $\lambda = 1$ 或 $x_1 = x_2$ 时才成立, 则称函数 g 是严格上凸的 (若不等式相反称是下凸的)。



注:

- (1) 如果 g 是上凸 (严格上凸) 的, 那么 $-g$ 必是下凸 (严格下凸) 的;
- (2) 函数是上凸的, 那么它的函数值总是位于任意弦的上方. 函数是下凸的, 那么它的函数值总是位于任意弦的下面. 如 $x^2, |x|, e^x, x \log x$ 等是下凸函数, $\sqrt{x}, \log x$ 是上凸函数. 注意 $ax + b$ 是上凸也是下凸的.
- (3) 像许多信息量, 如熵和互信息都具有上凸性.

定理 2.5.1 (判定定理)

如果函数 g 在任意处都有非负 (正) 二阶导数, 则函数 g 是下凸 (严格下凸的)



证明 函数 g 在的处的泰勒展开式

$$g(x) = g(x_0) + g'(x_0)(x - x_0) + \frac{1}{2}g''(x_*)(x - x_0)^2$$

其中 x_* 在 x 与 x_0 之间.

由假设 $g''(x_*) \geq 0$. 则有 $g(x) \geq g(x_0) + g'(x_0)(x - x_0)$. 令 $x_0 = \lambda x_1 + (1 - \lambda)x_2$.

当 $x = x_1$ 时有 $g(x_1) \geq g(x_0) + g'(x_0)(x_1 - x_0)$, 而 $x_1 - x_0 = x_1 - [\lambda x_1 + (1 - \lambda)x_2] = (x_1 - x_2)(1 - \lambda)$, 因而有

$$g(x_1) \geq g(x_0) + g'(x_0)(1 - \lambda)(x_1 - x_2) \quad (1)$$

当 $x = x_2$ 时有 $g(x_2) \geq g(x_0) + g'(x_0)(x_2 - x_0)$ 而 $x_2 - x_0 = x_2 - [\lambda x_1 + (1 - \lambda)x_2] = \lambda(x_2 - x_1)$, 因而有

$$g(x_2) \geq g(x_0) + g'(x_0)\lambda(x_2 - x_1) \quad (2)$$

$$\begin{aligned} (1) \times \lambda + (2) \times (1 - \lambda) &= \lambda g(x_1) + (1 - \lambda)g(x_2) \\ &\geq \lambda g(x_0) + \lambda g'(x_0)(1 - \lambda)(x_1 - x_2) \\ &\quad + (1 - \lambda)g(x_0) + \lambda g'(x_0)(1 - \lambda)(x_2 - x_1) \\ &= g(x_0) \end{aligned}$$

即证下凸性, 同理可证严格下凸性.

易知 $x \geq 0$ 时, $x^2, |x|, e^x, x \log x$ 均是严格下凸的. $\sqrt{x}, \log x$ 是严格上凸的. 如 $g(x) = x \log x$, 有 $g'(x) = 1 + \log x$, $g''(x) = \frac{1}{x} > 0$

2.5.2 熵函数的凸性

若 \mathcal{X} 是一个固定的集合, 记

$$\tilde{\mathcal{P}} = \left\{ \bar{p} = p(x) \mid x \in \mathcal{X}, p(x) \geq 0, \sum_{x \in \mathcal{X}} p(x) = 1 \right\}$$

是 \mathcal{X} 上的全体概率分布, 那么 $H(\bar{p}), H(p||q)$ 都是 $\tilde{\mathcal{P}}$ 上的函数.

定理 2.5.2

- (1) 熵函数 $H(\bar{p})$ 是 $\tilde{\mathcal{P}}$ 上的上凸函数, 即对 $\forall 0 \leq \lambda \leq 1, \forall \bar{p}_1, \bar{p}_2 \in \tilde{\mathcal{P}}$, 总有 $H(\lambda \bar{p}_1 + (1 - \lambda)\bar{p}_2) \geq \lambda H(\bar{p}_1) + (1 - \lambda)H(\bar{p}_2)$ 等号成立的充要条件为 $\lambda = 0$ 或 1 , 或 $\bar{p}_1 = \bar{p}_2$;
- (2) 在 q 固定时, 互熵函数 $H(p||q)$ 是 p 的下凸函数, 在 p 固定时, 互熵函数 $H(p||q)$ 是 q 的下凸函数.



2.6 连续型随机变量的信息量

2.6.1 连续型随机变量的 Shannon 熵

1. 连续型随机变量

定义 2.6.1

设 ξ 是一个随机变量, 在 $\mathbb{R} = (-\infty, +\infty)$ (或 \mathbb{R} 的某个区域 \mathcal{X}) 中取值, 称 ξ 是一个连续型随机变量, 它的概率分布函数定义为

$$F(x) = P_r\{\xi \leq x\}, x \in \mathbb{R}$$

如果 $F(x)$ 是连续函数, 那么称随机变量 ξ 具有连续分布, 若 $F(x)$ 的导数存在, 则 $f(x) = F'(x) = \frac{dF(x)}{d(x)}$ 是 ξ 的概率分布密度函数



注: 对于概率分布密度函数 $f(x)$, 有 $f(x) \geq 0$, 对 $\forall x \in \mathcal{X}$,

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

2. 连续型随机变量的 Shannon 熵

定义 2.6.2

一个概率密度函数为 $f(x)$ 的连续型随机变量 ξ 的熵 $H(\xi)$ 定义为

$$H(\xi) = - \int_{\mathcal{X}} f(x) \log f(x) dx.$$

规定 $0 \log 0 = 0$, 也称为微分熵



例题 2.6.1 (一致分布) 随机变量 ξ 在 0 到 a 之间为一致分布, 于是在 0 到 a 之间的密度为 $\frac{1}{a}$, 而在其它处的密度均为 0, 它的熵为

$$H(\xi) = - \int_0^a \frac{1}{a} \log \frac{1}{a} dx = \log a$$

注: 当 $a < 1$ 时, $\log a < 0$, 此时熵为负值, 此时并不代表信息的不确定性的, 而在物理学中有其意义.

例题 2.6.2 (指数分布) 设 $\xi \sim p_\lambda(x) = \lambda e^{-\lambda x}, x, \lambda \geq 0$, 它的熵为

$$\begin{aligned} H(\xi) &= - \int_0^\infty p_\lambda(x) \log p_\lambda(x) dx = - \int_0^\infty \lambda e^{-\lambda x} \log (\lambda e^{-\lambda x}) dx \\ &= -\lambda \int_0^\infty e^{-\lambda x} (\log \lambda - \lambda x) dx \\ &= -\lambda \int_0^\infty e^{-\lambda x} (-\lambda x) dx - \int_0^\infty \lambda e^{-\lambda x} \log \lambda dx \\ &= \int_0^\infty e^{-\lambda x} (-\lambda x) d(-\lambda x) - \log \lambda \\ &= 1 - \log \lambda \end{aligned}$$

例题 2.6.3 (正态分布) 如果 $\xi \sim N(\mu, \sigma^2)$, 其中 $N(\mu, \sigma^2)$ 表示期望为 μ , 均方为 σ^2 的正态分

布, 那么它的分布密度为

$$\phi_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

则它的熵为

$$\begin{aligned} H(\phi_{\mu, \sigma^2}) &= - \int \phi_{\mu, \sigma^2}(x) \log \phi_{\mu, \sigma^2}(x) dx \\ &= - \int \phi_{\mu, \sigma^2}(x) \left(-\frac{(x-\mu)^2}{2\sigma^2} - \log \sqrt{2\pi\sigma^2} \right) dx \\ &= \frac{E\{(\xi-\mu)^2\}}{2\sigma^2} + \frac{1}{2} \log 2\pi\sigma^2 \\ &= \frac{1}{2} + \frac{1}{2} \log 2\pi\sigma^2 \\ &= \frac{1}{2} \log 2\pi e\sigma^2 \end{aligned}$$

2.6.2 连续型随机变量的联合熵与条件熵

(ξ, η) 是一对连续型随机变量, 联合分布为 $p(x, y)$, 边际分布为 $p(x), p(y)$, 则有如下定义:

(1) 联合熵

$$H(\xi, \eta) = - \int p(x, y) \log p(x, y) dx dy$$

(2) 条件熵

$$H(\eta | \xi) = - \int p(x, y) \log \frac{p(x, y)}{p(x)} dx dy$$

(3) 互信息

$$I(\xi; \eta) = \int p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy$$

2.7 最大熵原理

连续型随机变量 ξ , 它的分布密度为 $f(x)$, $H(\xi) = - \int_{\mathcal{X}} f(x) \log f(x) dx$, ξ 服从何种概率分布密度函数时, $H(\xi)$ 最大?

即取合适的概率分布密度函数, 使 $H(\xi)$ 最大.

2.7.1 有限区间情形的最大熵 (均匀分布时取最大熵)

设 ξ 是有限区间 $\mathcal{X} = (a, b)$ 上取值的随机变量, 约束条件为

$$\int_a^b f(x) dx = 1$$

则 $H(\xi)$ 的最大值为 $\log(b-a)$, 即 $H(\xi) \leq \log(b-a)$.

2.7.2 半开区间情形的最大熵（指数分布时取最大熵）

设 ξ 是在半区间 $\mathcal{X} = (0, \infty)$ 上取值的随机变量, 约束条件为

$$\int_0^{\infty} f(x) dx = 1$$

其期望固定值为 $\int_0^{\infty} x f(x) dx = \mu > 0$, 则 $H(\xi)$ 的最大值为 $1 + \log \mu$, 即 $H(\xi) \leq 1 + \log \mu$, ($\mu = \frac{1}{\lambda}$)

2.7.3 全区间情形的最大熵（正态分布时取最大熵）

设 ξ 是在全区间 $(-\infty, +\infty)$ 上取值的随机变量, 约束条件为

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

其期望和方差分别固定为 $E(\xi) = \mu$ 和 $D(\xi) = \sigma^2 > 0$, 即 $\int_{-\infty}^{+\infty} x f(x) dx = \mu$ (数学期望定义),

$$\int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \sigma^2 \text{ (方差定义),}$$

则 $H(\xi)$ 的最大值为 $\frac{1}{2} \log (2e\pi\sigma^2)$,


$$\text{即 } H(\xi) \leq \frac{1}{2} \log (2e\pi\sigma^2)$$

2.8 习题课

 **练习 2.8.1** 计算 $H\left(\frac{1}{a}, \frac{1}{a}, \dots, \frac{1}{a}, \frac{2}{a}, \frac{2}{a}\right)$


解: 由 $\sum P_i = 1$ 知, 含 $(a-4)$ 个 $\frac{1}{a}$, 2 个 $\frac{2}{a}$, 总共 $(a-2)$ 项, 于是

$$\begin{aligned} H\left(\frac{1}{a}, \frac{1}{a}, \dots, \frac{1}{a}, \frac{2}{a}, \frac{2}{a}\right) &= \sum_{i=1}^{a-2} p_i \cdot \log \frac{1}{p_i} \\ &= \sum_{i=1}^{a-4} \frac{1}{a} \log a + 2 \cdot \frac{2}{a} \log \frac{a}{2} \\ &= \frac{a-4}{a} \cdot \log a + \frac{4}{a} \log \frac{a}{2} \\ &= \frac{a-4}{a} \cdot \log a + \frac{4}{a} \log a - \frac{4}{a} \log 2 \\ &= \log a - \frac{4}{a} \log 2 \end{aligned}$$

 **练习 2.8.2** 计算 $H'(p)$, 其中 $H(p)$ 为熵函数

解:

$$\begin{aligned}
 H(p) &= -p \log p - (1-p) \log(1-p) \\
 H'(p) &= -\log p - p \cdot \frac{1}{\ln 2 \cdot p} + \log(1-p) + \frac{1}{1-p} \cdot \frac{1-p}{\ln 2} \\
 &= -\log p - \frac{1}{\ln 2} + \log(1-p) + \frac{1}{\ln 2} \\
 &= -\log p + \log(1-p) \\
 &= \log \frac{1-p}{p}
 \end{aligned}$$

 **练习 2.8.3** 设两只口袋中各有 20 个球, 第一支口袋中有 10 个白球, 5 个黑球和 5 个红球; 第二只口袋中有 8 个白球, 8 个黑球和 4 个红球, 从每只口袋中各取一个球, 试判断哪一个结果的不肯定性更大.

解: 当我们要判断哪个结果的不确定性更大时, 可以使用熵来衡量. 首先, 我们将第一只口袋的球的颜色作为随机变量 ξ_1 , 它的概率分布为:

$$\xi_1 \sim \begin{pmatrix} \text{白} & \text{黑} & \text{红} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

其中, $\frac{1}{2}$ 表示白球的概率, $\frac{1}{4}$ 表示黑球的概率, $\frac{1}{4}$ 表示红球的概率.

计算第一只口袋的熵 $H(\xi_1)$:

$$H(\xi_1) = \frac{1}{2} \log 2 + 2 \times \frac{1}{4} \log 4 = \frac{1}{2} + 1 = 1.5 \text{ bits}$$

接下来, 我们将第二只口袋的球的颜色作为随机变量 ξ_2 , 它的概率分布为:


$$\xi_2 \sim \begin{pmatrix} \text{白} & \text{黑} & \text{红} \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \end{pmatrix}$$

其中, $\frac{2}{5}$ 表示白球的概率, $\frac{2}{5}$ 表示黑球的概率, $\frac{1}{5}$ 表示红球的概率.

计算第二只口袋的熵 $H(\xi_2)$:

$$\begin{aligned}
 H(\xi_2) &= \frac{4}{5} \log \frac{5}{2} + \frac{1}{5} \log 5 \\
 &= \frac{4}{5} (\log 5 - \log 2) + \frac{1}{5} \log 5 \\
 &= \frac{4}{5} \log 5 + \frac{1}{5} \log 5 - \frac{4}{5} \\
 &= \log 5 - \frac{4}{5} \approx 2.32 - 0.8 \\
 &= 1.52 \text{ 比特}
 \end{aligned}$$

比较 $H(\xi_1)$ 和 $H(\xi_2)$ 的值, 我们可以得出结论: 第二只口袋的结果的不确定性更大, 因为它的熵值更大.

 **练习 2.8.4** 设 ξ 和 η 联合分布 $p(0,0) = \frac{1}{3}, p(0,1) = \frac{1}{3}, p(1,0) = 0, p(1,1) = \frac{1}{3}$, 试求:

(1) $H(\xi), H(\eta)$;

(2) $H(\xi | \eta), H(\eta | \xi)$

- (3) $H(\xi, \eta)$;
 (4) $H(\eta) - H(\eta | \xi)$;
 (5) $I(\xi; \eta)$;
 (6) 画出上述各信息之间关系的韦恩图.

解:

$\eta \backslash \xi$	0	1	Σ
0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
Σ	$\frac{2}{3}$	$\frac{1}{3}$	1

$$\xi \sim \begin{pmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} \quad \eta \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

(1)

$$\begin{aligned} H(\xi) &= \sum_{i=0}^1 p_i \log \frac{1}{p_i} = \frac{2}{3} \log \frac{3}{2} + \frac{1}{3} \log 3 \\ &= \frac{2}{3} \log 3 - \frac{2}{3} \log 2 + \frac{1}{3} \log 3 \\ &= \log 3 - \frac{2}{3} \end{aligned}$$

$$\begin{aligned} H(\eta) &= \sum_{i=0}^1 p_i \log \frac{1}{p_i} = \frac{1}{3} \log 3 + \frac{2}{3} \log \frac{3}{2} \\ &= \frac{1}{3} \log 3 + \frac{2}{3} \log 3 - \frac{2}{3} \log 2 \\ &= \log 3 - \frac{2}{3} \end{aligned}$$

(3)

$$\begin{aligned} H(\xi, \eta) &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{1}{p(x, y)} \\ &= 3 \cdot \frac{1}{3} \log 3 = \log 3 \end{aligned}$$

(2)


$$\begin{aligned} H(\eta | \xi) &= H(\xi, \eta) - H(\xi) \\ &= \log 3 - \left(\log 3 - \frac{2}{3} \right) = \frac{2}{3} \\ H(\xi | \eta) &= H(\xi, \eta) - H(\eta) \\ &= \log 3 - \left(\log 3 - \frac{2}{3} \right) = \frac{2}{3} \end{aligned}$$

(4)

$$\begin{aligned} H(\eta) - H(\eta | \xi) &= \log 3 - \frac{2}{3} - \frac{2}{3} \\ &= \log 3 - \frac{4}{3} \end{aligned}$$

(5)

$$\begin{aligned} I(\xi; \eta) &= H(\xi) + H(\eta) - H(\xi, \eta) \\ &= 2 \left(\log 3 - \frac{2}{3} \right) - \log 3 \\ &= \log 3 - \frac{4}{3} \end{aligned}$$

 **练习 2.8.5** 设 $\{p_1, p_2, \dots, p_a\}$ 是一个概率分布, 并令 $q_m = p_{m+1} + \dots + p_a$, 证明: $H(p_1, \dots, p_a) \leq H(p_1, \dots, p_m, q_m) + q_m \log(a - m)$, 并指出等号何时成立.

解: 证明: 对于序列 p_1, \dots, p_m, q_m , 有 $\sum_{i=1}^m p_i + q_m = 1$, 其中 $q_m = \sum_{j=m+1}^a p_j$, 也即

$$p_1, \dots, p_m, p_{m+1}, \dots, p_a \quad \sum_{i=1}^a p_i = 1 \quad (i \leq m, q_i = p_i)$$

由熵函数可加性知:

$$\begin{aligned} &H(p_1, \dots, p_m, p_{m+1}, \dots, p_a) \\ &= H(p_1, \dots, p_m, q_m) + \sum_{i=1}^m p_i H\left(\frac{p_i}{q_m}\right) + q_m \cdot H\left(\frac{p_{m+1}}{q_m}, \dots, \frac{p_a}{q_m}\right) \\ &\leq H(p_1, \dots, p_m, q_m) + q_m \log(a - m) \quad (\text{最大值定理}) \\ &\quad \frac{p_{m+1}}{q_m} = \dots = \frac{p_a}{q_m} \Rightarrow p_{m+1} = \dots = p_a \text{ 时等号成立.} \end{aligned}$$

证明二: 首先证明 $H(p_1, \dots, p_a) = H(p_1, \dots, p_m, q_m) + q_m \cdot H\left(\frac{p_{m+1}}{q_m}, \frac{p_{m+2}}{q_m}, \dots, \frac{p_a}{q_m}\right)$.

$$\begin{aligned} \text{右边} &= H(p_1, \dots, p_m, q_m) + q_m \cdot H\left(\frac{p_{m+1}}{q_m}, \frac{p_{m+2}}{q_m}, \dots, \frac{p_a}{q_m}\right) \\ &= \left(- \sum_{i=1}^m p_i \log p_i - q_m \log q_m \right) - q_m \cdot \sum_{i=m+1}^a \frac{p_i}{q_m} \log \frac{p_i}{q_m} \\ &= - \sum_{i=1}^m p_i \log p_i - q_m \log q_m - \sum_{i=m+1}^a p_i \log p_i + \log q_m \cdot \sum_{i=m+1}^a p_i \\ &= - \sum_{i=1}^a p_i \log p_i = \text{左边} \end{aligned}$$


由离散最大熵定理有

$$H\left(\frac{p_{m+1}}{q_m}, \frac{p_{m+2}}{q_m}, \dots, \frac{p_a}{q_m}\right) \leq \log(a - m)$$

因此有

$$H(p_1, \dots, p_a) \leq H(p_1, \dots, p_m, q_m) + q_m \log(a - m)$$

等式成立的条件是 $p_{m+1} = p_{m+2} = \cdots = p_a = \frac{q_m}{a-m}$

 **练习 2.8.6** 设 ξ 是取 m 个值 x_1, x_2, \cdots, x_m 的随机变量, $p(\xi = x_m) = a$, 证明: $H(\xi) = a \log \frac{1}{a} + (1-a) \log \frac{1}{1-a} + (1-a)H(\eta)$, 其中 η 是取 $m-1$ 个值 $x_1, x_2, \cdots, x_{m-1}$ 的随机变量,

$$p(\eta = x_j) \stackrel{\text{def}}{=} p(\xi = x_j) / (1-a), 1 \leq j \leq m-1.$$

进一步证明: $H(\xi) \leq a \log \frac{1}{a} + (1-a) \log \frac{1}{1-a} + (1-a) \log(m-1)$, 并确定其中等号成立的条件.

解: 证明: 由 $p(\eta = x_j) \stackrel{\text{def}}{=} p(\xi = x_j) / (1-a) \Rightarrow p(\xi = x_j) = (1-a)p(\eta = x_j), j = 1, \cdots, m-1$.


$$\begin{aligned} \text{故 } H(\xi) &= a \log \frac{1}{a} + \sum_{j=1}^{m-1} p(\xi = x_j) \log \frac{1}{p(\xi = x_j)} \\ &= a \log \frac{1}{a} + \sum_{j=1}^{m-1} (1-a)p(\eta = x_j) \log \frac{1}{(1-a)p(\eta = x_j)} \\ &= a \log \frac{1}{a} + \sum_{j=1}^{m-1} (1-a)p(\eta = x_j) \log \frac{1}{1-a} + \sum_{j=1}^{m-1} (1-a)p(\eta = x_j) \log \frac{1}{p(\eta = x_j)} \\ &= a \log \frac{1}{a} + (1-a) \log \frac{1}{1-a} \sum_{j=1}^{m-1} p(\eta = x_j) + (1-a) \sum_{j=1}^{m-1} p(\eta = x_j) \log \frac{1}{p(\eta = x_j)} \\ &= a \log \frac{1}{a} + (1-a) \log \frac{1}{1-a} + (1-a)H(\eta) \end{aligned}$$

根据熵的最大值定理有 $H(\eta) \leq \log(m-1)$ 等号成立的条件为 $p(\eta = x_j)$ 为等概率分布 $p(\eta = x_j) = p(\xi = x_j) / (1-a)$

$$(m-1)p(\eta = x_j) = 1, \quad p(\eta = x_j) = \frac{1}{m-1}$$

即

$$p(\xi = x_j) = \frac{1-a}{m-1}, \quad j = 1, 2, \cdots, m-1.$$

 **练习 2.8.7** 设 $\tilde{p} = \{p_1, p_2, \cdots, p_a\}$ 是一个概率分布, 满足 $p_1 \geq p_2 \geq \cdots \geq p_a$, 假设 $\varepsilon \geq 0$, 使得 $p_1 - \varepsilon \geq p_2 + \varepsilon$ 成立, 证明:

$$H(p_1, p_2, \cdots, p_a) \leq H(p_1 - \varepsilon, p_2 + \varepsilon, p_3, \cdots, p_a)$$


解: 证明: $H(p_1, p_2, \cdots, p_a) = -p_1 \log p_1 - p_2 \log p_2 - \sum_{i=3}^a p_i \log p_i$

$$\begin{aligned} &H(p_1 - \varepsilon, p_2 + \varepsilon, p_3, \cdots, p_a) \\ &= -(p_1 - \varepsilon) \log(p_1 - \varepsilon) - (p_2 + \varepsilon) \log(p_2 + \varepsilon) - \sum_{i=3}^a p_i \log p_i \\ &= -p_1 \log(p_1 - \varepsilon) - p_2 \log(p_2 + \varepsilon) - \sum_{i=3}^a p_i \log p_i + \varepsilon \log \frac{p_1 - \varepsilon}{p_2 + \varepsilon} \end{aligned}$$

由引理2.3.2知: $-p_1 \log(p_1 - \varepsilon) - p_2 \log(p_2 + \varepsilon) - \sum_{i=3}^a p_i \log p_i$

$$\geq -p_1 \log p_1 - p_2 \log p_2 - \sum_{i=3}^a p_i \log p_i$$

而 $\varepsilon \log \frac{p_1 - \varepsilon}{p_2 + \varepsilon} \geq 0$, 故 $H(p_1, p_2, \dots, p_a) \leq H(p_1 - \varepsilon, p_2 + \varepsilon, p_3, \dots, p_a)$

 **练习 2.8.8** 设 ξ_1, ξ_2 具有相同的分布, 但它们不需要是独立的, 令 $\rho = 1 - \frac{H(\xi_1 | \xi_2)}{H(\xi_1)}$

(1) 证明: $\rho = \frac{I(\xi_1; \xi_2)}{H(\xi_1)}$;

(2) 证明: $0 \leq \rho \leq 1$;

(3) 何时 $\rho = 0$?

(4) 何时 $\rho = 1$?


解: 证明:

$$(1) \rho = 1 - \frac{H(\xi_1 | \xi_2)}{H(\xi_1)} = \frac{H(\xi_1) - H(\xi_1 | \xi_2)}{H(\xi_1)} = \frac{I(\xi_1; \xi_2)}{H(\xi_1)};$$

(2) 因 $I(\xi_1; \xi_2) \geq 0$, $H(\xi_1) \geq H(\xi_1 | \xi_2)$. 所以 $0 \leq \frac{H(\xi_1 | \xi_2)}{H(\xi_1)} \leq 1$ 因此 $0 \leq \rho \leq 1$;

(3) $\rho = 0$ 即: $H(\xi_1 | \xi_2) = H(\xi_1)$, 即 ξ_1 与 ξ_2 相互独立;


(4) $\rho = 1$ 即: $H(\xi_1 | \xi_2) = 0$, 或 $I(\xi_1; \xi_2) = H(\xi_1)$. 即 $\xi_1 = \xi_2$.

 **练习 2.8.9** 居住某地区的女孩子中有 25% 是大学生, 在女大学生中有 75% 是身高为 1.6 米以上的, 而女孩中身高 1.6 米以上的占总数的一半, 假如我们得知身高 1.6 米以上的某女孩是大学学生的消息, 问获得多少信息量.

解: 设事件 A 为女孩是大学生, 事件 B 为女孩子身高 1.6 米以上, 知: $p(A) = 0.25$, $p(B) = 0.5$, $p(B | A) = 0.75$. 身高 1.6 米以上的某女孩是大学生这消息表明在事件 B 的条件下 A 事件发生, 可得:

$$p(A | B) = \frac{p(AB)}{p(B)} = \frac{p(A)p(B | A)}{p(B)} = \frac{0.25 \times 0.75}{0.5} = 0.375$$

因而 $I(A | B) = -\log p(A | B) = \log \frac{1}{0.375}$ 比特

 **练习 2.8.10** 设某一彩色电视机分辨率为 500×1000 , 灰度为 10, 不同的色彩度为 30, 求一幅电视画面所含信息量的大小.

解: ξ_1 为灰度信源的随机变量

$$\xi_1 \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_{10} \\ \frac{1}{10} & \frac{1}{10} & \cdots & \frac{1}{10} \end{pmatrix}$$

每个像素灰度含有的信息量为 $H(\xi_1) = 10 \cdot \frac{1}{10} \log 10 = \log 10 \approx 3.32$ 比特则每副电视画面含信息量为

$$5 \times 10^5 \times \log 10 \approx 1.66 \times 10^6 \text{ 比特}$$

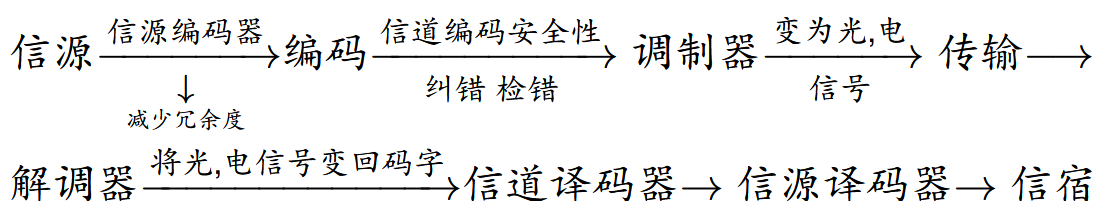
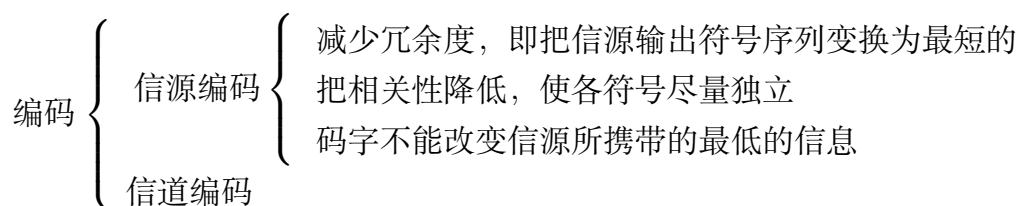
设 ξ_2 为色彩度信源的随机变量

$$\xi_2 \sim \begin{pmatrix} y_1 & y_2 & \cdots & y_{30} \\ \frac{1}{30} & \frac{1}{30} & \cdots & \frac{1}{30} \end{pmatrix}$$

每个色彩度含有的信息量为 $H(\xi_2) = \log 30 \approx 4.91$ 像素亮度与色彩互相独立, 故每个像素含有的信息量为

$$H(\xi_1, \xi_2) = H(\xi_1) + H(\xi_2) = (\log 10 + \log 30) \times 5 \times 10^5 \text{ 比特}.$$

第3章 信源编码



3.1 信源编码问题

3.1.1 信源编码

定义 3.1.1

信源序列: 设 $\mathcal{S} = \{\mathcal{X}, p(x)\}$ 为信源, 其中 \mathcal{X} 为信息字母表, $p(x), x \in \mathcal{X}$ 为消息字母表上的概率分布, 将 \mathcal{X} 中的信源字母进行分组, 记为 $x_1^{(n)}, x_2^{(n)}, \dots, x_L^{(n)}$. 其中 $x_i^{(n)} = \{x_{i1}, x_{i2}, \dots, x_{in}\}, i = 1, \dots, L, x_{ij} \in X, j = 1, \dots, n$. 每个 $x_i^{(n)}$ 的概率分布记为 $p_i^{(n)}$, 记 $\mathcal{S}^n = \{\mathcal{X}^n, p^{(n)}(x^{(n)})\}$ 称其为信源序列.



注:

(1) 发送消息时通常不是一个一个信源字母进行发送, 而是把它们分组变为信源序列后进行发送, 可以减少发送次数.

(2) 例: $\mathcal{X} = \{0, 1\}, p(0) = p, p(1) = 1 - p, x_1^{(3)} = 001, x_2^{(3)} = 111, x_3^{(3)} = 101$, 则 $p_1^{(3)} = (p, p, 1 - p), p_2^{(3)} = (1 - p, 1 - p, 1 - p), p_3^{(3)} = (1 - p, p, 1 - p)$. 于是信源序列为 $\{\mathcal{X}^3, p^{(3)}(x^{(3)})\}$.

定义 3.1.2

记 \mathcal{U} 和 \mathcal{U}^m 分别为信道的输入信号字母表和输入信号字母表序列, $\mathcal{U}^* = \bigcup_{m=1}^{\infty} \mathcal{U}^m$ 为全体信号字母串所组成的集合, 其中 \mathcal{U}^m 是全体在 \mathcal{U} 上取值的 m 维向量的集合, 则称映射 $f: \mathcal{X}(\mathcal{X}^n) \rightarrow \mathcal{U}^*$ 为信源编码 f .



注:

- (1) $f: \mathcal{X}^n \rightarrow \mathcal{U}^*$ 即 f 可把 n 长的消息序列变为信号序列 (可以不等长);
- (2) 信源编码的基本要求是编码运算的可还原性, 即可唯一的把码字还原为消息.
- (3) 编码可还原性的一个必要条件是编码运算 f 的 1-1 性 (显然, 后面举例说明 1-1 性是必要条件而不是充要条件)

定义 3.1.3

1-1 码: 如果对于 $\forall x \neq x' \in \mathcal{X}$, 都有 $f(x) \neq f(x')$, 则称 f 是一个 1-1 编码, 也称 f 满足 1-1 性.



注: 1-1 码不一定是可还原码, 后面举例说明.

3.1.2 定长编码与变长编码

1. 定长编码与变长编码

定义 3.1.4

若编码 $f: \mathcal{X}^n \rightarrow \mathcal{U}^m$, 其中 m, n 是两个固定的正整数, 则称 f 是一个定长编码, 即把 $(x_1, \dots, x_n) \rightarrow (u_1, \dots, u_m)$ (码字长度固定) 若 $f: \mathcal{X} \rightarrow \mathcal{U}^*$, 称 f 为变长编码 (码字长度不固定)



2. 扩张编码与唯一可译码

定义 3.1.5

(1) 设 $f: \mathcal{X}^n \rightarrow \mathcal{U}^m$ 定长编码. 记 $x^{(kn)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_k^{(n)})$, 其中 $x_j^{(n)} = (x_{j1}, \dots, x_{jn})$ $j = 1, 2, \dots, k$. 对任何的 $k = 1, 2, 3, \dots$, 定义

$$f^*(x^{(kn)}) = (f(x_1^{(n)}), f(x_2^{(n)}), \dots, f(x_k^{(n)}))$$

则称 f^* 是 f 的扩张编码. 这时 f^* 是一个从 $(\mathcal{X}^n)^*$ 到 \mathcal{U}^* 的映射.

(2) 设 f 是 \mathcal{X} 到 \mathcal{U}^* 的变长码, $x^{(k)} = (x_1, x_2, \dots, x_k)$ 是一个消息字母串, 对任何 $k = 1, 2, 3, \dots$, 定义

$$f^*(x^{(k)}) = (f(x_1), f(x_2), \dots, f(x_k))$$

称 f^* 是 f 的扩张编码. 这时 f^* 是一个从 \mathcal{X}^* 到 \mathcal{U}^* 的映射.

(3) 无论是定长编码还是变长编码, 它们的扩张 f^* 都是从 $(\mathcal{X}^n)^*$ (或 \mathcal{X}^*) 到 \mathcal{U}^* 的映射. 如果 f^* 是一个 1-1 映射, 那么我们称 f 是一个唯一可译码 (或可还原码).



3. 唯一可译码与 1-1 码的关系

定理 3.1.1

- (1) 无论是定长编码还是变长编码, 唯一可译码必是 1-1 码.
- (2) 如果 f 是一个 1-1 定长码, 那么 f 是一个唯一可译码.
- (3) 唯一可译的变长码 f 一定是 1-1 码, 反之不然.



证明 (1) 用反证法. 若 f 不是 1-1 码, 由扩张编码的定义 f^* 也不是一一映射, 与 f 是唯一可译码矛盾, 故 f 必是 1-1 码.

(2) 证明 f 是唯一可译码, 只需证明 f^* 是一一映射 (f 是定长码), 即 f^* 是单射. 设 $x^{(kn)} \neq y^{(kn)}$, 那么必有一个 $j \in \{1, 2, \dots, k\}$ 使 $x_j^{(n)} \neq y_j^{(n)}$, 因为 f 是 1-1 码, 所以必有 $f(x_j^{(n)}) \neq f(y_j^{(n)})$, 从而

$$\left(f(x_1^{(n)}), f(x_2^{(n)}), \dots, f(x_k^{(n)})\right) \neq \left(f(y_1^{(n)}), f(y_2^{(n)}), \dots, f(y_k^{(n)})\right),$$

即 $f^*(x^{(kn)}) \neq f^*(y^{(kn)})$, f^* 是 1-1 映射 (单射). 从而 f 是唯一可译码.

(3) 由 (1) 知唯一可译的变长码 f 一定是 1-1 码. 下面举例说明对于变长码来说, f 是 1-1 码, f 不一定是唯一可译码, 即 f^* 不一定是 1-1 映射.

取 $\mathcal{X} = \{a, b, c\}$, $\mathcal{U} = \{0, 1\}$, 令 $f(a) = 0, f(b) = 01, f(c) = 001$. 则 f 是一个 1-1 变长码, 但 f 不是唯一可译码.

事实上, $f^*(c) = (0, 0, 1), f^*(a, b) = (f(a), f(b)) = (0, 0, 1), (a, b) \neq c$, 但 $f^*(a, b) = f^*(c)$, 因此 f^* 不是 \mathcal{X}^* 到 \mathcal{U}^* 的 1-1 映射, 故 f 不是唯一可译码.

4. 一些概念

(1) 二源码: 如果信号字母集 $\mathcal{U} = \{0, 1\}$, 则称相应的码 (定长或者变长) 为二源码, 称 $\mathcal{U}_f = \{f(x) \mid x \in \mathcal{X}\}$ 为码元集, 记作 $C = \mathcal{U}_f = \{c_1, c_2, \dots, c_a\}$, 其中 $c_i = f(x_i)$.

(2) 记 $\ell_f(x)$ 为码字 $f(x)$ 的长度, $\ell_i = \ell_f(x_i)$ 为 x_i 对应的码字 $f(x_i)$ 的长度, 此时码元集 $C = \mathcal{U}_f = \{u_i^{(\ell_i)}, i = 1, 2, \dots, a\}$,

$$u_i^{(\ell_i)} = (u_{i1}, u_{i2}, \dots, u_{i\ell_i}) = f(x_i), u_{ij} \in \mathcal{U}$$

此时有 $C = \{c_1, c_2, \dots, c_a\} = \{u_1^{(\ell_1)}, u_2^{(\ell_2)}, \dots, u_a^{(\ell_a)}\}$, 如上例 $C = \mathcal{U}_f = \{0, 01, 001\}, \ell_1 = 1, \ell_2 = 2, \ell_3 = 3$, 因此有

$$C = \{u_1^{(1)}, u_2^{(2)}, u_3^{(3)}\}.$$

3.1.3 信源变长码的编码问题

定义 3.1.6 (变长编码 f 的平均码长)

设 $\mathcal{S} = \{\mathcal{X}, p(x)\}$ 是一个信源, f 是一个变长编码, 对于 $\forall x \in \mathcal{X}, f(x) \in \mathcal{U}^*$, 记 $\ell_f(x)$ 是 $f(x)$ 的向量长度, 定义

$$L(\mathcal{S}, f) = \sum_{x \in \mathcal{X}} p(x) \ell_f(x)$$

为变长编码 f 的平均码长.



注: (1) 平均码长小占存储空间少, 易于传输. (2) 平均码长与概率分布密切相关.

例题 3.1.1 考虑信源 $\mathcal{S} = \begin{pmatrix} a & b & c & d \\ \frac{2}{17} & \frac{2}{17} & \frac{9}{17} & \frac{4}{17} \end{pmatrix}$ $\mathcal{X} = \{a, b, c, d\}$.

$\mathcal{U} = \{0, 1\}$ 是二进制信号字母表, f_1, f_2 是两个编码方案,

$$f_1: f_1(a) = 11, f_1(b) = 0, f_1(c) = 100, f_1(d) = 10$$

$$f_2: f_2(a) = 01010, f_2(b) = 00, f_2(c) = 10, f_2(d) = 11$$

计算它们的平均码长

$$L(\mathcal{S}, f_1) = \frac{2}{17} \times 2 + \frac{2}{17} \times 1 + \frac{9}{17} \times 3 + \frac{4}{17} \times 2 = \frac{41}{17}$$

$$L(\mathcal{S}, f_2) = \frac{2}{17} \times 5 + \frac{2}{17} \times 2 + \frac{9}{17} \times 2 + \frac{4}{17} \times 2 = \frac{40}{17}$$

定义 3.1.7 (变长信源编码问题)

如果 $\mathcal{S} = \{\mathcal{X}, p(x)\}$ 是一个给定的信源, 变长信源编码问题是: 求一个唯一可译的变长码 f , 使 $L(\mathcal{S}, f)$ 最小, 即求唯一可译的变长码 f_0 , 使得 f_0 相对于其他唯一可译变长码 f , 总有

$$L(\mathcal{S}, f_0) \leq L(\mathcal{S}, f)$$

这时, 称 f_0 为 \mathcal{S} 的最优变长码.



3.1.4 信源序列的定长编码问题

记 $\mathcal{S}^n = \{\mathcal{X}^n, p^{(n)}(x^{(n)})\}$, $n = 1, 2, 3, \dots$, 是信源序列. \mathcal{U} 是输入信号字母表, $\mathcal{U}^{(m)}$ 是 \mathcal{U} 的 m 维乘积空间, 即 $\mathcal{U}^{(m)} = \{(u_1, \dots, u_m) \mid u_i \in \mathcal{U}\}$. 定长编码 (f, g) 分别是 $f: \mathcal{X}^n \rightarrow \mathcal{U}^m$, $g: \mathcal{U}^m \rightarrow \mathcal{X}^n$

记 $\xi^{(n)}$ 是由 \mathcal{S}^n 决定的随机变量.

定义 3.1.8 (编码的平均误差和可达速率)

(1) 对于固定的信源 \mathcal{S}^n , 与编、译码函数 (f, g) , 它们的平均误差为

$$e_n(f, g) = P_r \{ \xi^{(n)} \neq g(f(\xi^{(n)})) \}$$

(2) 记 $C = \mathcal{U}_f^{(m)} = \{f(x^n) \mid x^n \in \mathcal{X}^n\}$ 为定长编码码字的集合, 称 $V_n = |\mathcal{U}_f^{(m)}|$ 为编码的信号体积, 而称

$$R_n = \frac{1}{n} \log(v_n) = \frac{1}{n} \log |\mathcal{U}_f^{(m)}|$$

为编码 f 的码率.



注:

- (1) 记 $\mathcal{U} = \{u_1, u_2, \dots, u_k\}$, 码字长度为 m . 对于 $u_i \in \mathcal{U}$, u_i 能携带的最大信息量为 $\log_2 k$.
- (2) 则 m 长码字所提供的最大信息量为 $m \log_2 k$.

定义 3.1.9 (信源序列编码的可达速率)

称 R 是信源序列 \mathcal{S}^n 的一个可达速率, $n = 1, 2, \dots$, 如果存在一个数列 $\varepsilon_n \rightarrow 0$, 当 $n \rightarrow \infty$ 时, 存在一组编码序列 $(f^{(n)}, g^{(n)})$ 使得以下条件成立.

(1) 对任何 $n = 1, 2, 3, \dots$, $e(f^{(n)}, g^{(n)}) \leq \varepsilon_n \rightarrow 0$

(2) 对任何 $n = 1, 2, 3, \dots$, $R_n \leq R(1 + \varepsilon_n) \rightarrow R$. 其中 $R_n = \frac{1}{n} \log M_n$, $M_n = |\mathcal{U}_f^{(n)}|$.



可达速率是指对每组编码序列 $(f^{(n)}, g^{(n)})$, 在误差范围内传输的平均信息量, 故定长编码问题是考虑最小可达速率, 即在误差范围内传输的平均信息量最小值.

定义 3.1.10 (信源序列的最小可达速率和它的编码问题)

对已给的信源序列 \mathcal{S}^n , 全体可达速率的最小值, 称为该信源序列的最小可达速率, 信源序列的编码问题就是求它的最小可达速率. (注: 每取一组 $(f^{(n)}, g^{(n)})$ 就有一个可达速率, 去找所有最小值)



例题 3.1.2 \mathcal{S} 是信源, $\mathcal{X} = \{x_1, x_2, \dots, x_5\}$ 为信源字母表, 对应的概率分布为 $p_1 = 1 - \varepsilon$, $p_2 = p_3 = p_4 = p_5 = \frac{\varepsilon}{4}$, $\mathcal{U} = \{0, 1\}$, 利用定长编码, 则码字长度至少为 3. (否则最多有 4 个码字, 不能建立映射). 码字长度为 n 的二进制编码可以有 2^n 个不同的码字, $4 = 2^2 < 5 < 2^3 = 8$.

如果采用变长编码, 令

$$f(x_1) = 0, (f(x_2), f(x_3), f(x_4), f(x_5)) = (100, 101, 110, 111)$$

它的平均码长为

$$L(\mathcal{S}, f) = (1 - \varepsilon) + 4 \times \frac{\varepsilon}{4} \times 3 = (1 - \varepsilon) + 3\varepsilon = 1 + 2\varepsilon,$$

只要 $\varepsilon < 1$, 就会有 $L(\mathcal{S}, f) = 1 + 2\varepsilon < 3$. 如果 $\varepsilon < \frac{1}{2}$, 则有 $L(\mathcal{S}, f) = 1 + 2\varepsilon < 2$. 因此利用变长码可以大大压缩信源的编码长度.

f 是唯一可译码只需验证 f^* 是 1-1 映射 (显然).

3.2 前缀码和即时码

变长码的讨论 $\left\{ \begin{array}{l} (1) \text{ 唯一可译变长码的构造} \\ (2) \text{ 平均码长的估计与优化} \end{array} \right.$

3.2.1 唯一可译变长码的构造

例题 3.2.1 考虑一个信源字母表 $\mathcal{X} = \{a, b, c, d\}$, 它的编码函数为 $f(a) = 0$, $f(b) = 01$, $f(c) = 011$, $f(d) = 0111$, 则码元集 $C = \{0, 01, 011, 0111\}$, 则码 C 是唯一可译码. 事实上, 假设我们收到的码字符串为 01101001110010. 注意到每个码字都是以 0 开始的, 故按此规则将码字符串进行分组, 可得到信源字母串为 cbadaba.

再如收到码字符串为 0100111011, 则可知信源字母串为 badc. 码 C 为唯一可译码, 但注意到我们收到的码字符串中有码元 011, 我们还不能确定它的信源字母, 因为在 011 后面可能出现 0 或 1, 如果是 0, 那么就可把 011 译出为 c , 如果是 1, 那么就可把 0111 译出为 d . 即译码不是即时的, 依赖于它后面的字符, 故称这种码为**非即时码**.

例题 3.2.2 考虑信源字母表 $\mathcal{X} = \{a, b, c, d\}$, 编码函数为 f :

$$f(a) = 1, f(b) = 01, f(c) = 001, f(d) = 0001,$$

它的码元集为

$$C = \{1, 01, 001, 0001\}$$

它把 1 作为两个码字的分隔号 (以 1 为结尾), 收到码字串 10010110001101 时, 译为信源字母串 acbadab; 又如收到码字串 0010001101, 译为信源字母串 cdab, 这种码可直接译码, 不用考虑码字符后面出现什么数字, 当我们收到一个字符串时, 就可从左向右读, 只要码元一出现, 我们就可译出相应的信源字母, 这种码我们称为**即时码**.

定义 3.2.1 (即时码与非即时码)

如果在任意的码串中, 从左到右, 只要一个码字出现, 就可唯一译出这个码字所对应的信源字母, 则称这种码为即时码, 否则为非即时码.



注:

- (1) 即时码和非即时码均是对唯一可译码而言.
- (2) 唯一可译码不一定是即时码, 如第一个例子.

定义 3.2.2 (前缀码)

设有两个字符串 $a^{(k)} = (a_1, a_2, \dots, a_k)$, $b^{(k')} = (b_1, b_2, \dots, b_{k'})$, 如果 $k \leq k'$, 且有 $(a_1, a_2, \dots, a_k) = (b_1, b_2, \dots, b_k)$, 则称 $a^{(k)}$ 是 $b^{(k')}$ 的前缀. 如果码元集 C 中任何一个码字都不能是另一个码字的前缀, 即在码元集 C 中, 任何一个码元 c_i 都不能是另一个码元 $c_j (i \neq j)$ 的前缀, 就称码 C 为前缀码.



如上面第二个例.

$$f(a) = 1, f(b) = 01, f(c) = 001, f(d) = 0001,$$

每个码字都不是另一个码字的前缀. 但第一个例子中. $f(a) = 0, f(b) = 01, f(c) = 011, f(d) = 0111$, 就不是前缀码. 码元集 $C = \{0, 01, 011, 0111\}$.

如何判断一个码元集是否具有前缀性? 只要把每个码元按照它们的长度由小到大排列, 把每个码元与它后面的码元进行比较, 看它与后面的码元的前边部分是否相同即可. 只要能找到一个码元能与它后面的码元的前边部分相同, 那么这种码就不是前缀码, 否则就是前缀码.

定理 3.2.1 (即时码与前缀码的关系)

前缀码一定是即时码, 反之亦然, 即即时码也是前缀码.



证明 \Leftarrow : 假设 C 是一个码元集, 若 C 不是前缀码, 则存在码字 c_i, c_j , 使得 c_i 是 c_j 的前缀, 在一个含有 c_i 的码字串中, 从左到右, 当 c_i 出现时, 只有当 c_i 后面出现部分, 连同 c_i 不是 c_j 时才能把 c_i 还原; 若 c_i 以及连同后面部分是 c_j 时, 不能把 c_i 还原, 应该把 c_j 还原, 因此 C 不是即时码, 矛盾. 故即时码一定为前缀码.

\Rightarrow : 若 C 不是即时码, 则从左到右, 出现一个码字 c_i , 还原为消息字母时, 依赖于后面的字符串, 即存在另一个码字 c_j , 使得 c_i 是 c_j 的前缀, 从而 C 不是前缀码.

3.2.2 Kraft 不等式

$\mathcal{S} = \{\mathcal{X}, p(x)\}$ 是信源, \mathcal{U} 输入信号字母集, $f: \mathcal{X} \rightarrow \mathcal{U}^*$, 考虑即时码存在的条件.

定理 3.2.2 (Kraft 不等式)

如果 f 是一个变长码, 它的码元集为 C , 它的码字长度分别为 $\{\ell_1, \ell_2, \dots, \ell_a\}$, 记 $r = |\mathcal{U}|$, 如果 f 是一个即时码, 那么它必满足 Kraft 不等式

$$\sum_{k=1}^a \frac{1}{r^{\ell_k}} \leq 1.$$

反之, 如果有一组数 $\{\ell_1, \ell_2, \dots, \ell_a\}$ 满足 Kraft 不等式, 那么必存在一个码长为 $\{\ell_1, \ell_2, \dots, \ell_a\}$ 的即时码.



注: 反之的含义并不是若 $C = \{c_1, c_2, \dots, c_a\}$, ℓ_1, \dots, ℓ_a 满足 $\sum_{k=1}^a \frac{1}{r^{\ell_k}} \leq 1$, 则 C 一定是即时码.

证明 记 $C = \mathcal{U}_f = \{u_1^{(\ell_1)}, u_2^{(\ell_2)}, \dots, u_a^{(\ell_a)}\}$ 是一个即时码的码元集, 其中它的码字分别为

$$u_i^{(\ell_i)} = \{u_{i1}, u_{i2}, \dots, u_{i\ell_i}\}, i = 1, 2, \dots, a$$

令 $\ell = \max \{\ell_i \mid i = 1, \dots, a\}$, 对 $\forall j = 1, 2, \dots, a$, 记

$$\mathcal{U}_j = \left\{ \left(u_j^{(\ell_j)}, z^{(\ell-\ell_j)} \right) \mid z^{(\ell-\ell_j)} \in \mathcal{U}^{(\ell-\ell_j)} \right\} \subseteq \mathcal{U}^{(\ell)}$$

(因 $u_j^{(\ell_j)}$ 固定, $z^{(\ell-\ell_j)} = (*, *, \dots, *)$, $* \in \mathcal{U}$, $|\mathcal{U}| = r$. 每个位置 r 种取法) 有 $|\mathcal{U}_j| = r^{\ell-\ell_j}$, 且 $i \neq j$ 时, $\mathcal{U}_i \neq \mathcal{U}_j$, $\mathcal{U}_i \cap \mathcal{U}_j = \emptyset$

事实上, 若 $\left(u_i^{(\ell_i)}, z_1^{(\ell-\ell_i)} \right) = \left(u_j^{(\ell_j)}, z_2^{(\ell-\ell_j)} \right)$, 不妨设 $\ell_i \leq \ell_j$, 则 $u_i^{(\ell_i)}$ 是 $u_j^{(\ell_j)}$ 的前缀, 与 f 是即时码矛盾. 由 $\mathcal{U}_j \subseteq \mathcal{U}^{(\ell)}$ 知则有 $\bigcup_{j=1}^a \mathcal{U}_j \subseteq \mathcal{U}^{(\ell)}$

$$\left| \bigcup_{j=1}^a \mathcal{U}_j \right| \leq |\mathcal{U}^{(\ell)}|$$

$$\left| \bigcup_{j=1}^a \mathcal{U}_j \right| = \sum_{j=1}^a |\mathcal{U}_j| = \sum_{j=1}^a r^{\ell-\ell_j}, \quad |\mathcal{U}^{(\ell)}| = r^\ell$$

即

$$\sum_{j=1}^a r^{\ell-\ell_j} \leq r^\ell \Rightarrow \sum_{j=1}^a \frac{1}{r^{\ell_j}} \leq 1$$

反之, 假设 $\ell_1, \ell_2, \dots, \ell_a$ 和 r 满足 Kraft 不等式. 下面证明存在即时码 f , 使它的码字长度为 $\ell_1, \ell_2, \dots, \ell_a$. (下面通过例子来说明构造方法).

例题 3.2.3 令 $\mathcal{U} = \{0, 1, 2\}$, 且 $\ell_1 = \ell_2 = 1, \ell_3 = 2, \ell_4 = \ell_5 = 4, \ell_6 = 5$, 是否可构造出具有上述

码字长度的即时码, 若有, 构造出一个这样的码.

$$\begin{aligned}\sum_{i=1}^6 \frac{1}{r^{\ell_i}} &= \frac{1}{3} + \frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^4} + \frac{1}{3^4} + \frac{1}{3^5} \\ &= \frac{3^4 + 3^4 + 3^3 + 3 + 3 + 1}{3^5} = \frac{196}{243} < 1\end{aligned}$$

满足 Kraft 不等式, 故这样的即时码存在. 构造如下: 设 α_i 为码字长度为 i 的码字个数.

1. 挑选码字长度最小的码字. $\ell_1 = \ell_2 = 1, \alpha_1 = 2$ 所以取 $u_{1,1} = 0, u_{1,2} = 1$ ($u_{i,j}$ 长度为 i 的第 j 个码字)
2. $\ell_3 = 2, \alpha_2 = 1$, 取长度为 2 的码字不能以 0,1 开头 (为保证是前缀码), 故取 $(u_{2,1,1}, u_{2,1,2}) = (2, 0)$. 长度为 2 的码字的第 1 个分量不能是码字 $u_{1,1}, u_{1,2}$.
3. $\ell_4 = \ell_5 = 4$, 这两个码字不能以 0,1 开头, 只能以 2 开头. 同时它的前两个分量不能取 $(2, 0)$, 于是可以取

$$\begin{aligned}(u_{4,1,1}, u_{4,1,2}, u_{4,1,3}, u_{4,1,4}) &= (2, 1, 0, 0), \\ (u_{4,2,1}, u_{4,2,2}, u_{4,2,3}, u_{4,2,4}) &= (2, 1, 0, 1).\end{aligned}$$

4. 最后再挑选长度为 5 的码字

$$(u_{5,1,1,1}, u_{5,1,1,2}, u_{5,1,1,3}, u_{5,1,1,4}, u_{5,1,1,5}) = (2, 1, 1, 0, 0).$$

最后我们构造出满足即时性的码为

$$\mathcal{U}_f = \{0, 1, 20, 2100, 2101, 21100\}$$

注:

(1) 若 $\ell_1, \ell_2, \dots, \ell_a$ 满足 Kraft 不等式, 则必存在码字长度为 $\ell_1, \ell_2, \dots, \ell_a$ 的即时码. 如果一个码的码字长度满足 Kraft 不等式, 但它不一定是即时码.

如: 考虑二源码 $C = \{0, 11, 100, 110\}$, 码字长度分别为 1, 2, 3, 3, 因为 $|\mathcal{U}| = 2$, 我们有

$$\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^3} = 1,$$

所以, 它的码字长度满足 kraft 不等式. 但这个码并不是即时的 (不是前缀码), 因为码字 11 是码字 110 的前缀. 但根据 1, 2, 3, 3 可构造一个即时码, 如 $\{0, 10, 110, 111\}$ 或 $\{1, 01, 001, 000\}$.

(2) 不满足 Kraft 不等式的码字长度为 $\ell_1, \ell_2, \dots, \ell_a$ 的即时码一定不存在.

3.3 信源变长码的编码定理

这一节讨论最优变长码平均码长的上、下界估计问题.

3.3.1 最优变长码平均码长的下界估计

记 $\mathcal{S} = \{\mathcal{X}, p(x)\}$ 为信源, 其中 $\mathcal{X} = \{x_1, x_2, \dots, x_a\}$ 是信源字母表, 对应的概率分布为 $P = (p_1, p_2, \dots, p_a)$. 它的编码方案为 (C, f) , 其中 C 是一个码元集, 简记为 $C = \{c_1, c_2, \dots, c_a\}$,

码字长度分别是 $\{\ell_1, \ell_2, \dots, \ell_a\}$, 这时平均码长为

$$L(\mathcal{S}, f) = \sum_{i=1}^a p_i \ell(f(x_i)) = \sum_{i=1}^a p_i \ell_i$$

定理 3.3.1

如果 $\mathcal{S} = \{\mathcal{X}, p(x)\}$ 是给定信源, f 是即时码, 那么

$$H_r(p_1, \dots, p_a) \leq L(\mathcal{S}, f),$$

其中 $H_r(\cdot)$ 是取对数 r 为底的熵函数, 而等号成立的条件是 $\ell_i = -\log_r p_i$, $r = |\mathcal{U}|$, \mathcal{U} 为信号字母表.



证明 因为 f 是即时码, 根据 Kraft 不等式, 我们有

$$q_0 = \sum_{i=1}^a \frac{1}{r^{\ell_i}} \leq 1$$

令 $q_i = \frac{1}{(q_0 r^{\ell_i})}$, 则有 $q_i \geq 0, i = 1, \dots, a$,

$$\sum_{i=1}^a q_i = \sum_{i=1}^a \frac{1}{\left(\sum_{k=1}^a \frac{1}{r^{\ell_k}}\right) r^{\ell_i}} = \frac{\sum_{i=1}^a \frac{1}{r^{\ell_i}}}{\sum_{k=1}^a \frac{1}{r^{\ell_k}}} = 1$$

结合引理 2.3.2 则有

$$\begin{aligned} H_r(p_1, \dots, p_a) &= \sum_{i=1}^a p_i \log_r \frac{1}{p_i} \leq \sum_{i=1}^a p_i \log_r \frac{1}{q_i} \quad (p_i = q_i \text{ 等号成立}) \\ &= \sum_{i=1}^a p_i \log_r (q_0 r^{\ell_i}) \\ &= \sum_{i=1}^a p_i \ell_i + \log_r q_0 \quad (q_0 \leq 1) \\ &\leq \sum_{i=1}^a p_i \ell_i = L(\mathcal{S}, f), \end{aligned}$$

等号成立的充要条件为 $p_i = q_i, q_0 = 1$. 即 $p_i = \frac{1}{r^{\ell_i}}$ 或 $\ell_i = -\log_r p_i$.

注: 等号成立的条件为 $\ell_i = -\log_r p_i$. 即要求 $\ell_i = -\log_r p_i$ 必须是个整数, 而这个条件并不是总能满足, 所以定理中的等号一般不成立.

3.3.2 最优变长码平均码长的上界估计

定理 3.3.2

对于已给信源 $\mathcal{S} = \{\mathcal{X}, p(x)\}$, 它的 r 元最优变长即时码 f_0 有

$$L(\mathcal{S}, f_0) < H_r(p_1, \dots, p_a) + 1$$



证明 记 $\text{Int}(z)$ 是 z 的整数部分, 而

$$\text{Int}_+(z) = \begin{cases} z, & \text{如果 } z \text{ 是整数} \\ \text{Int}(z) + 1, & \text{如果 } z \text{ 不是整数} \end{cases}$$

如果 $\bar{p} = (p_1, p_2, \dots, p_a)$ 是固定信源的概率分布, 那么我们取 $\ell_i = \text{Int}_+(-\log_r p_i)$ 大于或等于 $\log_r \frac{1}{p_i}$ 的最小正整数. 那么有 $\log_r \frac{1}{p_i} \leq \ell_i < \log_r \frac{1}{p_i} + 1$, 且由 $\log_r \frac{1}{p_i} \leq \ell_i$ 可得 $\frac{1}{p_i} \leq r^{\ell_i}$, 因此有 $\frac{1}{r^{\ell_i}} \leq p_i$ 成立

从而 $\sum_{i=1}^a \frac{1}{r^{\ell_i}} \leq \sum_{i=1}^a p_i = 1$. 即 Kraft 不等式成立. 因此存在码长为 $\{\ell_1, \ell_2, \dots, \ell_a\}$ 的即时码. 这时它的平均码长为

$$\begin{aligned} L(\mathcal{S}, f) &= \sum_{i=1}^a p_i \ell_i \\ &< \sum_{i=1}^a p_i \left(\log_r \frac{1}{p_i} + 1 \right) \\ &= \sum_{i=1}^a p_i \log_r \frac{1}{p_i} + \sum_{i=1}^a p_i \\ &= H_r(p_1, \dots, p_a) + 1 \end{aligned}$$

于是我们找到了一个平均码字长度小于 $H_r(p_1, \dots, p_a) + 1$ 的即时编码方案.

定理 3.3.3

定理: 对于已给信源 $\mathcal{S} = \{\mathcal{X}, p(x)\}$, 它的 r 元最优变长即时码 f_0 有

$$H_r(p_1, \dots, p_a) \leq L(\mathcal{S}, f_0) < H_r(p_1, \dots, p_a) + 1$$

成立.



3.4 Huffman 信源编码算法

3.4.1 Huffman 编码的实例分析

变长最优码就是平均码字长度取最小值的即时码. 本节我们给出一种有效地构造最优编码方案的算法——Huffman 编码.

例题 3.4.1 已知信源概率分布为 $\bar{p} = \{0.24, 0.20, 0.18, 0.13, 0.10, 0.06, 0.05, 0.03, 0.01\}$, 如取信号字母表 $\mathcal{U} = \{0, 1, 2, 3\}$, 求信源的 Huffman 编码.

Huffman 编码方案的主要运算步骤是构造 Huffman 数据压缩表与 Huffman 编码表. 它的运算步骤如下.

1. 先构造 Huffman 数据压缩表

(1) 首先把概率分布 \bar{p} 的概率按大小降序排列, 并作为把它们作为 Huffman 数据压缩表第 1 列, 该列长度为 $a = 9$.

(2) 把第一列的 3 个最小的概率相加, 得到新的一列概率分布, 重新按降序排列, 成为 Huffman

数据压缩表中的第三列. 这时第三列的长度为 7, 相加后的数为 0.09, 我们用方框标出.

(3) 把第三列的 4 个最小的概率相加, 得到新的一列概率分布, 重新按降序排列, 成为 Huffman 数据压缩表中的第五列. 这时第五列的长度为 4, 相加后的数为 0.38, 我们用方框标出.

2. Huffman 数据压缩表构造 Huffman 编码表

(1) 因 Huffman 数据压缩表的第五列的概率分布只有 4 个行, 因此它们的编码正好是 0, 1, 2, 3. 把这 4 个数填入 Huffman 编码表的第六列. 因此, 第六列是一个码长为 1 的编码.

(2) 在第五列中, 带方框的概率为 0.38, 它对应的第六列的编码为 0, 而 0.38 是由第三列的 0.13, 0.10, 0.09, 0.06 这 4 个数相加而成. 这样我们构造第三列概率分布的编码为: 第三列的 0.13, 0.10, 0.09, 0.06 这 4 个数的编码是在 0.38 的编码 0 后边延长 1 个数, 它们分别为 00, 01, 02, 03. 第三列中 0.24, 0.20, 0.18 这 3 个数的编码与第五列的编码相同, 仍为 1, 2, 3. 把 0.24, 0.20, 0.18, 0.13, 0.10, 0.09, 0.06 这 7 个数的编码 1, 2, 3, 00, 01, 02, 03 列入 Huffman 编码表的第四列.

(3) 在表的第三列中, 带方框的概率为 0.09, 它对应的第四列的编码为 02, 而 0.09 是由第一列的 0.05, 0.03, 0.01 这 3 个数相加而成. 这样我们构造第一列概率分布的编码为第一列的前 6 个数的编码与第三列所对应的编码相同. 第一列中 0.05, 0.03, 0.01 这 3 个数的编码是在第 3 列的 0.09 的编码 02 后边延长 1 个数, 因此它们分别为 020, 021, 022. 把第一列 9 个概率的编码列入 Huffman 编码表的第二列. 最终完成 Huffman 编码表. 各项计算结果见表.

以上过程为 Huffman 编码方案. 主要原则是概率大的信源字母对应长度短的码字, 概率小的信源字母对应长度大的码字以保证平均码长尽量小. 频率高的元素使用较短的编码, 频率低的元素使用较长的编码. 这样做可以减少整体编码的平均长度, 从而达到压缩数据的目的.

概率	码	概率	码	概率	码
0.24	1	0.24	1	0.38	0
0.20	2	0.20	2	0.24	1
0.18	3	0.18	3	0.20	2
0.13	00	0.13	00	0.18	3
0.10	01	0.10	01		
0.06	03	0.09	02		
0.05	020	0.06	03		
0.03	021				
0.01	022				

3.4.2 Huffman 编码的一般算法

考虑一般情形. 令码字母表为 $\mathcal{U} = \{0, 1, \dots, r-1\}$. 信源概率分布为 $\bar{p} = (p_1, p_2, \dots, p_a)$. 构造 Huffman 编码的步骤:

1. 若 $a \leq r$, 则取 $f(x_i) = i-1$ 即可, $i = 1, \dots, a$

2. 对于 $a > r$, 我们观察上述例子的编码过程, 知编码表的最后一列序含 r 个元素. 设编码表为 $2k$ 列, 则有第 $2k$ 列序含 r 个元, 第 $2k-2$ 列序含 $2r-1$ 个元, $2k-4$ 列序含 $3r-2$ 个元,

即是一个以 $r-1$ 为公差的等差数列, 故编码表中的第 4 列序含 $(k-1)r - (k-2)$ 个元. 第 2 列序应含 $kr - (k-1)$ 个元 (确定 k 的值) 因此:

(1) 确定 k 的值. $kr - k + 1 \geq a$, 故 $k \geq \frac{a-1}{r-1}$, k 取 $\geq \frac{a-1}{r-1}$ 的最小正整数. 故令 $k = \text{Int}_+ \left(\frac{a-1}{r-1} \right)$.

(2) 确定第 1 列中应把最后的几个分量相加, 再按大小顺序排列, 应该为 $a - [(k-1)r - (k-2)] + 1$ 个分量相加.

(3) 从第 3 列开始每次将最后 r 个分量相加, 并按大小排序放入下一奇数列, 并将相加所得的概率用框标出. 于是得到了 Huffman 压缩表.

(4) 按例中编码规则去写出编码表即可.

上述算法称为 Huffman 编码算法, 得到的编码 $C = \{c_1, c_2, \dots, c_a\}$ 称为 Huffman 码.

例题 3.4.2 设 $\mathcal{S} = (\mathcal{X}, p(x))$, $\bar{p} = \{0.3, 0.1, 0.1, 0.1, 0.1, 0.06, 0.05, 0.05, 0.05, 0.04, 0.03, 0.02\}$, 试构造 \mathcal{S} 上的 3 元 Huffman 码并求平均码长.

解: (1) 先确定 k 的值

$$k = \text{Int}_+ \left(\frac{a-1}{r-1} \right) = \frac{11}{2} = 6$$

(2) 再确定第 1 列最后几个分量相加

$$\begin{aligned} & a - (k-1)r + k - 1 \\ &= 12 - (6-1) \times 3 + 6 - 1 \\ &= 12 - 15 + 6 - 1 = 2 \end{aligned}$$

于是我们构造 Huffman 编码为

概率	码	概率	码	概率	码	概率	码	概率	码	概率	码
0.3	1	0.3	1	0.3	1	0.3	1	0.3	1	0.4	0
0.1	02	0.1	02	0.14	01	0.16	00	0.3	2	0.3	1
0.1	20	0.1	20	0.1	02	0.14	01	0.16	00	0.3	2
0.1	21	0.1	21	0.1	20	0.1	02	0.14	01		
0.1	22	0.1	22	0.1	21	0.1	20	0.1	02		
0.06	000	0.06	000	0.1	22	0.1	21				
0.05	001	0.05	001	0.06	000	0.1	22				
0.05	002	0.05	002	0.05	001						
0.05	010	0.05	010	0.05	002						
0.04	012	0.05	011								
0.03	0110	0.04	012								
0.02	0111										

$$L(\mathcal{S}, f) = 0.3 \times 1 + 0.4 \times 2 + 0.25 \times 3 + 0.05 \times 4 = 2.05.$$

例题 3.4.3 $\bar{p} = (0.2, 0.1, 0.1, 0.3, 0.1, 0.2)$, $r = 3$, 求 Huffman 编码

解: (1) $k = \text{Int}_+ \left(\frac{a-1}{r-1} \right) = \text{Int}_+ \left(\frac{6-1}{3-1} \right) = 3$

(2) 确定第 1 列后几个分量相加 (最后分量相加的个数): $a - (k-1)r + k - 1 = 6 - (3-1) \times 3 + 3 - 1 = 2$. 于是有

概率	码	概率	码	概率	码
0.3	1	0.3	1	0.5	0
0.2	2	0.2	2	0.3	1
0.2	00	0.2	00	0.2	2
0.1	02	0.2	01		
0.1	010	0.1	02		
0.1	011				

3.5 Huffman 编码性能分析

本节对由 Huffman 信源编码算法所产生的 Huffman 码的性能进行分析。

3.5.1 Huffman 编码的前缀性

定理 3.5.1 (Huffman 码的前缀性)

由 Huffman 编码算法得到的 Huffman 码是前缀码。



证明 由 Huffman 编码算法中的各步骤可知, 第 $2k$ 列中各码元各不相同, 且第 $2k-2$ 列中各码元与第 $2k$ 列中码元相同或是某个码元的延伸, 因此第 $2k-2$ 列中各码元互不相同, 且每个码元不能成为另一码元的前缀。

一般情形, 如果第 $2j$ 列中各码元各不相同, 且第 $2j-2$ 列中各码元与第 $2j$ 列中码元相同或是某个码元的延伸, 那么第 $2j-2$ 列中各码元互不相同, 且每个码元不能成为另一码元的前缀。

由此递推, 最后得到第一列中各码元互不能成为另一码元的前缀。由此定理得证。

3.5.2 Huffman 编码的最优性

定理 3.5.2 (Huffman 码的最优性定理)

由 Huffman 算法构造的 Huffman 码是最优码, 即对固定信源 \mathcal{S} , 如果我们记 f_0 和 f 分别是 Huffman 码与任一前缀码, 那么必有

$$L(\mathcal{S}, f_0) \leq L(\mathcal{S}, f)$$



为证明这个定理我们先做以下讨论。设 $\mathcal{S}, \mathcal{S}'$ 是两个信源, 我们分别记之为

$$\mathcal{S} = (\mathcal{X}, \bar{p}), \quad \mathcal{S}' = (\mathcal{X}', \bar{p}'),$$

其中

$$\begin{aligned} \mathcal{X} &= \{x_1, x_2, \dots, x_a\}, \quad \bar{p} = (p_1, p_2, \dots, p_a), \\ \mathcal{X}' &= \{x'_1, x'_2, \dots, x'_a\}, \quad \bar{p}' = (p'_1, p'_2, \dots, p'_{a'}), \end{aligned}$$

且

$$p_1 \geq p_2 \geq \cdots \geq p_a, \quad p'_1 \geq p'_2 \geq \cdots \geq p'_{a'},$$

引理 3.5.1

如果 $C = \{c_1, c_2, \dots, c_a\}$ 是 \mathcal{S} 的最优前缀码, 那么必有

$$\ell(c_1) \leq \ell(c_2) \leq \cdots \leq \ell(c_a)$$

成立



证明 由最优码的定义, 若 f_0 是信源 \mathcal{S} 的最优码, 则

$$L(\mathcal{S}, f_0) = \sum_{k=1}^a p_k \ell(c_k)$$

最小. 若存在 $j < i$ 使得 $\ell(c_j) > \ell(c_i)$, 此时

$$\begin{aligned} L(\mathcal{S}, f_0) &= \sum_{k=1}^a p_k \ell(c_k) \\ &> p_1 \ell(c_1) + \cdots + p_{j-1} \ell(c_{j-1}) + p_j \ell(c_i) + p_{j+1} \ell(c_{j+1}) \\ &\quad + \cdots + p_i \ell(c_j) + \cdots + p_a \ell(c_a) \end{aligned}$$

事实上

$$p_j \ell(c_j) + p_i \ell(c_i) - p_j \ell(c_i) - p_i \ell(c_j) = (p_j - p_i)(\ell(c_j) - \ell(c_i)) > 0$$

从而 f_0 不是最优码.

定义 3.5.1

设 $\mathcal{S}, \mathcal{S}'$ 是两个信源, 如上所记. 我们有以下定义.

(1) 称 \mathcal{S}' 是 \mathcal{S} 的 r Huffman 扩张信源, 如果下面两个条件成立.

(i) $a < a' \leq a + r - 1$.

(ii) 对 p 和 p' , 存在一个正数 $1 \leq s \leq a$, 满足

$$\begin{cases} p_j = p'_j, & \text{当 } j < s \text{ 时,} \\ p_j = p'_{j-1}, & \text{当 } s < j < a \text{ 时,} \\ p_j = \sum_{j=1}^{a'-a+1} p_{a+1} & \text{当 } j = s \text{ 时,} \end{cases}$$

(2) 如果

$$C = \{c_1, c_2, \dots, c_a\}, \quad C' = \{c'_1, c'_2, \dots, c'_{a'}\}$$

分别是 \mathcal{S} 和 \mathcal{S}' 的编码, 称 C' 是 C 的 r Huffman 扩张编码, 如果以下两个条件成立.


(i) \mathcal{S}' 是 \mathcal{S} 的 r Huffman 扩张信源.

(ii) 码元集合 C' 和 C 满足

$$\begin{cases} c'_j = c_j, & \text{当 } j < s \text{ 时,} \\ c'_j = c_{j+1}, & \text{当 } s < j < a \text{ 时,} \\ c'_{a+i-1} = (c_s, i-1) & \text{当 } i = 1, 2, \dots, a' - a + 1 \text{ 时.} \end{cases}$$



引理 3.5.2

如果 C' 与 C 分别是 \mathcal{S}' 与 \mathcal{S} 的编码, \mathcal{S} 中的消息个数 $a = kr - k + 1$, 且 C' 是 C 的 r Huffman 扩张编码, 那么当 C 是 \mathcal{S} 的最优前缀码时, C' 一定是 \mathcal{S}' 最优前缀码. 

Huffman 编码的最优性定理的证明: 该定理的证明由 Huffman 编码表的定义与第二个引理即得, 因为有以下结论成立.

(1) 在 Huffman 编码表中, 如果记它的第 $2j - 1$ 列

$$\bar{p}_j = (p_{ji}, p_{j2}, \dots, p_{jt_j})$$

为信源 \mathcal{S}_j , 而记它的第 $2j$ 列

$$C_j = (c_{ji}, c_{j2}, \dots, c_{jt_j})$$

为信源 \mathcal{S}_j 的一个编码, 那么 \mathcal{S}_j 是 \mathcal{S}_{j+1} 的 Huffman 扩张信源, C_j 是 C_{j+1} Huffman 扩张 (见定义).

(2) 在 Huffman 扩张编码表的最后两列第 $2k - 1, 2k$ 列中, 因为码长 $\ell(c_{ki}) = 1$, 所以一定是最优前缀码.

(3) 由上面引理的递推法可得在 Huffman 编码表的 $2j - 1, 2j$ 列中, C_j 一定是 \mathcal{S}_j 的最优前缀码. 因此 $C = C_1$ 一定是 $\mathcal{S} = \mathcal{S}_1$ 的最优前缀码.

例题 3.5.1 $\bar{p} = (0.32, 0.19, 0.19, 0.11, 0.10, 0.09)$, $\mathcal{U} = \{0, 1\}$

$$k = \text{Int}_+ \left(\frac{a-1}{r-1} \right) = \text{Int}_+ \left(\frac{6-1}{2-1} \right) = 5$$

$$a = (k-1)r + k - 1 = 6 - (5-1) \times 2 + 5 - 1 = 2$$

编码一:

概率	码	概率	码	概率	码	概率	码	概率	码
0.32	00	0.32	00	0.32	00	0.38	1	0.62	0
0.19	10	0.19	10	0.30	01	0.32	00	0.38	1
0.19	11	0.19	11	0.19	10	0.30	01		
0.11	011	0.19	010	0.19	11				
0.10	0100	0.11	011						
0.09	0101								

码字总长度 17,

$$L(\mathcal{S}, f_1) = 4 \times 0.19 + 3 \times 0.11 + 2 \times 0.7 = 2.49.$$

编码二:

概率	码	概率	码	概率	码	概率	码	概率	码
0.32	00	0.32	00	0.32	00	0.38	1	0.62	0
0.19	11	0.19	10	0.30	01	0.32	00	0.38	1
0.19	010	0.19	11	0.19	10	0.30	01		
0.11	011	0.19	010	0.19	11				
0.10	100	0.11	011						
0.09	101								

码字总长度 16, $L(\mathcal{S}, f_2) = 3 \times 0.49 + 2 \times 0.51 = 2.49$. 信源固定的最优码的平均码长的定值 2.49. 虽编码方法不同, 但平均码长相同.

上面两个表显示的是构造一个二元 Huffman 码的过程, 其中 $r = 2$. 对相等的概率可有不同的排列, 这时所得的 Huffman 码可能不同, 但它们都是最优码, 因此它们的平均码长相等.

对相等的概率可有不同的排列, 这时由 Huffman 算法生成的 Huffman 编码不同. 经过计算可知, 第一个码的总码字长度是 17, 而第二个码的总码字长度却是 16, 但它们的平均码字长度都是 2.49, 这是为前缀码平均码长的最小值.

3.6 信源定长码的编码定理

信源定长码的编码问题我们已在 3.1 节中给出, 现在讨论它的编码定理, 为了简单起见, 我们只讨论无记忆信源的情形. 记 \mathcal{S}^n 是一个由 \mathcal{S} 决定的无记忆信源. 现在讨论它的定长编码问题, 求它的最小可达速率. 为证明定长编码定理, 我们先给出以下引理.

引理 3.6.1

R 是 \mathcal{S}^n 可达速率的充分与必要条件是存在一列 \mathcal{X}^n 的子集 \mathcal{X}_1^n , 与一正数列 $\varepsilon_n \rightarrow 0$, 使

$$M_n = |\mathcal{X}_1^n| < 2^{nR(1+\varepsilon_n)}$$

且

$$p(\mathcal{X}_1^n) = \Pr\{\xi^n \in \mathcal{X}_1^n\} > 1 - \varepsilon_n$$

引理给出了可达速率的一个充分必要条件, 证明过程中给出了如何构造可达速率的定义中需要的编译码序列 $(f^{(n)}, g^{(n)})$. 利用该引理可证明下面的信源定长码的编码定理.

定理 3.6.1

设 \mathcal{S}^n 是一个由 \mathcal{S} 决定的无记忆信源, 即

$$p(x^{(n)}) = p(x_1)p(x_2)\cdots p(x_n), \text{ 对 } \forall x^{(n)} = (x_1, x_2, \cdots, x_n) \in \mathcal{X}^n,$$

那么它的最小可达速率为

$$R_0 = H(\xi) = H(p_1, p_2, \cdots, p_a),$$

其中 $H(\xi)$ 是 ξ 的熵.

该定理不证明, 掌握结论.

3.7 习题课

3.7.1 基本概念

1. 信源编码: $f: \mathcal{X}$ (或 \mathcal{X}^n) $\rightarrow \mathcal{U}^*$ 的映射称为一个信源编码.
2. 1-1 码: f 是一个信源编码. 若对 $\forall x, x' \in \mathcal{X}, x \neq x'$ 都有 $f(x) \neq f(x')$, 称为 f 为一个 1-1 码, 或称 f 具有 1-1 性.
3. 扩张编码 (变长和定长)
4. 唯一可译码: 若 f 的扩张编码 f^* 是 1-1 映射, 则称 f 是唯一可译码.
5. 即时码: 如果对于任意一个码字串, 从左到右. 如果码字 c_i 出现就译为其相应的消息字母, 就称这种码为即时码, 否则称为非即时码.
6. 前缀码: 设 $a^{(k)} = (a_1, a_2, \dots, a_k)$, $b^{(k')} = (b_1, b_2, \dots, b_{k'})$ 若 $k \leq k'$ 且有 $(a_1, \dots, a_k) = (b_1, \dots, b_k)$, 则称 $a^{(k)}$ 为 $b^{(k')}$ 的前缀. 若在一个码中, 任何一个码字都不是另一个码字的前缀, 即任意码字 c_i , 都不是 c_j 的前缀, 称这种码为前缀码.
7. 平均码长: $\mathcal{S} = (\mathcal{X}, p(x)), \bar{p} = \{p_1, p_2, \dots, p_a\}, \ell_f(x)$ 为码字 $f(x)$ 的长度, 称 $L(\mathcal{S}, f) = \sum_{i=1}^a p(x) \ell_f(x)$ 为 f 的平均码长.

3.7.2 编码问题与基本结论

1. 变长码的编码问题

信源固定, 构造变长即时码 f_0 , 使得对任意变长即时码 f 有

$$L(\mathcal{S}, f_0) \leq L(\mathcal{S}, f)$$

此时, 称 f_0 为最优变长即时码.

最优码构造方法: Huffman 编码算法.

2. 定长编码 (信源序列) 的编码问题

$\mathcal{S}^n = (\mathcal{X}^n, p^{(n)}(x^{(n)}))$, 求最小可达速率.

信源定长码的编码定理: 设 \mathcal{S}^n 是一个由 \mathcal{S} 决定的无记忆信源, 那么它的最小可达速率为

$$R_0 = H(\xi) = H(p_1, p_2, \dots, p_a),$$

其中 $H(\xi)$ 是 ξ 的熵.

3. Kraft 不等式 (即时码存在的必要条件)

$$\sum_{k=1}^a \frac{1}{r^{\ell_k}} \leq 1 \quad C = \{c_1, c_2, \dots, c_a\}$$

$\ell_1, \ell_2, \dots, \ell_a$ 为码字的长度, $r = |\mathcal{U}|$. 反之, 若 $\{\ell_1, \ell_2, \dots, \ell_a\}$ 满足上式, 则一定存在以 $\ell_1, \ell_2, \dots, \ell_a$ 为码字长度的即时码 (构造方法).

4. 满足 Kraft 不等式的码不一定是即时码, 反例需知道.

5. 最优变长码平均码长的上、下界

定理: 对于已给信源 $\mathcal{S} = \{\mathcal{X}, p(x)\}$, 它的 r 元最优变长即时码 f_0 有

$$H_r(p_1, \dots, p_a) \leq L(\mathcal{S}, f_0) < H_r(p_1, \dots, p_a) + 1$$

成立.

3.7.3 课后习题

 **练习 3.7.1** 下面的码是否是即时码? 是否是唯一可译码?

(1) $C = \{0, 10, 1100, 1101, 1110, 1111\}$.


(2) $C = \{0, 10, 110, 1110, 1011, 1101\}$.

解: (1) C 是前缀码, 故是即时码, 从而是唯一可译码.

(2) C 不是前缀码, 因为码字 10 是码字 1011 的前缀, 故 C 不是即时码. C 不是唯一可译码. 字符串

$$\begin{array}{cccccc} 0 & 10 & 110 & 1110 & 1011 & 1101 \\ \hline a & b & c & d & e & f \\ 0 & 1011 & 0 & 1110 & 1011 & 1101 \\ \hline a & e & a & d & e & f \end{array}$$

同一字符串的还原消息为两个, 不是唯一可译码.

 **练习 3.7.2** 判断是否存在即时码具有以下的基数和码字长度, 如果有, 试构造出一个这样的码.

(1) $r = 2$, 长度: 1, 3, 3, 3, 4, 4.

(2) $r = 3$, 长度: 1, 1, 2, 2, 3, 3, 3.

(3) $r = 5$, 长度: 1, 1, 1, 1, 1, 8, 9.

解: (1) $\frac{1}{2} + 3 \times \frac{1}{2^3} + 2 \times \frac{1}{2^4} = 1$, 故满足 Kraft 不等式. 即时码存在.

$$\begin{array}{ll} u_{1,1} = 0 & 0 \\ u_{3,1,1} = 1 & u_{3,1,2} = 0 \quad u_{3,1,3} = 0 \quad (1, 0, 0) \\ u_{3,2,1} = 1 & u_{3,2,2} = 0 \quad u_{3,2,3} = 1 \quad (1, 0, 1) \\ u_{3,3,1} = 1 & u_{3,3,2} = 1 \quad u_{3,3,3} = 0 \quad (1, 1, 0) \\ u_{4,1,1} = 1 & u_{4,1,2} = 1 \quad u_{4,1,3} = 1 \quad u_{4,1,4} = 0 \quad (1, 1, 1, 0) \\ u_{4,2,1} = 1 & u_{4,2,2} = 1 \quad u_{4,2,3} = 1 \quad u_{4,2,4} = 1 \quad (1, 1, 1, 1) \end{array}$$

故此即时码为

$$\{0, 100, 101, 110, 1110, 1111\}$$

(2) $2 \times \frac{1}{3} + 2 \times \frac{1}{3^2} + 3 \times \frac{1}{3^3} = 1$ 故满足 Kraft 不等式, 即时码存在.

$$u_{1,1} = 0 \quad u_{1,2} = 1 \quad 0, 1$$

$$\begin{aligned}
u_{2,1,1} &= 2 & u_{2,1,2} &= 0 & (2, 0) \\
u_{2,2,1} &= 2 & u_{2,2,2} &= 1 & (2, 1) \\
u_{3,1,1} &= 2 & u_{3,1,2} &= 2 & u_{3,1,3} &= 0 & (2, 2, 0) \\
u_{3,2,1} &= 2 & u_{3,2,2} &= 2 & u_{3,2,3} &= 1 & (2, 2, 1) \\
u_{3,3,1} &= 2 & u_{3,3,2} &= 2 & u_{3,3,3} &= 2 & (2, 2, 2)
\end{aligned}$$

故此即时码为

$$\{0, 1, 20, 21, 220, 221, 222\}$$

(3) $5 \times \frac{1}{5} + \frac{1}{5^8} + \frac{1}{5^9} > 1$, 故这样的即时码不存在.


 **练习 3.7.3** 在证明 Kraft 不等式中, 我们说

$$\frac{\alpha_1}{r} + \frac{\alpha_2}{r^2} + \cdots + \frac{\alpha_n}{r^n} \leq 1$$

等价于 Kraft 不等式. 证明这个结果.

解: 令 α_j 表示 $\ell_i = j$ 的 i 的个数, 即长度为 j 的码字个数为 α_j 个, 故长度为 1 的个数为 α_1 个, 长度为 2 的个数为 α_2 个, \cdots , 长度为 n 的个数为 α_n 个.

$$\begin{aligned}
\sum_{i=1}^n \frac{1}{r^{\ell_i}} &= \alpha_1 \cdot \frac{1}{r^1} + \alpha_2 \cdot \frac{1}{r^2} + \cdots + \alpha_j \cdot \frac{1}{r^j} + \cdots + \alpha_n \cdot \frac{1}{r^n} \\
&= \sum_{i=1}^n \frac{\alpha_i}{r^i} \leq 1
\end{aligned}$$

 **练习 3.7.4** 令 C 是一个即时码, 试证明下列命题等价.

- (1) C 是最大即时码, 即没有码字能够添入 C 中而令 C 仍保持即时性.
- (2) 任意码元素的有限串都是某个码字串的前缀.
- (3) Kraft 不等式中的等号成立.


解: 只证明 (1) \iff (3).

\Leftarrow : 先证明若 C 不是最大即时码, 则 Kraft 不等式中等号不成立. 若 C 不是最大即时码, 则在码 C 中可至少添入一个码字, 成为一个新的即时码 C_1 , 假设 $C = \{c_1, c_2, \cdots, c_a\}$, c_i 的码长为 ℓ_i , $i = 1, \cdots, a$, 设添入的码字为 c_{a+1} , 长度为 ℓ_{a+1} , 因 C_1 仍为即时码, 故有

$$\sum_{i=1}^{a+1} \frac{1}{r^{\ell_i}} = \sum_{i=1}^a \frac{1}{r^{\ell_i}} + \frac{1}{r^{\ell_{a+1}}} \leq 1$$

从而有 $\sum_{i=1}^a \frac{1}{r^{\ell_i}} < 1$. 故 Kraft 不等式中等号不成立.

\Rightarrow : 若 Kraft 不等式中等号不成立, 令 $C = \{c_1, c_2, \cdots, c_a\}$, 对应码字长为 $\{\ell_1, \ell_2, \cdots, \ell_a\}$. 此时有 $\sum_{k=1}^a \frac{1}{r^{\ell_k}} < 1$. 令 $\ell = \lceil \log_r \left(1 - \sum_{k=1}^a \frac{1}{r^{\ell_k}} \right) \rceil + 1$, 则有 $\sum_{k=1}^a \frac{1}{r^{\ell_k}} + \frac{1}{r^{\ell}} \leq 1$ 故 $\{\ell_1, \ell_2, \cdots, \ell_a, \ell\}$ 满足 Kraft 不等式于是可构造即时码 $C' = \{c_1, c_2, \cdots, c_a, c'\}$, 与 C 是最大即时码矛盾.

 **练习 3.7.5** 对下面给定的概率分布和基数, 找出一个 Huffman 编码, 并求平均码长.

$$p = \{0.3, 0.1, 0.1, 0.1, 0.1, 0.06, 0.05, 0.05, 0.05, 0.04, 0.03, 0.02\}, \quad r = 2.$$

解: (1) 先确定 k 的值.

$$k = \text{Int}_+ \left(\frac{a-1}{r-1} \right) = \frac{11}{1} = 11.$$

(2) 再确定第 1 列最后几个分量相加.

$$\begin{aligned} & a - (k-1)r + k - 1 \\ &= 12 - (11-1) \times 2 + 11 - 1 = 2. \end{aligned}$$

于是我们构造 Huffman 编码为

概率 码	概率 码	概率 码	概率 码	概率 码	概率 码	概率 码
0.3 00	0.3 00	0.3 00	0.3 00	0.3 00	0.3 00	0.3 00
0.1 111	0.1 111	0.1 111	0.1 111	0.1 110	0.11 011	0.19 010
0.1 100	0.1 100	0.1 100	0.1 100	0.1 111	0.1 110	0.11 011
0.1 101	0.1 101	0.1 101	0.1 101	0.1 100	0.1 111	0.1 110
0.1 0100	0.1 0100	0.1 0100	0.1 0100	0.1 101	0.1 100	0.1 111
0.06 0110	0.06 0110	0.09 0101	0.1 0100	0.1 101	0.1 101	0.1 100
0.05 1100	0.05 0111	0.06 0110	0.09 0101	0.09 0101	0.1 0100	0.1 101
0.05 1101	0.05 1100	0.05 0111	0.06 0110	0.09 0101		
0.05 01010	0.05 1101	0.05 1100	0.05 0111			
0.04 01011	0.05 01010	0.05 1101				
0.03 01110	0.04 01011					
0.02 01111						
概率 码	概率 码	概率 码	概率 码	概率 码	概率 码	
0.3 00	0.3 00	0.3 00	0.3 00	0.4 1	0.6 0	
0.2 10	0.2 10	0.3 01	0.3 00	0.3 01	0.4 1	
0.19 010	0.2 11	0.2 10				
0.11 011	0.19 010	0.2 11				
0.1 110	0.11 011					
0.1 111						

$$L(\mathcal{S}, f) = 0.14 \times 5 + 0.26 \times 4 + 0.3 \times 3 + 0.3 \times 2 = 3.24$$

等价证明

练习题等价命题的证明:

从 (1) 到 (2): 假设 C 是最大即时码, 意味着不能向 C 添加更多的码字而保持其即时性. 我们需要证明, 任意码元素的有限串都是某个码字串的前缀.

由于 C 是最大即时码, 如果存在一个码元素的串, 它不是任何码字串的前缀, 那么我们可以将这个串作为一个新的码字添加到 C 中, 而不违反即时性. 这与 C 是最大即时码的假设矛盾. 因此, 任意码元素的有限串必须是某个码字串的前缀.

从 (2) 到 (3): 假设任意码元素的有限串都是某个码字串的前缀. 这意味着码字集覆盖了所有可能的码元组合, 形成了一种“完整”的编码系统, 没有未被利用的“空间”.

Kraft 不等式表达了对于长度可变的即时码, 其码字长度的集合 $\{l_1, l_2, \dots, l_n\}$ 必须满足以下条件:

$$\sum_{i=1}^n r^{-l_i} \leq 1$$

如果任意码元素的有限串都是某个码字串的前缀, 这意味着编码系统利用了所有可用的编码“空间”, 不留下任何“空隙”. 因此, **Kraft 不等式**中的等号必须成立, 否则还有“空间”可以添加更多的码字而不违反即时性, 这与假设矛盾.

从 (3) 到 (1): 假设 **Kraft 不等式**中的等号成立, 即

$$\sum_{i=1}^n r^{-l_i} = 1$$

这表明编码系统完美地匹配了编码空间的容量, 没有未被利用的部分. 在这种情况下, 不能添加更多的码字而不增加现有码字的长度, 因为这将违反 **Kraft 不等式**, 导致总和超过 1. 因此, 如果 **Kraft 不等式**中的等号成立, 那么 C 必须是最大即时码, 因为没有余地添加更多的码字而保持即时性.

第 4 章 信道编码定理

信道编码的目的：在信源编码的基础上，增加一些适量的冗余信息，保证信号的安全传输，最终考虑在保证传输质量的前提下，最多传输多少数量的信号（信道容量）。

4.1 信道编码问题

信道编码问题 $\left\{ \begin{array}{l} \text{线路编码（变信号）} \\ \text{差错控制编码（信息论主要讨论的问题）} \end{array} \right.$

信道的分类

离散信道：输入输出均为离散事件集。

连续信道：输入输出空间均为连续事件集。

半连续信道：输入和输出一个是离散的，一个是连续的。

信道的概率统计模型（表示参数）

信道的主要参数为转移概率 $p(v | u)$

信道由输入信号集合 \mathcal{U} ，输出信号集合 \mathcal{V} ，及转移概率 $p(v | u)$ 构成。

注： $p(v | u)$ 转移概率：当输入信号字母是 u 时，输出信号字母为 v 的概率。因此 $p(v | u) \geq 0$, $\sum_{v \in \mathcal{V}} p(v | u) = 1$, 对 $\forall u \in \mathcal{U}, v \in \mathcal{V}$ 。信道记为 $\mathcal{C} = \{\mathcal{U}, p(v | u), \mathcal{V}\}$ 。

4.1.1 通信系统的编码误差

通信系统 \mathcal{E} 与编码 (f, g) 给定后，输入消息，输入信号，输出信号与输出消息的随机变量 $(\tilde{\xi}, \xi, \eta, \tilde{\eta})$ 也确定，我们称

$$e(f, g) = P_r\{\tilde{\xi} \neq \tilde{\eta}\}$$

为通信系统 $\mathcal{E}(f, g)$ 所产生的编码误差。

例题 4.1.1 一个简单的通信系统 \mathcal{E} 给定如下。取 \mathcal{S} 和 \mathcal{C} 都是四元信源与信道。这时取

$$\mathcal{X} = \mathcal{U} = \mathcal{V} = \mathcal{Y} = \{0, 1, 2, 3\}$$

它们的信源分布概率为

$$p(0) = p(1) = p(2) = p(3) = 0.25$$

信道的转移概率 $p(v | u)$ 为

$u \backslash v$	0	1	2	3
0	0.64	0.16	0.16	0.04
1	0.16	0.64	0.04	0.16
2	0.16	0.04	0.64	0.16
3	0.04	0.16	0.16	0.64

我们取编码 (f, g) 为 $f(x) = x, g(v) = v, x, v = 0, 1, 2, 3$

$$f: \mathcal{X} \longrightarrow \mathcal{U} \quad \mathcal{U} \xrightarrow{\text{信道传输}} \mathcal{V} \quad g: \mathcal{V} \longrightarrow \mathcal{Y}$$

这时 $(\bar{\xi}, \xi, \eta, \bar{\eta})$ 的联合概率分布为

$$p(x, u, v, y) = \begin{cases} \frac{1}{4}, & \text{如果 } x = u, v = y, \\ 0, & \text{否则.} \end{cases}$$

则此通信系统的编码误差为

$$\begin{aligned} e(f, g) &= p(0)P_r(0 \neq y) + p(1)P_r(1 \neq y) + p(2)P_r(2 \neq y) + p(3)P_r(3 \neq y) \\ &= 0.25 \times (0.16 + 0.16 + 0.04) + 0.25 \times (0.16 + 0.04 + 0.16) + 0.25 \times \\ &\quad (0.16 + 0.04 + 0.16) + 0.25 \times (0.04 + 0.16 + 0.16) \\ &= 0.16 + 0.16 + 0.04 = 0.36 \end{aligned}$$

误差率达到 36%, 无法使用.

从这个例子中可以看出, 如果改变它的编码, 也不会降低它的编码误差. 一个通信系统如果它的传递误差达到 36%, 那么这个近信系统实际上是无法使用的. 因为改变它的编码方式, 不能降低它的编码误差, 因此如何提高通信系统的通信质量是首先要考虑的问题.

例题 4.1.2 为了在信道不变的条件下提高通信系统 \mathcal{C} 的通信质量, 我们减少信源的传输消息数, 如取

$$\mathcal{X} = \mathcal{Y} = \{0, 1\}, \quad p(0) = p(1) = 0.5,$$

而信道与例 4.1.1 相同. 如取编码方案为

$$f(0) = 0, f(1) = 3, \quad g(0) = g(1) = 0, \quad g(2) = g(3) = 1,$$

那么我们计算它的编码误差为

$$e(f, g) = 0.5 \times (0.16 + 0.04) + 0.5 \times (0.16 + 0.04) = 0.2.$$

由例 4.1.1 和例 4.1.2 可知, 在信道不变的条件下, 减少信源的传输消息数可以提高通信质量. 这就是信息论的一个基本原理: **在通信问题中, 牺牲数量可换取质量.**

因此, 信息与编码理论的核心问题是对通信系统寻找一个最佳的数量与质量平衡点. 这就是, **在通信质量到达一定标准条件下, 尽可能多的增加数量.** 在一般通信系统中, 质量标准要求差错率不超过十万分之一 $[e(f, g) \leq 10^{-5}]$. 在通信工程中, 数量与质量又称为有效性与可靠性.

4.1.2 信道序列的编码问题

信道序列的编码问题是在确保编码误差很小的条件下, 尽可能多地传递消息, 我们给出数学描述.

定义 4.1.1

称 R 为信道序列 $\mathcal{C}^n = \{\mathcal{U}^n, p(\mathcal{V}^n | \mathcal{U}^n), \mathcal{V}^n\}$, $n = 1, 2, 3, \dots$ 的一个可达速率, 如果存在一列正数 $\epsilon_n \rightarrow 0$, 一个信源序列 \mathcal{S}^n

$$\begin{cases} \mathcal{X}^n = \mathcal{Y}^n = \{1, 2, \dots, M_n\}, \\ p^{(n)}(x^{(n)}) = \frac{1}{M_n}, \text{ 对任何 } x^{(n)} \in \mathcal{X}^n, \end{cases}$$

以及一列编码函数 $\{f^{(n)}, g^{(n)}\}$, 满足以下条件

(1) 对任何 $n = 1, 2, 3, \dots$, $M_n > 2^{nR(1-\epsilon_n)}$

(2) 由 $\mathcal{S}^n, \mathcal{C}^n, (f^{(n)}, g^{(n)})$ 所确定的通信系统 $\mathcal{E}^n(f^{(n)}, g^{(n)})$ 的误差概率

$$e(f^{(n)}, g^{(n)}) = \Pr\{\tilde{\xi}^{(n)} \neq \tilde{\eta}^{(n)}\} < \epsilon_n.$$



定义 4.1.2

对已给信道序列 \mathcal{C}^n , 它的全体可达速率的最大值或上确界, 被称为信道序列的最大可达速率.



因此, 信道编码问题就是对已给信道序列 \mathcal{C}^n 求它的最大可达速率问题.

4.2 离散无记忆信道

信息论中最常用的是离散无记忆信道.

4.2.1 离散无记忆信道的一般定义

1. 离散无记忆信源

设 $\mathcal{S}^n = \{\mathcal{X}^n, p^{(n)}(x^n)\}$ 是由 $\mathcal{S} = (\mathcal{X}, p(x))$ 确定的信源序列, 如果对任何的 $n = 1, 2, 3, \dots$, $x^n \in \mathcal{X}^n = \{x^n = (x_1, x_2, \dots, x_n) | x_i \in \mathcal{X}\}$, 有

$$p(x^n) = p(x_1, x_2, \dots, x_n) = \prod_{k=1}^n p(x_k)$$

则称 $\mathcal{S}^n = (\mathcal{X}^n, p^n(x^n))$ 是由 $\mathcal{S} = (\mathcal{X}, p(x))$ 确定的无记忆信源序列.

2. 离散无记忆信道

记 \mathcal{C}^n 是一个信道序列, $\mathcal{C}^n = (\mathcal{U}^n, p(v^{(n)} | u^{(n)}, \mathcal{V}^n))$, 如果它的转移概率分布满足

$$p(v^{(n)} | u^{(n)}) = \prod_{i=1}^n p(v_i | u_i)$$

对 $\forall u^{(n)} = (u_1, u_2, \dots, u_n) \in \mathcal{U}^n, v^{(n)} = (v_1, v_2, \dots, v_n) \in \mathcal{V}^n$ 都成立, 其中 $\mathcal{C} = (\mathcal{U}, p(v|u), \mathcal{V})$ 是一固定的信道, 则称 \mathcal{C}^n 是由 \mathcal{C} 决定的无记忆信道序列, 或简称 $\mathcal{C}^n(\mathcal{C})$ 为无记忆信道.

注:

- (1) 离散信道指输入输出字母表均为离散事件集 (有限的);
- (2) 无记忆指当输入字母 u_i 固定时, 它接收信号字母 v_i 的概率与以前、以后输入输出信号无关.

4.2.2 几种特殊的离散无记忆信道

1. 二元对称信道 (无丢失)

记输入输出字母表 $\mathcal{U} = \mathcal{V} = \{0, 1\}$. 信道转移概率分布为

$$p(0|1) = p(1|0) = p, p(0|0) = p(1|1) = 1 - p$$

称 p 为交叉概率误差 (输入 0 输出 1 和输入 1 输出为 0)

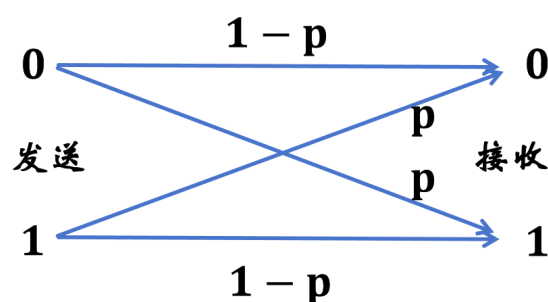


图 4.1: 二元对称信道

信道转移概率矩阵为 $\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$

2. 二元擦除信道 (M 信道)(有丢失)

下图所示的信道称为二元擦除信道, 输出 * 表示输入的丢失或擦除.

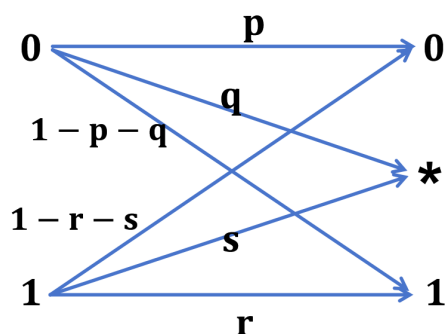


图 4.2: 二元擦除信道

二元擦除信道的一个特例如下图所示, 称之为 M 信道, 这个名字来源于它的图类似于字母 M .

$$p(0|0) = p(1|1) = 1 - p, p(*|0) = p(*|1) = p$$

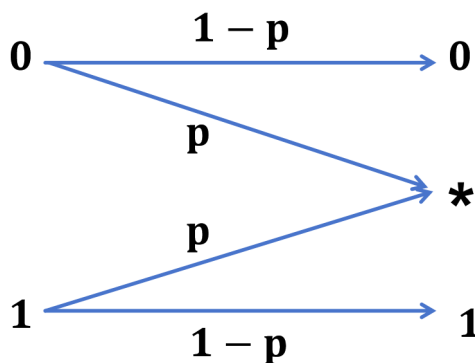


图 4.3: M 信道

3. 信道转移概率矩阵 (信道矩阵)

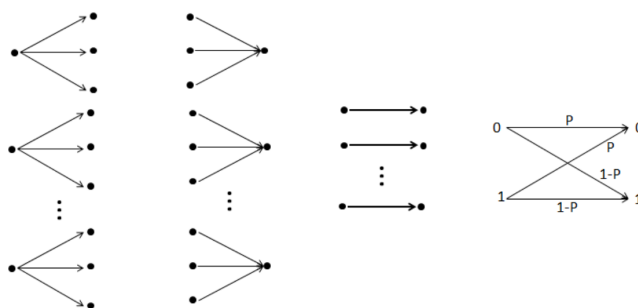
设 $\mathcal{U} = \{u_1, u_2, \dots, u_a\}$, $\mathcal{V} = \{v_1, v_2, \dots, v_b\}$, 定义信道矩阵为

$$\begin{pmatrix} p(v_1|u_1) & p(v_2|u_1) & \cdots & p(v_b|u_1) \\ p(v_1|u_2) & p(v_2|u_2) & \cdots & p(v_b|u_2) \\ \vdots & \vdots & & \vdots \\ p(v_1|u_a) & p(v_2|u_a) & \cdots & p(v_b|u_a) \end{pmatrix}$$

是一个 $a \times b$ 阶矩阵, 则该矩阵的每一行对应一个输入字母, 每一列对应一个输出字母, 该矩阵的每一行元素之和为 1, 即

$$\sum_{j=1}^b p(v_j|u_i) = 1, \quad i = 1, \dots, a$$

按照信道矩阵, 可定义下面几种典型的离散无记忆信道.



无丢失信道 确定的信道 无噪声信道 无用的信道
(有噪无损信道) (有噪有损信道) (无噪无损) (无噪有损)

定义 4.2.1

- (1) 如果输入 ξ 完全由输出 η 所决定, 称该信道是无丢失的.
- (2) 如果输出 η 完全由输入 ξ 所决定, 称该信道为决定的 (确定的).
- (3) 如果信道既是无丢失的也是决定的, 称信道是无噪声的.
- (4) 如果输入随机变量 ξ 的知识不能告诉我们任何关于输出 η 的知识, 称其为无用的信道.

**定义 4.2.2**

根据信道矩阵, 可以定义几种不同信道

- (1) 如果信道矩阵的每一行是另一行的置换, 则称这个信道是行对称的.
- (2) 如果信道矩阵的每一列是另一列的置换, 则称这个信道是列对称的.
- (3) 如果信道矩阵是行对称的, 也是列对称的, 则称这个信道是对称的.



例题 4.2.1 具有信道矩阵 $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ 和 $\begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{pmatrix}$ 的信道都是对称的.

信道矩阵为 $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} \end{pmatrix}$ 的信道是行对称的, 但不是列对称信道.

4. 行对称信道和列对称信道的主要特征**定理 4.2.1**

对于行对称信道, 知道 ξ 时 η 的不确定性与 ξ 分布无关, 即 $H(\eta | \xi)$ 与入口分布无关. 事实上, 对任意 $i = 1, 2, \dots, a$, 我们有

$$H(\eta | \xi) = \sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)}$$

换句话说, 一个信道是行对称的, 如果给定输入 ξ 时关于输出 η 的知识并不依赖于所使用的入口分布.

**证明**

$$\begin{aligned}
 H(\eta | \xi) &= \sum_{i=1}^a p(u_i) H(\eta | \xi = u_i) = \sum_{i=1}^a p(u_i) \left(\sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)} \right) \\
 &= p(u_1) \cdot \sum_{j=1}^b \left(p(v_j | u_1) \log \frac{1}{p(v_j | u_1)} \right) + p(u_2) \cdot \sum_{j=1}^b \left(p(v_j | u_2) \log \frac{1}{p(v_j | u_2)} \right) \\
 &\quad + \dots + p(u_a) \cdot \sum_{j=1}^b \left(p(v_j | u_a) \log \frac{1}{p(v_j | u_a)} \right)
 \end{aligned}$$

因行对称, 每行元素互为置换, 故

$$\sum_{j=1}^b \left(p(v_j | u_k) \log \frac{1}{p(v_j | u_k)} \right) = \sum_{j=1}^b \left(p(v_j | u_\ell) \log \frac{1}{p(v_j | u_\ell)} \right)$$

因此有

$$H(\eta | \xi) = [p(u_1) + \cdots + p(u_a)] \sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)}, \forall i = 1, 2, \dots, a.$$

可得

$$H(\eta | \xi) = \sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)}$$

定理 4.2.2

对于列对称信道, 一个均匀入口分布产生一个均匀出口分布.



证明 设转移概率矩阵 P : 一个 $a \times b$ 矩阵, 其中 a 是输入符号的数量, b 是输出符号的数量. P_{ij} 表示从输入 i 到输出 j 的概率. 均匀入口分布: 所有输入符号发生的概率相等, 即每个输入符号的概率为 $\frac{1}{a}$. 为了证明在均匀入口分布下, 输出分布也是均匀的, 即证明每个输出符号的概率为 $\frac{1}{b}$.

对于列对称信道, 考虑输出符号 j 的概率. 由全概率公式, 输出 j 的概率 $P(\eta = v_j)$ 可以表示为:

$$P(\eta = v_j) = \sum_{i=1}^a P(\xi = u_i) P(\eta = v_j | \xi = u_i)$$

其中, $P(\xi = u_i) = \frac{1}{a}$ (因为入口分布是均匀的). 由于信道是列对称的, 对于任何固定的 j , 所有 $P(\eta = v_j | \xi = u_i)$ 都相等, 因为每列的概率分布是通过相同的方式置换得到的. 设这个常数为 c_j , 那么有:

$$P(\eta = v_j) = \sum_{i=1}^a \frac{1}{a} c_j = c_j$$

因为 $\sum_{i=1}^a \frac{1}{a} = 1$. 要使输出分布均匀, 我们需要证明对于所有的 j , $P(\eta = v_j)$ 都是相等的. 由于列对称性, 所有 c_j 实际上都相等, 因为每一列都可以通过相同的置换得到其他列, 所以它们对应的条件概率相等. 由于有 b 个输出符号, 而且它们的概率之和必须为 1, 所以每个输出符号的概率必须是 $\frac{1}{b}$. 即, 对于所有的 j , 有 $P(\eta = v_j) = \frac{1}{b}$.

4.3 无记忆信道的信道容量

信道容量的概念

设 $\mathcal{C} = (\mathcal{U}, p(v | u), \mathcal{V})$ 是已给信道, 当信道入口分布 $p(u)$ 给定时, 信道的入口与出口联合分布 $p(u, v) = p(u)p(v | u)$ 确定. 因此, 相应的入口与出口随机变量 (ξ, η) 也就确定. 因此它

们的互信息由入口分布 $p(u)$ 与转移概率 $p(v | u)$ 决定, 因此互信息可写成

$$I(\xi; \eta) = I(p(u); p(v | u)).$$

因此当信道 \mathcal{C} 给定时 (也就是信道转移概率 $p(u | v)$ 确定), 互信息就是入口分布 $p(u)$ 的函数. 以下记

$$\mathcal{P}_{\mathcal{U}} = \left\{ \bar{p} \mid p(u) \geq 0, \sum_{u \in \mathcal{U}} p(u) = 1 \right\}$$


为入口分布全体.

4.3.1 信道容量的一般定义

定义 4.3.1

设 \mathcal{C} 是一个固定信道, 那么定义它的信道容量为互信息 $I(\xi; \eta)$ 的最大值, 对所有的入口分布 $\mathcal{P}_{\mathcal{U}}$ 中. 即

$$C = \max \{ I = (p(u); p(v | u)) \mid p(u) \in \mathcal{P}_{\mathcal{U}} \}.$$

如果入口分布 $p_0(u) \in \mathcal{P}_{\mathcal{U}}$, 使 $I(p_0(u); p(u | v)) = C$ 成立. 那么称 $p_0(u)$ 为互信息的最大的入口分布, 简称为最大入口分布. 

注:

(1) 在本章 4.2 节中, 我们已给出了信道序列的最大可达速率的定义. 它与容量在一定条件下可能相等, 但是它们的最初定义的含义是不同的.

(2) 对记号 C , 我们已给出了三种定义, 即: 码元集 C , 信道 \mathcal{C} 与信道容量 C . 注意它们的区别.

(3) 对 $\mathcal{P}_{\mathcal{U}}$ 中的元, 我们分别用 $\bar{p}, p(u), \xi$ 来表示, 它们都是 $\mathcal{P}_{\mathcal{U}}$ 中的概率分布或相应的随机变量.

例题 4.3.1 对于二元对称信道, 计算它的信道容量.

解: 对于二元对称信道, 它的交叉概率为 p .

$$\begin{aligned} I(\xi; \eta) &= H(\eta) - H(\eta | \xi) \\ &= H(\eta) - \sum_{u \in \mathcal{U}} p(u) H(\eta | \xi = u) \\ &= H(\eta) - \sum_{u \in \mathcal{U}} p(u) H(p) \\ &= H(\eta) - \left(\sum_{u \in \mathcal{U}} p(u) \right) H(p) \\ &= H(\eta) - H(p) \\ &\leq 1 - H(p) \end{aligned}$$

注意

$$\begin{aligned}
 H(\eta | \xi = 0) &= p(0 | 0) \log \frac{1}{p(0 | 0)} + p(1 | 0) \log \frac{1}{p(1 | 0)} \\
 &= (1 - p) \log \frac{1}{1 - p} + p \log \frac{1}{p} = H(p), \\
 H(\eta | \xi = 1) &= p(0 | 1) \log \frac{1}{p(0 | 1)} + p(1 | 1) \log \frac{1}{p(1 | 1)} \\
 &= p \log \frac{1}{p} + (1 - p) \log \frac{1}{1 - p} = H(p).
 \end{aligned}$$

如果我们取输入分布为等概分布: $p_0(0) = p_0(1) = \frac{1}{2}$, 那么输出分布为 $q_0(0) = q_0(1) = \frac{1}{2}$, 这时相应的输入、输出随机变量为 (ξ_0, η_0) ,

$$I(\xi_0, \eta_0) = 1 - H(p) \geq H(\eta) - H(p) = I(\xi; \eta)$$

因此二元对称信道的信道容量为

$$C = 1 - H(p)$$

它的最优输入分布为等概分布.

例题 4.3.2 对于 M 信道, 求它的信道容量.

解:

$$\begin{aligned}
 C &= \max \{I(\xi; \eta) | \xi \in \mathcal{P}_{\mathcal{U}}\} = \max \{H(\eta) - H(\eta | \xi) | \xi \in \mathcal{P}_{\mathcal{U}}\} \\
 H(\eta | \xi = 0) &= p(0 | 0) \log \frac{1}{p(0 | 0)} + p(* | 0) \log \frac{1}{p(* | 0)} + p(1 | 0) \log \frac{1}{p(1 | 0)} \\
 &= (1 - p) \log \frac{1}{1 - p} + p \log \frac{1}{p} = H(p) \\
 H(\eta | \xi = 1) &= p(0 | 1) \log \frac{1}{p(0 | 1)} + p(* | 1) \log \frac{1}{p(* | 1)} + p(1 | 1) \log \frac{1}{p(1 | 1)} \\
 &= p \log \frac{1}{p} + (1 - p) \log \frac{1}{1 - p} = H(p).
 \end{aligned}$$

因此

$$H(\eta | \xi) = \sum_{u \in \mathcal{U}} p(u) H(\eta | \xi = u) = \left(\sum_{u \in \mathcal{U}} p(u) \right) H(p) = H(p)$$

$$C = \max \{H(\eta) - H(p) | \xi \in \mathcal{P}_{\mathcal{U}}\}.$$

现在计算 $H(\eta)$ 的值, 如取 $p(0) = \theta, 0 \leq \theta \leq 1$, 那么 we 得到

$$\begin{aligned}
 (q(0), q(*), q(1)) &= (p(0), p(1)) \begin{pmatrix} 1 - p & p & 0 \\ 0 & p & 1 - p \end{pmatrix} \\
 &= (\theta, 1 - \theta) \begin{pmatrix} 1 - p & p & 0 \\ 0 & p & 1 - p \end{pmatrix} \\
 &= (\theta(1 - p), p, (1 - \theta)(1 - p)).
 \end{aligned}$$

代入熵的定义计算可得

$$H(\eta) = (1-p)H(\theta) + H(p).$$

因为 p 是固定常数, 所以 $H(\eta)$ 的最大值是当 $\theta = \frac{1}{2}$ 时取得

$$H(\eta_0) = (1-p)H(\theta) + H(p) = 1-p + H(p).$$

这时

$$C = H(\eta_0) - H(p) = 1-p.$$

且当 $\theta = \frac{1}{2}$ 时达到最大值, 它的最优输入分布为等概率分布.

对于前面定义的几种信道, 可计算它们的信道容量.

定理 4.3.1

- (1) 无丢失信道的容量是 $\log a$, 其中 a 是输入字母表的大小.
- (2) 决定信道的容量是 $\log b$, 其中 b 是集合 $\{v_j \mid \text{存在某个 } u_i, \text{ 使得 } q(v_j \mid u_i) = 1 \text{ 成立}\}$ 的元素个数.
- (3) 无噪声信道的容量是 $\log a$, 其中 a 是输入字母表的大小.
- (4) 无用信道的容量是 0.



证明 (1) 因 ξ 完全由 η 决定, 即 $H(\xi \mid \eta) = 0$.

$$I(\xi; \eta) = H(\xi) - H(\xi \mid \eta) = H(\xi) - 0 = H(\xi) \leq \log a$$

$H(\xi)$ 的最大值 $\log a$.

$$(2) I(\xi; \eta) = H(\eta) - H(\eta \mid \xi) = H(\eta) \leq \log b.$$

(3) 无噪声信道等价条件是, 存在一个 $\mathcal{U} \rightarrow \mathcal{V}$ 的 1-1 映射 ϕ , 使得 $p(\phi(u) \mid u) = 1$ 对所有 u 成立, 从而 $a = b$. 因此 $C = \log a = \log b$.

(4) 无用信道意味着输出不依赖于输入, 或者说输出对于输入的选择完全没有信息. 在这种情况下, 无论输入是什么, 输出的分布都保持不变, 因此 $H(\eta \mid \xi) = H(\eta)$. $I(\xi; \eta) = H(\xi) - H(\xi \mid \eta) = H(\xi) - H(\xi) = 0$. (注意 ξ 与 η 是相互独立的.)

下面计算对称信道的信道容量.

定理 4.3.2

对称信道的信道容量为

$$C = \log b - \sum_{j=1}^b p(v_j \mid u_i) \log \frac{1}{p(v_j \mid u_i)}$$

对 $\forall i = 1, 2, \dots, a$ 都成立, 而且它的最大输入分布为均匀分布 $p(u_i) = \frac{1}{a}$



证明

$$I(\xi; \eta) = H(\eta) - H(\eta \mid \xi) \leq \log b - H(\eta \mid \xi),$$

其中

$$H(\eta | \xi) = \sum_{i=1}^a p(u_i) \left(\sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)} \right)$$

因为信道的对称性, 括号里面的和与 i 无关, 所以

$$\begin{aligned} H(\eta | \xi) &= \left(\sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)} \right) \left(\sum_{i=1}^a p(u_i) \right) \\ &= \sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)}. \end{aligned}$$

可见 $H(\eta | \xi)$ 与 ξ 的分布无关. 因此当我们取入口分布为均匀分布 $p(u) = \frac{1}{a}$ 时, 相应的出口分布为

$$q(v) = \sum_{u \in \mathcal{U}} p(u) p(v | u) = \frac{1}{a} \sum_{u \in \mathcal{U}} p(v | u).$$

由对称信道的定义可知, $\sum_{u \in \mathcal{U}} p(v | u)$ 与 v 无关, 因此 $q(v)$ 与 v 无关, 是个均匀分布. 这时 $\log b$ 是 $H(\eta)$ 的最大值. 于是,

$$C = \log b - H(\eta | \xi) = \log b - \sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)}$$

为对称信道的信道容量, 它的最大入口分布为均匀分布.

例题 4.3.3 对于二元对称信道, 转移矩阵为 $\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$. 于是

$$C = \log 2 - (1-p) \log \frac{1}{1-p} - p \log \frac{1}{p} = 1 - H(p)$$

例题 4.3.4 信道的转移概率矩阵为 $P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$, 求其信道容量. $|\mathcal{U}| = 2, |\mathcal{V}| = 4$

解: 该信道为对称信道, 故

$$\begin{aligned} C &= \log b - \sum_{j=1}^b p(v_j | u_i) \log \frac{1}{p(v_j | u_i)} = \log 4 - \sum_{j=1}^4 p(v_j | u_i) \log \frac{1}{p(v_j | u_i)} \\ &= 2 - 2 \times \frac{1}{3} \log 3 - 2 \times \frac{1}{6} \log 6 = 2 - \frac{2}{3} \log 3 - \frac{1}{3} \log 6 \\ &= 2 - \frac{2}{3} \log 3 - \frac{1}{3} \log 3 - \frac{1}{3} = 2 - \log 3 - \frac{1}{3} \approx 0.082 \text{ 比特/符号}. \end{aligned}$$

4.3.2 无记忆信道序列的容量性质

设 \mathcal{C}^n 是离散无记忆信道序列, 同样可以定义 \mathcal{U}^n 上的全体概率分布 $\mathcal{P}_{\mathcal{U}^n}$, 那么对 $p^{(n)}(u^{(n)})$, 可以确定联合分布

$$p^{(n)}(u^{(n)}, v^{(n)}) = p^{(n)}(u^{(n)}) p(v^{(n)} | u^{(n)}) = p^{(n)}(u^{(n)}) \prod_{j=1}^n p(v_j | u_j)$$

及相应的入口与出口随机变量 $(\xi^{(n)}, \eta^{(n)})$, 于是可定义它们的互信息

$$\begin{aligned} I(\xi^{(n)}; \eta^{(n)}) &= I(p^{(n)}(u^{(n)}); p^{(n)}(v^{(n)} | u^{(n)})) \\ &= \sum_{(u^{(n)} \times v^{(n)}) \in \mathcal{U}^n \times \mathcal{V}^n} p^{(n)}(u^{(n)}, v^{(n)}) \log \left(\frac{p^{(n)}(u^{(n)}, v^{(n)})}{p^{(n)}(u^{(n)}) q^{(n)}(v^{(n)})} \right) \end{aligned}$$

其中 $q^{(n)}(u^{(n)}) = \sum_{v^{(n)} \in \mathcal{V}^n} p^{(n)}(u^{(n)}, v^{(n)})$

定义 4.3.2

信道序列 \mathcal{C}_n 的信道容量定义为

$$C_n = \max \{ I(p^{(n)}(u^{(n)}); p^{(n)}(v^{(n)} | u^{(n)})) \mid p^{(n)}(u^{(n)}) \in \mathcal{P}_{\mathcal{U}^n} \}$$



无记忆信道序列容量的性质

定义 4.3.3

设 \mathcal{Z}^n 是任一有限集合 \mathcal{Z} 上的 n 维乘积空间, $p(z)$ 是 \mathcal{Z} 上的一个概率分布. 如果 $p^{(n)}(z^{(n)}) = \prod_{i=1}^n p(z_i)$, 则称概率分布 $p^{(n)}(z^{(n)})$ 为由 $p(z)$ 确定的无记忆概率分布.



定理 4.3.3

如果 \mathcal{C}^n 是由 \mathcal{C} 确定的无记忆信道, 它们的信道容量分别为 C_n 与 C , 那么必有 $C_n = nC$ 成立, 且 \mathcal{C}^n 的最大入口分布为 $p_0^{(n)}(u^{(n)})$ 是 $p_0(u)$ 确定的无记忆分布, 其中 $p_0(u)$ 是信道 \mathcal{C} 的最大入口分布.



4.4 信道容量的计算

在 4.3 节我们给出了信道容量的定义并求出了一些特殊信道的容量, 但对于一般的信道, 求它的容量并不是件容易的事. 在这一节中我们给出一些计算信道容量的方法.

4.4.1 凸函数的极大值性质

因为信道容量是互信息的最大值, 所以为了求得某个信道的容量, 先研究这个信道的输入和输出之间的互信息. 如前所记, 信道的输入和输出字母表分别是

$$\mathcal{U} = \{u_1, u_2, \dots, u_a\} \quad \mathcal{V} = \{v_1, v_2, \dots, v_b\},$$

信道的转移概率矩阵 $p(v_j | u_i)$ 给定. 因此信道容量的计算问题就是求

$$I(\xi; \eta) = I(p(u); p(v | u)) \quad p(u) \in \mathcal{P}_{\mathcal{U}}$$

的最大值问题.

记入口分布 $\bar{p} = \{p(u_1), p(u_2), \dots, p(u_a)\}$, 我们可以把互信息写成如下形式:

$$I(\xi; \eta) = \sum_{i=1}^a \sum_{j=1}^b p(u_i) p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{\ell=1}^a p(u_\ell) p(v_j | u_\ell)}.$$

引理 4.4.1

对固定的转移概率 $p(v | u)$, 互信息

$$I(\bar{p}) = I(p(u); p(v | u))$$

是 \bar{p} 上的凸函数.



引理 4.4.2

设 $f(\bar{p})$ 是有 n 个输入变量的连续上凸函数, 输入 $\bar{p} = (p_1, p_2, \dots, p_a)$ 满足

$$\sum_{i=1}^a p_i = 1, \quad p_i \geq 0, \quad i = 1, 2, \dots, a$$

且函数的偏导数 $\partial f(\bar{p}) / \partial p_i, i = 1, 2, \dots, a$ 有定义且连续, 则函数 $f(\bar{p})$ 对于 $\bar{p}^* = (p_1^*, \dots, p_a^*)$ 取最大值的充要条件为存在某一实数 λ , 且满足

$$\begin{aligned} \frac{\partial f(\bar{p}^*)}{\partial p_i^*} &= \lambda, \quad \text{对 } p_i^* > 0 \text{ 的 } i, \\ \frac{\partial f(\bar{p}^*)}{\partial p_i^*} &\leq \lambda, \quad \text{对 } p_i^* = 0 \text{ 的 } i. \end{aligned}$$



注:

- (1) 只需证明对于其他输入 \bar{p} , 都有 $f(\bar{p}) \leq f(\bar{p}^*)$
- (2) 求函数极值的方法: 对 $p_i^* > 0$ 的 i , $\left. \frac{\partial f(\bar{p})}{\partial p_i} \right|_{\bar{p}=\bar{p}^*} = \lambda$ 时, $f(\bar{p})$ 取得最大值.

4.4.2 信道容量的计算

我们下面只讲用极值法求信道容量的方法, 利用迭代法求信道容量的方法不讲.

定理 4.4.1

一个信道的入口分布 $\bar{p}^* = (p^*(u_1), \dots, p^*(u_s))$ 使得输入和输出之间的互信息达到最大值的充要条件是存在一个常数 C , 满足

$$\begin{aligned} \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{\ell=1}^a p^*(u_\ell) p(v_j | u_\ell)} &= C, \quad \text{对于 } p^*(u_i) > 0, \\ \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{\ell=1}^a p^*(u_\ell) p(v_j | u_\ell)} &\leq C, \quad \text{对于 } p^*(u_i) = 0. \end{aligned}$$

这时 C 即为信道容量.



证明 令 $I(\xi; \eta) = I(p)$. 由引理, 互信息达到信道容量的充要条件为存在一常数 λ , 使得:

$$\begin{aligned} \left. \frac{\partial I(\bar{p})}{\partial p_i} \right|_{\bar{p}=p^*} &= \lambda, \quad p_i^* > 0, \\ \left. \frac{\partial I(\bar{p})}{\partial p_i} \right|_{\bar{p}=p^*} &\leq \lambda, \quad p_i^* = 0. \end{aligned}$$

$$I(\bar{p}) = \sum_{i=1}^a \sum_{j=1}^b p(u_i) p(v_j | u_i) \log p(v_j | u_i) - \sum_{i=1}^a \sum_{j=1}^b p(u_i) p(v_j | u_i) \log \left(\sum_{\ell=1}^a p(u_\ell) p(v_j | u_\ell) \right)$$

对 $p_i(p(u_i))$ 求偏导, 则其它的 $p(u_j)$ 视为常数, 求偏导为 0.

$$\begin{aligned} \frac{\partial I(\bar{p})}{\partial p_i} &= \sum_{j=1}^b p(v_j | u_i) \log p(v_j | u_i) - \sum_{j=1}^b p(v_j | u_i) \log \left(\sum_{\ell=1}^a p(u_\ell) p(v_j | u_\ell) \right) \\ &\quad - \sum_{k=1}^a \sum_{j=1}^b p(u_k) p(v_j | u_k) \frac{p(v_j | u_i)}{\sum_{\ell=1}^a p(u_\ell) p(v_j | u_\ell)} \\ &= \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{q(v_j)} - \sum_{j=1}^b \left(\sum_{k=1}^a p(u_k) p(v_j | u_k) \right) \frac{p(v_j | u_i)}{\sum_{\ell=1}^a p(u_\ell) p(v_j | u_\ell)} \\ &= \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{q(v_j)} - \sum_{j=1}^b p(v_j | u_i) \frac{\sum_{k=1}^a p(u_k) p(v_j | u_k)}{\sum_{\ell=1}^a p(u_\ell) p(v_j | u_\ell)} \\ &= \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{q(v_j)} - \sum_{j=1}^b p(v_j | u_i) \\ &= \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{q(v_j)} - 1 \\ &= \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{\ell=1}^a p(u_\ell) p(v_j | u_\ell)} - 1 \end{aligned}$$

令

$$\lambda = \left. \frac{\partial I(\bar{p})}{\partial p_i} \right|_{p=p^*},$$

则

$$\lambda = \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{k=1}^a p_k^* p(v_j | u_k)} - 1$$

两边同时乘以 p_i^* 再对 i 求和有

$$\begin{aligned}
 \lambda &= \sum_{i=1}^a p_i^* \frac{\partial I(\bar{p})}{\partial p_i} \Big|_{p=p^*} \\
 &= \sum_{i=1}^a \sum_{j=1}^b p_i^* p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{k=1}^a p_k^* p(v_j | u_k)} - \sum_{i=1}^a p_i^* \\
 &= \sum_{i=1}^a \sum_{j=1}^b p_i^* p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{k=1}^a p_k^* p(v_j | u_k)} - 1 \\
 &= I(\bar{p}^*) - 1 = C - 1
 \end{aligned}$$

根据第二个引理, $I(\bar{p})$ 在 \bar{p}^* 取最大值 $I(\bar{p}^*)$, $C = I(p^*) = \lambda + 1$.

$$C = \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{\ell=1}^a p^*(u_\ell) p(v_j | u_\ell)}, \quad \text{对于 } p^*(u_i) > 0.$$

例题 4.4.1 对于 M 信道, 它的信道矩阵为

$$\begin{pmatrix} 1-p & 0 & p \\ 0 & 1-p & p \end{pmatrix}, \quad 0 < p < 1.$$

设入口分布为 $\bar{p} = (p_0, p_1)$, 对应的出口分布为 $\bar{q} = \{q_0, q_1, q_2\}$, 而 $q(v_j) = \sum_{i=1}^a p(u_i) p(v_j | u_i)$. 由定理的结论知

$$\begin{aligned}
 C &= \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{\sum_{\ell=1}^a p(u_\ell) p(v_j | u_\ell)} \\
 &= \sum_{j=1}^b p(v_j | u_i) \log \frac{p(v_j | u_i)}{q(v_j)} \\
 &= (1-p) \log \frac{1-p}{q_0} + p \log \frac{p}{q_2} = (1-p) \log \frac{1-p}{q_1} + p \log \frac{p}{q_2}
 \end{aligned}$$

于是有 $q_0 = q_1$, 而

$$(q_0, q_1, q_2) = (p_0, p_1) \begin{pmatrix} 1-p & 0 & p \\ 0 & 1-p & p \end{pmatrix} = (p_0(1-p), p_1(1-p), p).$$

由 $q_0 = q_1$ 知, $p_0 = p_1 \Rightarrow p_0 = p_1 = \frac{1}{2}$, $q_0 = q_1 = \frac{1-p}{2}$, $q_2 = p$, 于是有 $C = 1 - p$.

4.5 信道的编码和译码问题

在本章 4.2 节中, 我们已给出了信道序列的可达速率的定义, 它的实质是存在适当的编码与译码算法, 使消息传递的误差很小, 而且可以转送一定数量的数据量. 我们现在给出一个关于信道序列的可达速率的等价条件. 我们仍记

$$\mathcal{C}^n = (\mathcal{U}^n, p(v^{(n)}) | u^{(n)}, \mathcal{V}^n)$$

为信道序列, 记 $\mathcal{S}^n = \{\mathcal{X}^n, p^{(n)}(x^n)\}$ 为信源序列, 在 4.1 节中我们已经给出可达速率 R 的定义, 且取信源序列为 \mathcal{S}^n

$$\begin{cases} \mathcal{X}^n = \mathcal{Y}^n = \{1, 2, \dots, M_n\}, \\ P^{(n)}(x^{(n)}) = \frac{1}{M_n}, \text{ 对任何 } x^{(n)} \in \mathcal{X}^n, \end{cases}$$

所给, 那么信道的编码问题就是求信道序列的最大可达速率问题. 为讨论这个问题, 我们先给出 R 是可达速率的一个等价条件, 由这个等价条件也可看到信道的编码问题的本质.

定义 4.5.1

信道序列仍记为 \mathcal{C}^n . 我们称

$$(u_i^{(n)}, \mathcal{B}_i^{(n)}), i = 1, 2, \dots, M_n$$

是 \mathcal{C}^n 的一组 ϵ 专线, 如果它满足以下条件.

(1) $u_i^{(n)} \in \mathcal{U}^n$, 它们互不相同.

(2) $\mathcal{B}_i^{(n)}$ 是 \mathcal{V}^n 的一组子集, 它们互不相交, 即

$$\mathcal{B}_i^{(n)} \cap \mathcal{B}_j^{(n)} = \emptyset, \text{ 当 } i \neq j \text{ 时},$$

(3) 对任何 $i = 1, 2, \dots, M_n$, 有 $P^{(n)}(\mathcal{B}_i^{(n)} | u_i^{(n)}) > 1 - \epsilon$ 成立. 称式中的 M_n 为 ϵ 专线的数目.



通过以下定理可以看到, ϵ 专线与可达速率的关系问题.

定理 4.5.1

R 是信道序列 \mathcal{C}^n 可达速率的充分与必要条件是存在数列 $\epsilon_n \rightarrow 0$, 使 \mathcal{C}^n 有 $M_n > 2^{nR(1-\epsilon_n)}$ 条 ϵ_n 专线.



该定理不证明, 掌握结论.

具有 ϵ_n 专线的编、译码的信息传输图如图所示.

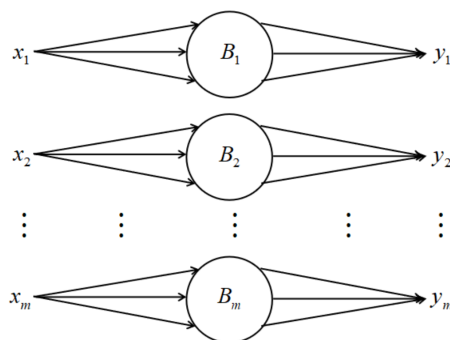


图 4.4: ϵ 专线的编码和译码判决方案表示图

可以看出, 专线对通信系统的编码和译码方式给出了一个形象的描述, 对我们理解通信编码问题很有帮助.

4.6 信道的正编码定理和反编码定理

在前几节中, 我们已给出了信道序列的最大可达速率与信道容量的定义. 我们现在讨论它们的相互关系问题. 这就是信道编码定理.

定理 4.6.1 (信道编码的正编码定理)

如果 C 是一个离散无记忆信道 \mathcal{C} 的信道容量, 那么 C 必是该离散无记忆信道序列的一个可达速率.



证明 关于信道编码定理的证明有许多种方法, 较为精确的证明方法为组合法, 但它涉及许多排列组合的计算. 另外, 还可以利用 ϵ 专线的方法给以证明, 该方法是编码定理的经典证明方法, 很为直观. 在本书中我们采用随机码的方法给以证明, 这是一种很为巧妙的证明方法, 在信息论中广为使用. 对此我们分以下几步给出证明思路.

1. 随机码的定义

因为 \mathcal{C}^n 是离散无记忆信道序列, 所以它的转移概率分布为

$$p^{(n)}(v^{(n)} | u^{(n)}) = \prod_{i=1}^n p(v_i | u_i)$$

它由 $\mathcal{C} = \{\mathcal{U}, P(v | u), \mathcal{V}\}$ 决定. 记 $p_0(u)$ 是信道的入口分布, 使 $I(p_0(u); p(v | u)) = C$ 成立. 设 \mathcal{S} 为具有均匀分布的信源,

$$\mathcal{X}^n = \{1, 2, \dots, M_n\}, P^{(n)}(i) = \frac{1}{M_n} \quad (4.6.1)$$

那么随机码的定义为 $\mathcal{X}^n \rightarrow \mathcal{U}^n$ 的一组随机映射

$$\bar{f}^* = \{f_1^*, f_2^*, \dots, f_{M_n}^*\} \quad (4.6.2)$$

它满足以下条件.

- (1) $M_n = |\mathcal{X}^n|$ 是信源字母表的元素个数.
- (2) $f_1^*, f_2^*, \dots, f_{M_n}^*$ 一组独立同分布的随机变量.
- (3) 每个 f^* 在 \mathcal{U}^n 中取值, 并具有分布为

$$p_0^{(n)}(u^{(n)}) = \prod_{i=1}^n p(u_i)$$

随机码 \bar{f}^* 的一个样本值记为

$$\bar{f}^{(n)} = \{f_1^{(n)}, f_2^{(n)}, \dots, f_{M_n}^{(n)}\} = \{u_1^{(n)}, u_2^{(n)}, \dots, u_{M_n}^{(n)}\} \quad (4.6.3)$$

它就是 $\mathcal{X}^n \rightarrow \mathcal{U}^n$ 的一个编码.

如果信道的输入概率分布为 $p_0^{(n)}(u^{(n)})$ 与信道转移概率分布 $p^{(n)}(v^{(n)} | u^{(n)})$, 那么信道输入与输出的联合概率分布密度与互信息密度函数确定, 它们分别为

$$p_0^{(n)}(u^{(n)}, v^{(n)}) = p_0^{(n)}(u^{(n)}) p^{(n)}(v^{(n)} | u^{(n)})$$

及

$$i_0(u^{(n)}; v^{(n)}) = \log \left(\frac{p_0^{(n)}(u^{(n)}, v^{(n)})}{p_0^{(n)}(u^{(n)}) q_0^{(n)}(v^{(n)})} \right)$$

由于输入概率分布 $p_0^{(n)}(u^{(n)})$ 、信道转移概率分布 $p^{(n)}(v^{(n)} | u^{(n)})$ 是无记忆的, 可知 $p_0^{(n)}(u^{(n)}, v^{(n)})$, $q_0^{(n)}(v^{(n)})$ 是无记忆的.

因此互信息密度函数为

$$i_0(u^{(n)}; v^{(n)}) = \sum_{j=1}^n \log \left(\frac{p_0(u_j, v_j)}{p_0(u_j) q_0(v_j)} \right) \quad (4.6.4)$$

2. 阈值译码算法与带随机编码的通信系统

阈值译码算法的定义如下.

定义 4.6.1

如果 \bar{f} 是由式 (4.6.3) 给定的随机编码样本, 那么它的码元集合也由式 (4.6.3) 给定. 如果 $v^{(n)}$ 是信道序列的一个输出向量, 那么以下的译码算法为阈值译码算法,

$$g(v^{(n)}) = \begin{cases} j, & \text{如果 } i_0(u_j^{(n)}; v^{(n)}) > K_n \\ 1, & \text{否则.} \end{cases} \quad (4.6.5)$$

其中 K_n 是一个适当的常数, 我们称之为阈值.



由以上定义, 我们得到一个带有随机编码的通信系统, 并记之为

$$\varepsilon^* = \{\bar{f}^*, \mathcal{C}\} = \{j, f_j^*, v_j^*, g(v_j^*) : j = 1, 2, \dots, M_n\}, \quad (4.6.6)$$

对此模型, 我们说明如下.

- (1) 在式 (4.6.6) 中, j 是消息字母, 它的取值概率为 $\Pr\{\tilde{\xi} = j\} = \frac{1}{M_n}$.
- (2) f_j^* 是由消息字母 j 决定的随机码, 它在 \mathcal{U}_n 上取值, 具有概率分布为 $p_0^{(n)}(u^{(n)})$.
- (3) 由随机码的假定, 当 $j \neq k$ 与时, f_j^* 与 f_k^* 相互独立.
- (4) v_j^* 是由输入信号 f_j^* 与信道 \mathcal{C}^n 决定的输出信号, 它在 \mathcal{V}^n 上取值, 当 $f_j^* = u^{(n)}$ 时, V_j^* 的概率分布为

$$\Pr\{v_j^* = v^{(n)} | f_j^* = u^{(n)}\} = p^{(n)}(v^{(n)} | u^{(n)})$$

因此 (f_j^*, v_j^*) 的联合概率分布为

$$p_0^{(n)}(u^{(n)}, v^{(n)}) = \Pr\{f_j^* = u^{(n)}, v_j^* = v^{(n)}\} = p_0^{(n)}(u^{(n)}) p^{(n)}(v^{(n)} | u^{(n)}),$$

- (5) 当 $k \neq j$ 与时, (f_k^*, v_k^*) 与 v_j^* 相互独立, 因此它们的联合概率分布为

$$\Pr\{f_j^* = u^{(n)}, v_k^* = v^{(n)}\} = p_0^{(n)}(u^{(n)}) q_0(v^{(n)}),$$

- (6) $g(v_j^*)$ 是由式 (4.6.5) 给定的译码函数.

3. 随机编码的误差概率

我们现在考虑由式 (4.6.3) 和式 (4.6.5) 给定的随机码的编、译码方案的误差问题. 它可能

出现两种不同类型的误差.

定义 4.6.2

如果 $\bar{f}(n)$ 是随机码 \bar{f}^* 确定的一个编码, 由式 (4.6.3) 给定, $g(v^{(n)})$ 是由式 (4.6.5) 给出的阈值译码, 那么通信系统 (4.6.6) 式所出现的两种不同类型的误差概率为

- (1) 第一类误差概率. 如果发送消息是 j , 但是最终还原消息 $g(v_j^*) \neq j$;
- (2) 第二类误差概率. 如果发送消息是某个 j , 但是有一个其他的发送消息 $k \neq j$, 使 k 的还原消息是 j , 也就是 $g(v_k^*) = j$.



这两类误差我们分别记为 $e_{1,j}(\bar{f}^{(n)}, g)$, $e_{2,j}(\bar{f}^*, g)$, 它们分别是

$$\begin{cases} e_{1,j}(\bar{f}^{(n)}, g) = \Pr\{g(v_j^*) \neq j\} \\ e_{2,j}(\bar{f}^*, g) = \Pr\{\text{有一个 } k \neq j, \text{ 使 } g(v_k^*) = j\}, \end{cases}$$

那么它们的平均误差分别为

$$e_\tau(\bar{f}^{(n)}, g) = \frac{1}{M_n} \sum_{j=1}^{M_n} e_\tau(\bar{f}^{(n)}, g(v_j^*)), \quad \tau = 1, 2.$$

而记

$$e_0(\bar{f}^{(n)}, g) = e_1(\bar{f}^{(n)}, g) + e_2(\bar{f}^{(n)}, g)$$

为总误差概率. 我们以下记

$$e_\tau(\bar{f}^*, g) = \sum_{\bar{f}^{(n)}} \Pr\{f^* = \bar{f}^{(n)}\} e_\tau(\bar{f}^{(n)}, g), \quad \tau = 1, 2$$

为随机码 \bar{f}^* 的平均概率误差, 其中 $\Pr\{f^* = \bar{f}^{(n)}\}$ 为随机码 f^* 取样本值 $\bar{f}^{(n)}$ 概率. 我们现在估计 $e_\tau(\bar{f}^*, g)$ 的值.

4. 关于随机编码的误差概率的估计

在对随机编码的误差概率进行估计时, 我们首先注意到在随机编码 f_j^* 与接受信号之间的对称性, 因此有

$$e_\tau(\bar{f}^*, g) = e_\tau(\bar{f}_1^*, g), \quad \tau = 1, 2$$

成立, 这样只要估计 $e_\tau(\bar{f}_1^*, g)$ 的值就可. 可以证明, 取 $\epsilon > 0$ 是任意小的正数, 当 n 充分大时, 必有

$$e_1(\bar{f}_1^*, g) < \frac{\epsilon}{2}, \quad e_2(\bar{f}_1^*, g) < \frac{\epsilon}{2}$$

成立. 由此可得, 对任何 $\epsilon > 0$, 对均匀分布的信源 \mathcal{S} , 当它的消息数 $M_n = 2^{(nR(1-\epsilon))}$ 时, 如果编码为式 (4.6.2) 的随机码, 而译码是式 (4.6.4) 的阈值 $K_n = nR(1 - \frac{\epsilon}{2})$, 则当 n 充分大时, 相应的平均误差概率

$$e_0(\bar{f}^*, g) \leq e_1(\bar{f}^*, g) + e_2(\bar{f}^*, g) \leq \epsilon$$

因为 ϵ 是任意取的, 所以必存在一系列 $\epsilon_n \rightarrow 0$, 使以上命题同样成立. 因此 $R = C$ 是无记忆信道序列 \mathcal{C}^n 的可达速率. 定理得证.

我们现在讨论无记忆信道编码的反编码定理, 这就是信道容量是无记忆信道序列的一个最大可达速率.

定理 4.6.2 (无记忆信道编码的反编码定理)

如果 C 是一个离散无记忆信道 \mathcal{C} 的信道容量, 那么对任何 $R > C$, 则 R 一定不是该离散无记忆信道序列的可达速率.



该定理不证明.

综合信道编码定理和逆定理, 我们可知码率小于信道容量是错误概率趋于 0 的充分必要条件.

信道编码定理的证明方法为随机码方法, 它首次由 Shannon 在他的原始论文中提出. 这种方法虽然巧妙, 但它却不是构造性的, 它说明存在许多满足定理要求的码, 但它并没有告诉我们具体的构造方法. 从 Shannon 发表他的文章到现在, 编码学者们一直在寻找构造满足信道编码定理条件的码的具体方法.

在实际的编码工作中, 一个码仅满足定理要求是远远不够的, 它的编码和译码计算必须快速实现, 使它的运算与通信同步, 这样才具有应用价值. 在本书后面的章节中, 我们将会看到, 为了使编码和译码计算快速实现, 需要借助于代数或几何的工具, 构造出各种有用码.

4.7 可加高斯 (Gaussian) 信道

在前几节讨论的信道都是离散信道, 实际通信的信号在许多情形下是连续的. 连续信号通过分层处理才变成离散信号, 因此对连续信道的研究也是十分重要的. 现代的调制解调码理论就是对连续信号的直接处理.

在连续信道中最重要的信道就是可加高斯信道, 它的定义如下.

定义 4.7.1

对连续型信道我们有以下定义.

- (1) 称一个信道 $\mathcal{C} = \{\mathcal{U}, P(v | u), \mathcal{V}\}$ 为连续信道, 如果 \mathcal{U}, \mathcal{V} 是连续型集合, 如取 $\mathcal{U} = \mathcal{V} = \mathbf{R}$ 是全体实数集合.
- (2) 在连续信道中, 如果存在一个随机变量 ζ , 与任何输入信号 ξ 的取值无关, 且输出信号总有 $\eta = \xi + \zeta$ 成立, 那么称这个信道为可加噪声信道, 称 ζ 为噪声随机变量.
- (3) 在可加噪声信道中, 如果噪声随机变量是一个均值为零的正态随机变量, 那么称这个信道为可加高斯信道.



在可加高斯信道中, 记 ζ 的方差为 σ_N^2 , 那么 ζ 具有正态分布 $N(0, \sigma_N^2)$. 这时信道 \mathcal{C} 的转移概率分布密度为

$$p(v | u) = \frac{1}{\sqrt{2\pi\sigma_N^2}} \exp\left(-\frac{(v-u)^2}{2\sigma_N^2}\right). \quad (4.7.1)$$

在连续信道的研究中, 对它的输入、输出及转移概率分布, 一般用分布密度来讨论, 如果记

$p(u)$ 是输入信号的概率分布密度, 那么它的输入、输出概率分布密度同样用 $p(u, v) = p(u)p(v | u)$ 来表示. 这时输入与输出信号的互信息为

$$I(\xi; \eta) = \int_{\mathcal{U}} \int_{\mathcal{V}} p(u, v) \log \frac{p(u, v)}{p(u)q(v)} dv du, \quad (4.7.2)$$

其中 $q(v) = \int_{\mathcal{U}} p(u, v) du$ 为输出信号的概率分布密度, 而 ξ, η 分别是信道的输入、输出随机变量.

对连续状态下的互信息, 同样可有关系式

$$\begin{aligned} I(\xi; \eta) &= H(\xi) + H(\eta) - H(\xi, \eta) \\ &= H(\xi) - H(\xi | \eta) \\ &= H(\eta) - H(\eta | \xi) \end{aligned}$$

成立, 其中 $H(\xi), H(\eta), H(\xi, \eta)$ 和 $H(\xi | \eta), H(\eta | \xi)$ 分别是 ξ, η 的熵、联合熵和条件熵, 它们的定义分别为

$$\begin{aligned} H(\xi) &= - \int_{\mathcal{U}} p(u) \log[p(u)] du, \\ H(\eta) &= - \int_{\mathcal{V}} q(v) \log[q(v)] dv \\ H(\xi, \eta) &= - \int_{\mathcal{U} \times \mathcal{V}} p(u, v) \log[p(u, v)] dv du, \end{aligned}$$

而

$$H(\xi | \eta) = H(\xi, \eta) - H(\eta), \quad H(\eta | \xi) = H(\xi, \eta) - H(\xi).$$

引理 4.7.1

在可加高斯信道中, 条件熵 $H(\eta | \xi)$ 与输入分布 $p(u)$ 无关, 且

$$H(\eta | \xi) = \frac{1}{2} \log(2\pi e \sigma_N^2)$$



该引理不证明. 在可加高斯信道中, 方差 σ_N^2 是干扰信号功率的强度. 在连续信道中, 一般对输入信号的功率强度应有限制. 因此, 连续信道的信道容量应定义为对输入信号功率强度限制在一定区域内的最大互信息.

定义 4.7.2

如果 $\mathcal{C} = \{\mathcal{U}, P(v | u), \mathcal{V}\}$ 为一个连续信道, 取 $\mathcal{U} = \mathcal{V} = \mathbf{R}$, 那么它的信道容量应定义为

$$C = \sup \{I(\xi; \eta) | \text{Var}(\xi) \leq \sigma_S^2\} \quad (4.7.3)$$

其中 σ_S^2 是个常数, 它代表信道输入信号功率的上限, $I(\xi; \eta)$ 是输入、输出信号的互信息, 由式 (4.7.2) 定义, 而

$$\text{Var}(\xi) = \int_{-\infty}^{\infty} (u - \mu)^2 p_{\xi}(u) du$$

是输入信号的方差, 其中 $\mu = \int_{-\infty}^{\infty} up_{\xi}(u)du$ 是输入信号的均值.



定理 4.7.1

在可加高斯信道中, 如果输入信号功率的上限与干扰信号功率的强度分别为 σ_S^2 与 σ_N^2 , 那么该信道的信道容量为

$$C = \frac{1}{2} \log \left(1 + \frac{\sigma_S^2}{\sigma_N^2} \right). \quad (4.7.4)$$

在信息论中, 称 $\frac{\sigma_S^2}{\sigma_N^2}$ 为信噪比.



证明 由可加信道的定义可知, 如记 ξ, η, ζ 分别为输入、输出信号与噪声的随机变量, 这时 $\eta = \xi + \zeta$. 因为 ξ 与 ζ 相互独立, 所以

$$\text{Var}(\eta) = \text{Var}(\xi) + \text{Var}(\zeta) = \sigma_S^2 + \sigma_N^2$$

由引理 4.7.1 和最大熵原理可得

$$\begin{aligned} I(\xi; \eta) &= H(\eta) - \frac{1}{2} \log(2\pi e \sigma_N^2) \\ &\leq \frac{1}{2} \log(2\pi e (\sigma_S^2 + \sigma_N^2)) - \frac{1}{2} \log(2\pi e \sigma_N^2) \\ &= \frac{1}{2} \log \left(\frac{\sigma_S^2 + \sigma_N^2}{\sigma_N^2} \right) \\ &= \frac{1}{2} \log \left(1 + \frac{\sigma_S^2}{\sigma_N^2} \right). \end{aligned}$$

另一方面, 如取输入随机变量 ξ 为正态分布 $N(0, \sigma_S^2)$ 时, 其中的等号成立, 因此式 (4.7.4) 成立. 定理得证. 可加高斯信道在调制解调码理论中得到应用.

4.8 习题课

4.8.1 基本概念

1. 通信编码误差: $\mathcal{E} = \{\mathcal{S}, \mathcal{C}\}$, (f, g) 固定, $e(f, g) = \Pr\{\tilde{\xi} \neq \tilde{\eta}\}$ 为通信系统 $\mathcal{E}(f, g)$ 所产生的编码误差.

2. 几种类型的无记忆信道:

- (1) 无丢失信道: ξ 完全由 η 决定
- (2) 确定 (决定) 信道: η 完全由 ξ 决定
- (3) 无噪声信道: 无丢失且决定
- (4) 无用信道: 由 ξ 得不到关于 η 的任何信息.

3. 信道矩阵


4. 行对称、列对称、对称信道

5. 信道容量: $C = \max I(p(u); p(v | u))$

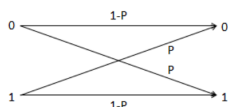
4.8.2 基本方法

1. 利用定义计算二元对称信道, M 信道, 四种典型无记忆信道的信道容量.
2. 利用极值法计算几种典型信道的信道容量 (二元对称信道, M 信道, Z 信道)

4.8.3 课后习题

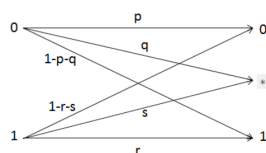
 **练习 4.8.1** 写出二元对称信道, 二元擦除信道及 M 信道的信道矩阵.

解:



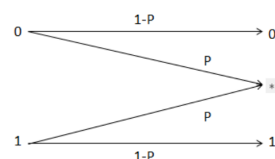
$$\begin{matrix} 0 & 1 \\ \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \\ 1 \end{matrix}$$

(a) 二元对称信道




$$\begin{matrix} 0 & * & 1 \\ \begin{pmatrix} p & q & 1-p-q \\ 1-r-s & s & r \end{pmatrix} \\ 1 \end{matrix}$$

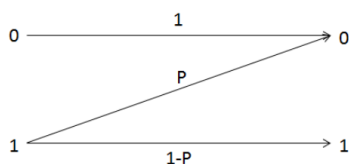
(b) 二元擦除信道



$$\begin{matrix} 0 & * & 1 \\ \begin{pmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{pmatrix} \\ 1 \end{matrix}$$

(c) M 信道

 **练习 4.8.2** Z 信道如下图所示, 求其信道容量



$$\begin{matrix} 0 & 1 \\ \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix} \\ 1 \end{matrix}$$

解: 根据信道容量的定义有

$$C = p(0 | 0) \log \frac{p(0 | 0)}{q_0} + p(1 | 0) \log \frac{p(1 | 0)}{q_1} = p(0 | 1) \log \frac{p(0 | 1)}{q_0} + p(1 | 1) \log \frac{p(1 | 1)}{q_1}$$

$$\text{左边} = \log \frac{1}{q_0},$$

$$\begin{aligned} \text{右边} &= p \log \frac{p}{q_0} + (1-p) \log \frac{1-p}{q_1} \\ &= p \log p - p \log q_0 + (1-p) \log(1-p) - (1-p) \log q_1 \\ &= -H(p) - p \log q_0 - (1-p) \log q_1. \end{aligned}$$

$$(q_0, q_1) = (\theta, 1-\theta) \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix} = (\theta + p(1-\theta), (1-\theta)(1-p))$$

$$-\log q_0 + p \log q_0 + (1-p) \log q_1 = -H(p)$$

$$(1-p) \log q_0 + (p-1) \log q_1 = H(p)$$

$$(1-p) \log \frac{q_0}{q_1} = H(p)$$

$$\log \frac{q_0}{q_1} = \frac{H(p)}{1-p}$$

$$\frac{q_0}{q_1} = 2^{\frac{H(p)}{1-p}}$$

$$\frac{\theta + p(1-\theta)}{(1-\theta)(1-p)} = 2^{\frac{H(p)}{1-p}}$$

$$\frac{\theta - 1 + p(1-\theta) + 1}{(1-\theta)(1-p)} = 2^{\frac{H(p)}{1-p}}$$

$$\frac{(1-\theta)(p-1) + 1}{(1-\theta)(1-p)} = 2^{\frac{H(p)}{1-p}}$$

$$-1 + \frac{1}{(1-\theta)(1-p)} = 2^{\frac{H(p)}{1-p}}$$

$$\frac{1}{(1-\theta)(1-p)} = 2^{\frac{H(p)}{1-p}} + 1$$

$$\frac{1}{1-\theta} = (1-p) \left[2^{\frac{H(p)}{1-p}} + 1 \right]$$

$$\theta = \frac{1}{1-p} \left[-p + \frac{2^{\frac{H(p)}{1-p}}}{1 + 2^{\frac{H(p)}{1-p}}} \right]$$

于是

$$C = \log \frac{1}{\theta + p(1-\theta)} = \log \frac{1}{\theta(1-p) + p} = \log \frac{1 + 2^{\frac{H(p)}{1-p}}}{2^{\frac{H(p)}{1-p}}} = \log \left[1 + 2^{\frac{-H(p)}{1-p}} \right].$$

 **练习 4.8.3** 考虑离散无记忆信道 $Y = (X + Z) \bmod 11$, 其中

$$Z = \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$X \in \{0, 1, \dots, 10\}$. 假设 X 和 Z 独立.

(1) 求这个信道的容量.

(2) 找出达到信道容量的入口分布.

解: (1) $Y = (X + Z) \bmod 11$, 输入为 X , 输出为 Y , Z 为噪声信道, 而

$$Z = \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}, \quad X \in \{0, 1, \dots, 10\}$$

所以 $Y \in \{0, 1, \dots, 10\}$,

因为 Z 可以取三个值 $(1, 2, 3)$, 每个都有 $\frac{1}{3}$ 的概率, 所以对于每个 X 的值, Y 可以是三个可能的结果之一, 这取决于 Z 的值. 信道矩阵 $P(Y | X)$ 将具有 11 行 (对应于 X 的可能值) 和 11 列 (对应于 Y 的可能值). 每个元素 P_{ij} 表示给定输入 $X = i$ 时输出 $Y = j$ 的概率.

由于 Z 的作用是加在 X 上然后对 11 取模, 每个 X 值将映射到三个不同的 Y 值, 每个的概率都是 $\frac{1}{3}$. 例如, 如果 $X = 0$, 则 Y 可以是 1, 2 或 3, 每个都有 $\frac{1}{3}$ 的概率, 因为 Z 分别加 1, 2 或 3. 因此, 信道矩阵的一般形式将是每行有三个 $\frac{1}{3}$ 的条目, 分别对应于 X 加上 1, 2, 3 和模 11 的结果, 而其他位置为 0. 对于 $X = 0$: Y 的可能值是 1, 2, 3, 每个概率为 $\frac{1}{3}$. 对于 $X = 1$: Y 的可能值是 2, 3, 4, 每个概率为 $\frac{1}{3}$. 以此类推, 直到 $X = 10$. 每行的具体值会随着 X 的增加而“滚动”, 并在达到 10 并绕回 0 时循环. 这种模式的重复构成了完整的信道矩阵:

$$\begin{array}{c} \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{matrix} \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

可见 Z 对输入 X 的每个取值的作用效果一样, 这是一个对称信道, 因此

$$\begin{aligned} C &= \log 11 - \left(\frac{1}{3} \log 3 + \frac{1}{3} \log 3 + \frac{1}{3} \log 3 \right) \\ &= \log 11 - \log 3 = \log \frac{11}{3}. \end{aligned}$$

(2) 对于对称信道, 最大化信道容量的输入分布是等概率分布, 即

$$p(X) = \frac{1}{11}, \quad X \in \{0, 1, \dots, 10\}.$$

这个等概率分布的选择反映了信息论中的一个基本原则: 当所有可能的事件 (在这个情况下是输入信号的选择) 都具有相同的概率时, 不确定性 (和因此信息熵) 最大. 在一个对称信道中, 由于所有的输入对输出的影响是相同的, 等概率分布确保了输出分布也尽可能均匀, 从而最大化了从输入到输出的信息传递.

第5章 抽象代数的基本知识

5.1 群

5.1.1 群的概念

定义 5.1.1

设 $G \times G \rightarrow G$ 是 G 上的一个二元运算, 若“ \cdot ”满足

- (1) $\forall a, b, c \in G, (a \cdot b) \cdot c = a \cdot (b \cdot c)$;
- (2) $\exists e \in G$, 有 $a \cdot e = e \cdot a = a$;
- (3) $\forall a \in G, \exists a^{-1} \in G$, 有 $a \cdot a^{-1} = a^{-1} \cdot a = e$.

则称 G 关于二元运算“ \cdot ”作成一个群. $a \cdot b$ 简记作 ab .



例题 5.1.1 $(\mathbb{Z}, +)$, 单位元 $0, \forall a \in \mathbb{Z}, -a = a^{-1}$.

例题 5.1.2 (\mathbb{Z}_p^*, \odot) , 单位元 $1, \mathbb{Z}_p^* = \{1, 2, \dots, p-1\}$. p 为素数.

$a \odot b = ab(\text{mod } p)$, 若 $(a, p) = 1$ 即 $\exists s, t$, 使得 $as + pt = 1$, 因此 $as \equiv 1(\text{mod } p)$, 故 $a^{-1} = s$. 则 (\mathbb{Z}_p, \odot) 是一个群.

$$a \oplus b = (a + b)(\text{mod } p)$$

$\mathbb{Z}_p = \{0, 1, 2, \dots, p-1\}$, 根据上面的定义 (\mathbb{Z}_p, \oplus) 构成一个含有 p 个元素的交换群.

注: G 是一个群, 若对 $\forall a, b \in G$, 有 $ab = ba$, 则 G 为交换群 (Abel 群, 加法群)

此时: $a \cdot b = a + b, a^n = a + a + \dots + a = na, 0 \cdot a = 0$, 单位元用 0 表示

5.1.2 子群及判定

定义 5.1.2

设 G 是一个群, $S \subseteq G, S \neq \emptyset$, 若 S 关于 G 的运算也构成一个群, 则称 S 是 G 的子群.



如: $S = \{nk \mid \forall k \in \mathbb{Z}, n \text{ 是固定的整数}\}$

判定 1: 设 G 是一个群, $S \subseteq G, S \neq \emptyset$, 若 S 满足:

- (1) $\forall a, b \in S, ab \in S$; (2) $\forall a \in S$, 有 $a^{-1} \in S$; 则称 S 是 G 的子群.

判定 2: 设 G 是一个群, $S \subseteq G, S \neq \emptyset$, 若对 $\forall a, b \in S$, 有 $ab^{-1} \in S$, 则称 S 是 G 的一个子群.

5.1.3 群中元素的阶

定义 5.1.3

设 G 是一个群, $a \in G$, 使得 $a^m = e$ 的最小正整数 m 称为 a 的阶, 记作 $o(a) = m$ 或 $|a| = m$, 若这样的正整数 m 不存在, 则称 a 的阶为无穷大 (无限的).



例题 5.1.3 $(\mathbb{Z}, +)$, $\forall a \in \mathbb{Z}, a \neq 0$, a 的阶无限. $G = \{A = (a_{ij})_{2 \times 2} \mid A \text{ 可逆}, a_{ij} \in \mathbb{R}\}$, 则 G 关于乘法构成群,

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in G \quad T^2 = E, \quad o(T) = 2.$$

性质:

- (1) 设 $o(a) = m$, 则 $a^n = e \Leftrightarrow m \mid n$;
- (2) $o(a) = m, o(b) = n$, 且有 $ab = ba$, 若 $(m, n) = 1$, 则 $o(ab) = mn$;
- (3) $o(a) = m$, 则 $o(a^k) = \frac{m}{(m, k)}$, 且 $o(a^k) = m \Leftrightarrow (k, m) = 1$ (k 为正整数).

5.1.4 循环群

定义 5.1.4

设 G 是一个群, $o(a) = n$, 若 $G = \{e, a, a^2, \dots, a^{n-1}\}$, 则称 G 为 n 阶循环群. 若 $o(a) = \infty$, 则 $G = \{e, a^{\pm 1}, a^{\pm 2}, \dots, a^{\pm n}, \dots\} = \{a^n \mid n \in \mathbb{Z}\}$, 则称 G 为无限阶循环群, 记作 $G = \langle a \rangle$.



例题 5.1.4 $(\mathbb{Z}, +)$ 是无限循环群; (\mathbb{Z}_n, \oplus) 为 n 阶循环群.

结论: 循环群的子群也是循环群.

5.2 环与域

5.2.1 环的概念

定义 5.2.1

$R \neq \emptyset, +$, 满足

- (1) R 对加法构成一个交换群; 0 是零元;
- (2) $\forall a, b, c \in R, (ab)c = a(bc)$;
- (3) $a(b+c) = ab+ac, (b+c)a = ba+ca$.

则称 R 关于二元运算 $+$, 运算构成一个环, 记作 $(R, +, \cdot)$.

若对 $\forall a, b \in R$, 有 $ab = ba$, 则称 R 为交换环.



例题 5.2.1 $(\mathbb{Z}, +, \cdot), (\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ 都是交换环. (无限)

例题 5.2.2 $(\mathbb{Z}_n, \oplus, \odot)$ 构成一个交换环.(有限)

$$a \oplus b = (a + b)(\text{mod } n) \quad a \odot b = ab(\text{mod } n)$$

5.2.2 域

定义 5.2.2

R 是一个交换环, $R^* = \{a \in R \mid a \neq 0\}$, 若 R^* 关于环 R 的乘法运算作成一个交换群, 则称 R 为一个域. 加法的单位元称为零元 0 , R^* 乘法的单位元 1 .



例题 5.2.3 $(\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ 都是域. $(\mathbb{Z}_n, \oplus, \odot)$ 是域 $\Leftrightarrow n = p$ (p 是素数)

定义 5.2.3

设 $(R, +, \cdot)$ 是环, $(F, +, \cdot)$ 是域,

- (1) $R_1 \subseteq R, R_1 \neq \emptyset$, 若 R_1 构成一个环, 则称 R_1 为 R 的子环, R 称为 R_1 的扩环.
- (2) $F_1 \subseteq F, F_1 \neq \emptyset$, 若 F_1 构成一个域, 则称 F_1 为 F 的子域, F 称为 F_1 的扩域.



定理 5.2.1

设 $(R, +, \cdot)$ 为环, $(F, +, \cdot)$ 为域, 则

- (1) R_1 为 R 的子环 $\Leftrightarrow \forall a, b \in R_1, a - b \in R_1, a \cdot b \in R_1$;
- (2) F_1 为 F 的子域 $\Leftrightarrow \forall a, b \in F_1, a - b \in F_1, ab^{-1} \in F_1$.



5.3 理想和商环

5.3.1 理想

定义 5.3.1

设 I 是环 R 的子环, 若对 $\forall a \in I, x \in R$, 有 $a \cdot x \in I, x \cdot a \in I$, 则称 I 是 R 的理想.



例题 5.3.1 $(\mathbb{Z}, +, \cdot) \quad I = \{nk \mid k \in \mathbb{Z}, n \text{ 是固定整数}\}, I$ 是 \mathbb{Z} 的理想. 平凡理想: $\{0\}, R$.

判定: 设 $(R, +, \cdot)$ 是一个环, $I \subseteq R, I \neq \emptyset, I$ 是 R 的理想 \Leftrightarrow

- (1) $\forall a, b \in I, a - b \in I$;
- (2) $\forall a \in I, x \in R$, 有 $a \cdot x \in I, x \cdot a \in I$.

定义 5.3.2

设 $(R, +, \cdot)$ 为环, $a \in R$, R 的包含 a 的最理想称为由 a 生成的主理想, 记作 $\langle a \rangle$.



定理 5.3.1

若 $(R, +, \cdot)$ 是有单位元的交换环, 则 $\langle a \rangle = \{ra \mid \forall r \in R\}$.



5.3.2 商环

定义 5.3.3

设 $(R, +, \cdot)$ 为一个环, I 是 R 的一个理想, 定义 $a + I = \{a + i \mid \forall i \in I\}$, 称为 $/$ 的一个陪集, a 为代表元. 记 $R/I = \{a + I \mid \forall a \in R\}$



注:

- (1) $a + I = b + I \Leftrightarrow a - b \in I$;
- (2) $a + I$ 与 $b + I$ 要么相等, 要么交为 \emptyset .

定义 5.3.4

在 R/I 上定义两个二元运算:

- (1) $(a + I) \oplus (b + I) = (a + b) + I$;
- (2) $(a + I) \odot (b + I) = ab + I$;

则 R/I 关于定义的 \oplus, \odot 作成环, 称为商环.



5.3.3 环 (域) 的同构

设 R, S 是两个环 (或域), 若存在一个 1-1 映射 $f: R \rightarrow S$ 满足:

- (1) $\forall a, b \in R$, 有 $f(a + b) = f(a) + f(b)$;
- (2) $\forall a, b \in R$, 有 $f(ab) = f(a)f(b)$; 则称 f 为同构映射, 此时称 S 与 R 同构, 记作 $R \cong S$.

例题 5.3.2 $(\mathbb{Z}, +, \cdot)$ 是环, $I = \{nk \mid \forall k \in \mathbb{Z}, n \text{ 为固定整数}\}$, 则: $\mathbb{Z}/I \cong (\mathbb{Z}_n, \oplus, \odot)$

$$f: \mathbb{Z}_n \rightarrow \mathbb{Z}/I$$

$$x \mapsto x + I$$

$$f(x) = f(y) \Rightarrow x + I = y + I \Rightarrow x - y \in I,$$

即 $x - y = nk_1, n \mid x - y \Rightarrow x = y$ f 是单射, 显然 f 是满射.

$$f(x + y) = (x + y) + I = (x + I) + (y + I) = f(x) + f(y)$$

$$f(x \cdot y) = (x \cdot y) + I = (x + I) \cdot (y + I) = f(x) \cdot f(y)$$

故 f 为同构映射.

5.4 域上的多项式环

5.4.1 域上的多项式环

1. 域上的多项式

$(F, +, \cdot)$ 是一个域, x 是一个文字 (未定元), 称 $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$,

$(a_i \in F, i = 0, 1, \cdots, n)$ 为域 F 上的一元多项式, 若 $a_n \neq 0$, 定义 $\deg(f(x)) = n$, a_n 为首项系数.

2. 域上的多项式环

记 $F[x] = \{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \mid (a_i \in F, i = 0, \cdots, n)\}$, 在 F 上定义 $+$, 则 $F[x]$ 构成一交换环 (不是域) .

5.4.2 带余除法

定理 5.4.1

$f(x), g(x) \in F[x], \exists$ 唯一的 $q(x), r(x)$ 有

$$f(x) = g(x)q(x) + r(x)$$

其中 $\partial(r(x)) < \partial(g(x))$ 或 $r(x) = 0$, 记 $r(x) = [f(x)]_{g(x)}$



定义 5.4.1 (同余)

$a(x), b(x), p(x) \in F[x]$, 若 $p(x) \mid a(x) - b(x)$, 则称 $a(x)$ 与 $b(x)$ 模 $p(x)$ 同余, 记作 $a(x) \equiv b(x) \pmod{p(x)}$.



例题 5.4.1 $F = F_2 = \{0, 1\}$, 取 $a(x) = x^5 + x^4 + x^2 + 1, b(x) = x^3 + x + 1$, 则 $q(x) = x^2 + x + 1, r(x) = x^2$.

5.4.3 最大公因式与最小公倍式

1. 最大公因式;
2. 求法: 辗转相除法;
3. 互素;
4. 最小公倍式.

5.4.4 不可约多项式

定义 5.4.2

$p(x) \in F[x], \deg(p(x)) \geq 1$, 若 $p(x)$ 的因式只能为 $a \in F$ 或 $cp(x)$, 则称 $p(x)$ 为不可约多项式, 否则称其为可约多项式.



性质:

- (1) $p(x)$ 不可约, $p(x) \mid a(x)b(x)$, 则有 $p(x) \mid a(x)$ 或 $p(x) \mid b(x)$;
- (2) $p(x)$ 不可约, 则对 $\forall a(x) \in F[x]$, 有 $p(x) \mid a(x)$ 或 $(p(x), a(x)) = 1$;

标准分解式: $f(x) \in F[x], \partial(f(x)) \geq 1$, 则 $f(x)$ 可以唯一地表示成

$$f(x) = bp_1(x)^{k_1} \cdots p_s(x)^{k_s}$$

其中 $p_i(x) (i = 1, \cdots, s)$ 为不可约多项式, 且首项系数为 1 .

5.4.5 不可约多项式与有限域的构造

设 $p(x)$ 为不可约多项式, $\langle p(x) \rangle = \{p(x)a(x) \mid \forall a(x) \in F[x], \deg(p(x)) = n, F[x]/\langle p(x) \rangle$ 是商环, $\forall f(x) \in F[x], f(x) = g(x)p(x) + r(x) (\partial(r(x)) < \partial(p(x)) \text{ 或 } r(x) = 0)$, 记 $F[x]_{p(x)} = \{a_0 + a_1x + \cdots + a_{n-1}x^{n-1} \mid \forall a_i \in F, i = 0, \cdots, n-1\}$, 则 $F[x]_{p(x)}$ 关于下述定义加法和乘法作成交换环,

$$a(x) \oplus b(x) = a(x) + b(x)$$

$$a(x) \odot b(x) = (a(x) \cdot b(x))_{p(x)}$$

则 $F[x]/\langle p(x) \rangle \cong F[x]_{p(x)}$

证明

$$\varphi: F[x]/\langle p(x) \rangle \rightarrow F[x]_{p(x)}$$

$$f(x) + \langle p(x) \rangle \mapsto (f(x))_{p(x)}$$

$$\begin{aligned} \varphi(f(x) + \langle p(x) \rangle + g(x) + \langle p(x) \rangle) &= \varphi(f(x) + g(x) + \langle p(x) \rangle) \\ &= (f(x) + g(x))_{p(x)} = (f(x))_{p(x)} + (g(x))_{p(x)} \end{aligned}$$

φ 是一个同构映射.

$F[x]/\langle p(x) \rangle$ 是域 $\Leftrightarrow p(x)$ 不可约; $F[x]_{p(x)}$ 是域 $\Leftrightarrow p(x)$ 不可约; $\forall a(x) \in F[x]_{p(x)}, a(x) \neq 0$, 即 $a(x) \nmid p(x), (a(x), p(x)) = 1$. $(a(x), p(x)) = 1 \Leftrightarrow \exists u(x), v(x)$ 有 $a(x)u(x) + p(x)v(x) = 1 \Leftrightarrow$ 有 $(a(x)u(x))_{p(x)} = 1, a(x)^{-1} = u(x)$. 故 $F[x]_{p(x)}$ 是一个域, 即 $F[x]/\langle p(x) \rangle$ 是一个域.

$$F[x]_{p(x)} = \{a_0 + a_1x + \cdots + a_{n-1}x^{n-1} \mid \forall a_i \in F, i = 0, \cdots, n-1\},$$

若 $|F| = q$, 则 $|F[x]_{p(x)}| = q^n$; 特别地 $|F| = p$, 则 $|F[x]_{p(x)}| = p^n$.

例题 5.4.2 $p(x) = x^2 + x + 1, p(x)$ 是 F_2 上的不可约多项式, $F[x]_{p(x)} = \{a_0 + a_1x \mid a_0, a_1 \in F_2\}$ 是域. $F[x]_{p(x)} = \{0, 1, 1+x, x\}$

5.4.6 重因式及多项式的根

$p(x) = x^2 + x + 1, 0, 1$ 不是 $p(x)$ 的根.

5.5 有限域

5.5.1 有限域

定义 5.5.1

设 F 是一个域, 若 F 含有限个元素, 则称 F 为有限域, 若 $|F| = q$, 则记为 F_q .



例题 5.5.1 $(\mathbb{Z}_p, \oplus, \odot), \mathbb{Z}_p = \{0, 1, 2, \cdots, p-1\}, p$ 为素数, 构成一个有限域.

$$a \oplus b = (a + b) \pmod{p}$$

$$a \odot b = ab(\bmod p)$$

若 p 不是素数, 则 Z_p 不是域.

5.5.2 域的特征

定义 5.5.2

设 F 是一个域, e 为 F 的单位元, 若对任意的正整数 m , 有 $me \neq 0$, 则称 F 的特征为 0; 若存在正整数 m , 有 $me = 0$, 则满足该条件的最小正整数称为 F 的特征.



例题 5.5.2 Q, R, C 的特征为 0, Z_p (p 为素数) 特征为 p .

性质 1: 设 F 为有限域, 则 F 的特征为素数.

证明 设 p 为 F 的特征, 假设 p 不是素数, 则 $p = p_1 p_2$ ($p_1 < p, p_2 < p$), 于是 $pe = p_1 p_2 e = (p_1 e)(p_2 e) = 0 \Rightarrow p_1 e = 0$ 或 $p_2 e = 0$, 而 $p_1 < p, p_2 < p$, 与 p 是 F 的特征矛盾, 故 p 为素数.

性质 2: 设 F 为有限域, 若 F 的特征为 p , 则对 $\forall a \in F, a \neq 0$, 有 $pa = 0$ 且 p 为 a 的加法阶.

证明 设 p 为 F 的特征, 则有 $pe = 0$, 对 $\forall a \in F, a \neq 0, a = a \cdot e = e \cdot a$ 则有 $pa = pe \cdot a = 0 \cdot a = 0$. 假设 $o(a) = m$, 则 $m < p$, 有 $ma = 0, 0 = ma = mea = (me)a, a \neq 0 \Rightarrow me = 0$, 而 $m < p$ 与 p 是 e 的阶矛盾, 故对 $\forall a \in F, a \neq 0, pa = 0$.

5.5.3 素域

定义 5.5.3

设 F 为有限域, 称 F 的最小子域为 F 的素域, 即 F 的素域是 F 的所有子域的交集.



设 p 是有限域 F 的特征, 记 $\pi = \{0, e, 2e, \dots, (p-1)e\}$ 则 π 是 F 的最小子域, 事实上, 对 $\forall a, b \in \pi$,

$$a = ke, b = \ell e, \ell \neq 0$$

$$a - b = ke - \ell e = (k - \ell)e = (k - \ell)(\bmod p)e \in \pi$$

$$ab^{-1} = k\ell^{-1}e = k\ell^{-1}(\bmod p)e \in \pi$$

而 e 含于 F 的任意子域, 故 π 含于 F 的任一子域中, 即 $\forall F_1$ 为 F 的子域, $\pi \subseteq F_1$, 从而 π 是 F 的素域.

事实上, 因 π 包含于 F 的任一子域, 不妨设为 F_1, F_2, \dots, F_n , 故 $\pi \subseteq \bigcap_{i=1}^n F_i$, 又 $\pi \subseteq \bigcap_{i=1}^n F_i$ 为 F 的最小子域, 而 π 是 F 的子域, 不妨设 $\pi = F_j$, 故 $\bigcap_{i=1}^n F_i \subseteq \pi$, 故 $\pi = \bigcap_{i=1}^n F_i$.

结论:

$$\pi \cong Z_p = \{0, 1, 2, \dots, p-1\}$$

$$f: Z_p \rightarrow \pi$$

$$k \mapsto ke$$

5.5.4 有限域 F_q 的性质

运算性质:

(1) 设 p 是有限域 F 的特征, 则 $(a \pm b)^p = a^p \pm b^p$.

证明

$$(a+b)^p = a^p + b^p + \sum_{i=1}^{p-1} C_p^i a^i b^{p-i}$$

下面证明 $p \mid C_p^i$, $C_p^i = \frac{p!}{(p-i)!i!}$. 即 $p! = C_p^i (p-i)!i!$ $p \nmid (p-i)!, p \nmid i!$, 故 $p \nmid (p-i)!i!$ (p 是素数). 而 $p \mid p!$, 即 $p \mid C_p^i (p-i)!i!$, 故 $p \mid C_p^i$. 从而 $C_p^i a^i b^{p-i} = 0$ 成立.

(2) 设 p 是有限域 F 的特征, 则 $\left(\sum_{i=1}^m a_i\right)^p = \sum_{i=1}^m a_i^p, (a \pm b)^{p^n} = a^{p^n} \pm b^{p^n}$.

引理 5.5.1

设 (G, \cdot) 有限交换群, n 是 G 中所有元素阶数的最大值, 则 G 中所有元素的阶数是 n 的因子.



证明 设 $a \in G$, a 的阶为 n , 即 $a^n = e$, 对 $\forall b \in G$, 设 b 的阶为 m , 证明 $m \mid n$, 假设 $m \nmid n$

$$m = p_1^{e_1} p_2^{e_2} \cdots, p_s^{e_s}, n = p_1^{e'_1} p_2^{e'_2} \cdots, p_s^{e'_s}$$

$m \nmid n$, 则一定存在 p_i , 有 $e_i > e'_i$ 不妨设为 p_1 , 即 $e_1 > e'_1$

$$\text{令 } m = p_1^{e_1} m_1, \quad n = p_1^{e'_1} n_1$$

$$e_1 > e'_1 \quad (p_1, m_1) = 1 \quad (p_1, n_1) = 1$$

$a^{p_1^{e'_1}}$ 的阶数为 $\frac{n}{\binom{n}{n, p_1^{e'_1}}} = \frac{n}{p_1^{e'_1}} = n_1$, 又 $(n_1, p_1^{e_1}) = 1$

定理 5.5.1

设 F 为有限域, $F^* = F \setminus \{0\}$, 则 (F^*, \cdot) 是一个循环群.



证明 设 n 是 F^* 中元素的最大阶, 则对 $\forall a \in F^*, o(a) \mid n$, 故对 $\forall a \in F^*, a^n = 1$. 设 $\alpha \in F^*, \alpha$ 的阶数为 n , 令 $\langle \alpha \rangle = G = \{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$, 下面证明 $G = F^* = \langle \alpha \rangle$. 设 $|F| = q$, 则 $|F^*| = q-1, G \subseteq F^*$, 故 $q-1 \geq n$; 又令 $f(x) = x^n - 1, f(x)$ 在 F 上至多有 n 个根, 而 $\forall a \in F^*$, 均有 $a^n = 1$, 即 $a^n - 1 = 0$, 即 F^* 中的 $q-1$ 个元素均为 $f(x)$ 的根, 故有 $q-1 \leq n$, 从而 $q-1 = n, F^* = G, F^*$ 为循环群.

定义 5.5.4

设 F 为有限域, 乘法群 F^* 的生成元称为 F 的本原元.



定理 5.5.2

设 F_1 是有限域 F 的子域, 并且 $|F_1| = q$, 则一定存在正整数 n , 使得 $|F| = q^n$.



证明 设 $F_1 = F$, 则结论显然成立.

若 $F_1 \subset F, \exists e_1 \in F$, 但 $e_1 \notin F_1$, 令 $F_2 = \{a_1 + a_2 e_1 \mid a_1, a_2 \in F_1\}$, $|F_1| = q$, 故 $|F_2| = q^2$, 事实上, 只需说明 $a_1 + a_2 e_1 = b_1 + b_2 e_1 (b_1, b_2, a_1, a_2 \in F_1) \Leftrightarrow a_1 = b_1, a_2 = b_2$ 即可.

$$a_1 + a_2 e_1 = b_1 + b_2 e_1 \Rightarrow (a_1 - b_1) + (a_2 - b_2) e_1 = 0$$

即 $(a_2 - b_2) e_1 = b_1 - a_1 \Rightarrow (b_1 - a_1) (a_2 - b_2)^{-1} = e_1 \in F_1$ 矛盾;

若 $F_2 \neq F, \exists e_2 \in F$ 但 $e_2 \notin F_2$ 令 $F_3 = \{a_1 + a_2 e_1 + a_3 e_2 \mid a_1, a_2, a_3 \in F_1\}$, 则 $|F_3| = q^3$; 依次下去, 因 F 是有限域, 故必存在 n 使得 $|F| = q^n$.

推论 5.5.1

设 p 为有限域 F 的特征, 则必存在正整数 n , 使得 $|F| = p^n$.



证明 $\pi = \{0, e, 2e, \dots, (p-1)e\}$ 为 F 的子域.

定理 5.5.3

任意两个元素个数相同的有限域一定同构.

$$f: \rightarrow F'$$

$$0 \mapsto 0'$$

$$\alpha \mapsto \beta \text{ (其中 } \alpha \text{ 为 } F^* \text{ 的本原元, } \beta \text{ 为 } F'^* \text{ 的本原元)}$$



推论 5.5.2

设 F 是有限域, $Z_p = \{0, 1, \dots, p-1\}$, p 为素数, $p(x)$ 是 Z_p 上的不可约多项式,

(1) 如果 $|F| = p^n$, 则 $F \cong Z_p[x]/\langle p(x) \rangle$;

(2) 如果 $|F| = p$, 则 $F \cong Z_p$.



5.5.5 极小多项式与本原多项式

定义 5.5.5

设 F 为有限域, F_q 为 F 的含有 q 个元素的子域, $\alpha \in F, F_q$ 上的以 α 为根, 并且首项系数为 1 的次数最低的多项式称为 α 在 F_q 上的极小多项式.



定理 5.5.4

设 F 为有限域, F_q 为 F 的含有 q 个元素的子域, $\alpha \in F$, 则 α 在 F_q 上的极小多项式存在, 是唯一的, 并且是 F_q 上的不可约多项式.



例题 5.5.3 $f(x) = x^2 + x + 1$

$f(x)$ 在 F_2 上不可约 (无根), 在 F_4 中有两个根.

$f(x)$ 在 \mathbb{R} 中不可约, $\alpha = \frac{-1+\sqrt{3}i}{2}$

$f(\alpha) = 0 \quad \alpha \in \mathbb{C}$ 但 $\alpha \notin \mathbb{R}$, $f(x)$ 为 α 在 \mathbb{R} 上的极小多项式.

定义 5.5.6

定义: 设 F 为有限域, F_q 为 F 的含有 q 个元素的子域, $f(x)$ 为 F_q 上的不可约多项式. 如果 $f(x)$ 的根都是 F 的本原元, 则称 $f(x)$ 为本原多项式.



例题 5.5.4 $F_2[x]$ 中的 4 次本原多项式为 $f(x) = x^4 + x + 1$; $F_2[x]$ 中的 $f(x) = x^3 + x + 1$ 是本原多项式, $f(x)$ 的根在 F_8 中, 每个根的阶均为 7.

定理 5.5.5

设 F 为有限域, F_q 为 F 的含有 q 个元素的子域, α 是 F 的本原元, $|F| = q^n$, 则 α 在 F_q 上的极小多项式为 n 次多项式

$$f(x) = (x - \alpha)(x - \alpha^q)(x - \alpha^{q^2}) \cdots (x - \alpha^{q^{n-1}})$$

进一步 $\alpha, \alpha^q, \dots, \alpha^{q^{n-1}}$ 均为 F 的本原元.



注: 极小多项式不一定是本原多项式.

5.6 域上的线性代数

5.6.1 域上的向量空间

定义 5.6.1

设 F 是一个域, $V \neq \emptyset$,

$$+ : V \times V \rightarrow V$$

$$\cdot : F \times V \rightarrow V$$

且满足:

- (1) $(V, +)$ 是一个交换群;
- (2) 对 $\forall a \in F, v_1, v_2 \in V, a(v_1 + v_2) = av_1 + av_2$;
- (3) 对 $\forall a_1, a_2 \in F, v \in V, (a_1 + a_2)v = a_1v + a_2v$;
- (4) 对 $\forall v \in V$, 有 $1 \cdot v = v$;

则称 V 为 F 上的向量空间.

**向量空间的基与维数:**

V 是域 F 上的向量空间, $e_1, e_2, \dots, e_n \in V$, 若对 $\forall v \in V, v$ 可唯一地表示为 $v = c_1e_1 + c_2e_2 + \dots + c_ne_n$, 其中 $c_i \in F, i = 1, 2, \dots, n$, 则称 e_1, e_2, \dots, e_n 为 V 的一组基, V 称为 F 上的 n 维向量空间.

线性相关与线性无关:

设 V 是域 F 上的线性空间, $v_1, v_2, \dots, v_r \in V$, 如果存在不全为 0 的 c_1, c_2, \dots, c_r 使得:

$$c_1v_1 + c_2v_2 + \dots + c_rv_r = 0,$$

则称 v_1, v_2, \dots, v_r 为线性相关, 否则 v_1, v_2, \dots, v_r 线性无关.

定理 5.6.1

设 V 是域 F 上的 n 维向量空间, 则 V 的任意一组基都是线性无关的.



证明 设 e_1, \dots, e_n 是 V 的一组基, 假设 e_1, e_2, \dots, e_n 线性相关, 则存在不全为 0 的 c_1, c_2, \dots, c_n , 使得:

$$c_1 e_1 + c_2 e_2 + \dots + c_n e_n = 0$$

$0e_1 + 0e_2 + \dots + 0e_n = 0$ (0 向量有两种表达形式) 与 e_1, e_2, \dots, e_n 是基矛盾.

例题 5.6.1 $V(n, q) = \{(a_1, a_2, \dots, a_n) \mid a_i \in F_q, i = 1, 2, \dots, n\}$, F_q 为 q 元有限域.

$a = (a_1, \dots, a_n), b = (b_1, \dots, b_n)$. 定义: $a + b = (a_1 + b_1, \dots, a_n + b_n) \in V(n, q)$

$\lambda \cdot a = (\lambda a_1, \dots, \lambda a_n) \in V(n, q) \quad (\lambda \in F_q)$. $e_i = (0, \dots, 0, 1, 0, \dots, 0) i = 1, \dots, n$ 是它的一组基, $\forall a \in V(n, q), a = (a_1, \dots, a_n), a = \sum_{i=1}^n a_i e_i$. $V(n, q)$ 为 F_q 上的 n 维向量空间.

定理 5.6.2

设 F_q 是有限域 F 的含 q 个元素的子域, 且 $|F| = q^n$, 则 F 是 F_q 上的 n 维向量空间.



证明 设 α 是 F 的本原元, α 在 F_q 上的极小多项式为 $g(x)$ (不可约). $|F| = q^n$, 所以 $\deg(g(x)) = n$ ($F \cong F_q[x]/\langle g(x) \rangle$). 下面证明 $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$ 是 F 的一组基. 首先证 $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$ 线性无关; 假设存在不全为 0 的 $c_0, c_1, \dots, c_{n-1} \in F_q$ 使得:

$$c_0 + c_1 \alpha + c_2 \alpha^2 + \dots + c_{n-1} \alpha^{n-1} = 0$$

$$\text{令 } f(x) = c_0 + c_1 x + c_2 x^2 + \dots + c_{n-1} x^{n-1},$$

则 $f(x) \in F_q[x]$ 并且有 $f(\alpha) = 0$, 因此 $g(x) \mid f(x)$ (极小多项式必整除零化多项式), $f(x) \neq 0$ 且 $\deg(f(x)) = n-1 < \deg(g(x)) = n$ 与 $g(x)$ 是极小多项式矛盾.

下证 F 的任一元素可由其线性表示. α 是 F 的本原元, 并且 $|F| = q^n$, 则有:

$$F = \{0, 1, \alpha, \alpha^2, \dots, \alpha^{q^n-2}\}$$

$$\text{设 } x^i = q(x)g(x) + r(x), \quad q(x), r(x) \in F_q[x],$$

$$\deg(r(x)) < \deg(g(x)) = n, i = 0, 1, \dots, q^n - 2,$$

于是有 $\alpha^i = r(\alpha)$

$$\text{设 } r(x) = r_0 + r_1 x + r_2 x^2 + \dots + r_{n-1} x^{n-1} r_j \in F_q, j = 0, 1, \dots, n-1$$

则有:

$$\alpha^i = r(\alpha) = r_0 + r_1 \alpha + r_2 \alpha^2 + \dots + r_{n-1} \alpha^{n-1} (i = 0, 1, \dots, q^n - 2)$$

即 F^* 中的每个元可由 $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$ 线性表示, 0 显然可由 $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$ 线性表示, 故 $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$ 是 F 的一组基, F 是 F_q 上的 n 维向量空间.

5.6.2 极大线性无关组

5.6.3 域 F 上的 $m \times n$ 矩阵

1. 定理: $A_{m \times n}$ 的行秩, 列秩, 秩;
2. 运算: $A + B, AB, kA$
3. 初等行变换及性质;
4. 可逆矩阵;
5. 可逆充要条件: A 可逆 $\Leftrightarrow \text{rank}(A) = n$;
6. 线性方程组和行列式理论.

第 6 章 编码理论的基本知识

为了保证通信系统能够准确的传输信息,我们要对信源消息进行编码,代数编码能够起到降低通信中传输误差的作用.即可在信道接收端实现自动地纠错和检错.本章介绍编码理论的一些基本知识.

6.1 码的基本概念

记输入信号字母集和输出信号字母集为 $\mathcal{U} = \mathcal{V} = F_q$, F_q 是 q 元域, $q = p^m$, p 是素数. $V(n, q) = F_q^n$ 表示 q 元域 F_q 上的 n 维向量, 信息输入和输出字符串用 F_q^n 中的元素表示. F_q^n 中的元素 (z_1, z_2, \dots, z_n) 表示为 $z_1 z_2 \dots z_n$.

6.1.1 码的定义

定义 6.1.1

$C \subseteq V(n, q)$, $C \neq \emptyset$, 则称 C 为 q 元分组码, n 称为码长, C 中的向量称为码字. $M = |C|$ 为码字的个数. 此时称 C 为 q 元 (n, M) 码.



定义 6.1.2

设 C 为 q 元 (n, M) 码, 定义 $R(C) = \frac{\log_q M}{n}$, 称为码 C 的码率.



例题 6.1.1 $q = 2, n = 3, V(n, q) = \{000, 001, 010, 011, 100, 101, 110, 111\}$. 取 $C = \{000, 101, 111\}$, 显然 C 为二元分组码, 码长 $n = 3$. 且 $|C| = 3 = M$, 因此 C 为二元 $(3, 3)$ 码.

$$\text{码率 } R(C) = \frac{\log_2 3}{3}$$

注: $R(C)$ 的意义: 传输的最大平均信息量, n 长的码字所携带的最大信息量为 $(\mathcal{U} = F_q) n \log_q q = \log_q q^n \triangleq \log_q M$, 平均信息量 $\frac{\log_q q^n}{n}$, 将 q^n 换为 M , 即为 $\frac{\log_q M}{n}$.

6.1.2 Hamming 距离和 Hamming 重量

(用于后面的译码、纠错和检错)

定义 6.1.3

设 $x, y \in V(n, q)$, x, y 的 Hamming 距离 $d(x, y)$ 定义为 x 和 y 中不同分量的个数, 即

$$d(x, y) = \sum_{j=1}^n d(x_j, y_j)$$

其中

$$d(x_j, y_j) = \begin{cases} 0 & \text{如果 } x_j = y_j \\ 1 & \text{如果 } x_j \neq y_j \end{cases}$$



注: $d(x, y) : V(n, q) \times V(n, q) \rightarrow N$ 的映射, N 为全体非负整数的集合.

定义 6.1.4

设 $x \in V(n, q)$, 称 x 中非零分量的个数为 x 的 Hamming 重量, 记为 $\omega(x)$.



例题 6.1.2 $x = 12112 \in V(5, 3), y = 10201 \in V(5, 3)$, 则 $d(x, y) = 4, \omega(x) = 5, \omega(y) = 3$

Hamming 距离的性质

定理 6.1.1

对 $\forall x, y, z \in V(n, q)$, 即 Hamming 距离 $d(x, y)$ 满足下列性质:

- (1) 非负性 $d(x, y) \geq 0, d(x, y) = 0 \iff x = y$
- (2) 对称性 $d(x, y) = d(y, x)$
- (3) 三角不等式 $d(x, y) \leq d(x, z) + d(z, y)$



证明 (1)、(2) 显然成立, 只需证明 (3) 成立.

设 $x = x_1x_2 \cdots x_n, y = y_1y_2 \cdots y_n, z = z_1z_2 \cdots z_n$, 对 $\forall j = 1, 2, \cdots, n$

若 $x_j = y_j$, 则 $d(x_j, y_j) = 0$. 于是有 $d(x_j, y_j) \leq d(x_j, z_j) + d(z_j, y_j)$, 而 $d(x, y) = \sum_{j=1}^n d(x_j, y_j), d(x, z) = \sum_{j=1}^n d(x_j, z_j), d(y, z) = \sum_{j=1}^n d(y_j, z_j)$. 故 (3) 成立.

若 $x_j \neq y_j$ 则 $d(x_j, y_j) = 1$, 此时 $x_j \neq z_j$ 与 $y_j \neq z_j$ 至少有一个成立, 因此有 $d(x_j, y_j) \leq d(x_j, z_j) + d(z_j, y_j)$, 于是 (3) 成立.

注: $V(n, q)$ 中定义了 Hamming 距离, 称 $V(n, q)$ 为 Hamming 空间.

6.1.3 Hamming 距离 (重量) 与译码

(最大似然译码或最小 Hamming 距离译码)

分析: 对于二元对称信道

信道矩阵为 $\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$, 其中令 $p < \frac{1}{2}$, 设 x 为输入码字, y 为输出向量, 则信道发生

错误的字符个数等于 Hamming 距离 $d(x, y)$, 因此我们有

$$p(y | x) = p^{d(x, y)}(1-p)^{n-d(x, y)}$$

对于 Hamming 距离 d' , 若 $d < d'$, 则有 $(p < \frac{1}{2})$:

$$p^d(1-p)^{n-d} > p^{d'}(1-p)^{n-d'} \Rightarrow \text{(Hamming 距离小, 输出的概率大)}$$

对 $\forall y \in V(n, q)$, 存在 $x' \in C$, 使得 $d(x', y) \leq d(x, y)$. 此时有 $p(y | x') \geq p(y | x)$, 于是将收到的向量译为与其 Hamming 距离最近的码字, 即将 y 译为 x' .

例题 6.1.3 码长为 3 的二元重复码为 $C = \{000, 111\}$, 已知码字集合, 但不知发送的是哪一个, 设 000 是发送的码字, 则收到字 000, 100, 010, 001 时将被译成 000, 当收到字 111, 110, 101, 011 时将被译成 111, 这时对任意信道入口概率分布 $p(000) = p_0, p(111) = 1 - p_0$, 则译码错误概率为

$$\begin{aligned} e &= p_0(p(111 | 000) + p(110 | 000) + p(101 | 000) + p(011 | 000)) \\ &\quad + (1 - p_0)(p(000 | 111) + p(100 | 111) + p(001 | 111) + p(010 | 111)) \\ &= p_0(p^3 + 3p^2(1 - p)) + (1 - p_0)(p^3 + 3p^2(1 - p)) \\ &= 3p^2(1 - p) + p^3 = 3p^2 - 2p^3 \end{aligned}$$

6.1.4 系统码

设 $V(k, q) = \{(a_1, a_2, \dots, a_k)\}$, 则 q^k 个消息编码为 $V(q, k)$ 中的 q^k 个元素, 如此编码无纠错、检错能力, 故需编成 n 长的, 使其具有纠错、检错能力.

设 $\alpha = \{i_1, i_2, \dots, i_k\}$ 是一个正整数集合, $1 \leq i_1 < i_2 < \dots < i_k \leq n$. 设 $x = (x_1, x_2, \dots, x_n) \in V(n, q)$, 称 $x_\alpha = (x_{i_1}, x_{i_2}, \dots, x_{i_k})$ 为 x 的部分向量, 这时有 $x = (x_\alpha, x_{\alpha^c})$ 其中 $\alpha^c = N - \alpha$ 为 α 的余集, $N = \{1, 2, \dots, n\}$.

定义 6.1.5 (系统码)

设 C 是一个 q 元 (n, k) 码, 如果存在一个下标集合 $\alpha = \{i_1, i_2, \dots, i_k\}$, 使得码 C 中所有码字都去掉其余 $n - k$ 个位置后, 得到的向量的全体为 F_q 上长度为 k 的所有串的集合 $V(k, q)$, 即

$$C_\alpha = \{x_\alpha = (x_{i_1}, x_{i_2}, \dots, x_{i_k}) \mid x \in C\} = V(k, q)$$

则称 C 为具有 k 个信息位的 q 元系统码, $\{i_1, i_2, \dots, i_k\}$ 称为信息位, 其余 $n - k$ 个位置称为校验位或冗余位.



例题 6.1.4 二数码 $C = \{0000, 0110, 1001, 1010\}$ 是系统码

$$\begin{aligned} 00 &\longrightarrow \underline{0000} \\ 01 &\longrightarrow \underline{0110} \\ 10 &\longrightarrow \underline{1001} \\ 11 &\longrightarrow \underline{1010} \end{aligned} \quad i_1 = 1, i_2 = 3 \quad k = 2$$

译码过程可从码字的信息位上读出信源字符.

例题 6.1.5 二数码 $C = \{000, 100, 010, 001\}$ 不是系统码.

系统码的定义: 在错误更正编码中, 一个系统码通常被定义为能够将信息位 (或消息位) 和校验位 (或冗余位) 分开的编码, 其中信息位直接出现在码字中的固定位置上. 这样, 原始信息可以直接从码字的某个部分读取, 而无需进行任何转换.

解: 给定的码 $C = \{000, 100, 010, 001\}$ 包含 4 个码字, 每个码字由 3 位组成. 对于一个 q -元系统 (这里 $q = 2$, 即二进制), 如果我们尝试将 C 视为系统码, 我们需要确定信息位 k 和码字长度 n 的关系, 以及如何从码字中分离信息位和校验位.

$k = 1$ 的情况: 当我们假设信息位数 $k = 1$ 时, 理论上码 C 应包含 $q^k = 2^1 = 2$ 个码字, 因为一个信息位可以表示两个不同的值. 然而, 码 C 实际上包含 4 个码字, 这超出了 $k = 1$ 时的预期码字数量. 因此, 在 $k = 1$ 的情况下, 给定的码 C 不能满足系统码的要求. $k = 2$ 的情况: 对于 $k = 2$, 理论上码 C 应包含 $q^k = 2^2 = 4$ 个码字, 与 C 实际的码字数量匹配. 然而, 问题在于找不到一种合理的方式将任一位确定为校验位, 并使剩余的位作为信息位来满足系统码的要求. 无论选择哪个位作为校验位, 都无法创建一个满足所有情况下校验规则的系统, 特别是无法生成长度为 2 的信息位序列 (11), 同时满足给定的码字集.

因此, 无论是在 $k = 1$ 还是 $k = 2$ 的情况下, 给定的码 C 都不能被视为系统码. 在系统码中, 信息位和校验位之间应有明确的、固定的区分, 而给定的码 C 无法提供这样的区分, 因此它不符合系统码的定义.

6.2 码的检错和纠错能力

6.2.1 最小距离

定义 6.2.1

设 C 是一个 q 元 (n, M) 码, 码 C 的最小距离定义为

$$d(C) = \min\{d(x, y) \mid x, y \in C, x \neq y\}$$



注: 用记号 (n, M, d) 表示码长为 n , 码字个数为 M , 最小距离为 d 的码.

例题 6.2.1 码 $C = \{0000, 0110, 1001, 1010\}$, 则 $d(C) = 2$.

汉明距离 $d(C)$ 是码 C 中任意两个不同码字之间的最小汉明距离, 其中汉明距离是指两个码字在相同位置上不同符号的数量. 对于给定的码 $C = \{0000, 0110, 1001, 1010\}$, 我们需要计算所有码字对之间的汉明距离, 然后找出这些距离中的最小值来确定 $d(C)$. 计算码 C 中所有码字对之间的汉明距离: $d(0000, 0110) = 2$; $d(0000, 1001) = 3$; $d(0000, 1010) = 2$; $d(0110, 1001) = 4$; $d(0110, 1010) = 3$; $d(1001, 1010) = 2$. 从这些计算中, 我们可以看到最小的汉明距离是 2, 出现在多个码字对之间. 因此, 码 $C = \{0000, 0110, 1001, 1010\}$ 的最小汉明距离 $d(C)$ 是 2.

6.2.2 码的检错和纠错能力

定义 6.2.2

如果对于码 C 中的每一个码字, 当发生至多 t 个错误时, 检查出所产生的字不是码字, 则称码 C 为可检查 t 个错误的检错码. 如果能检查 t 个错误而不能检查 $t + 1$ 个错误, 则称码 C 为至多可检查 t 个错误的检错码.



定理 6.2.1

码 C 至多可检查 t 个错误的充分必要条件为 $d(C) = t + 1$



证明 \Leftarrow 充分性 ($d(C) = t + 1 \Rightarrow$ 码 C 至多可检查 t 个错误). 如果码 C 的最小汉明距离 $d(C) = t + 1$, 考虑以下情况:

对于码字 $x \in C$, 如果 x 发生的错误个数 $\leq t$, 即输入 x , 输出 y , 且 $d(x, y) \leq t$, 则 y 必定不在码 C 中, 因为从 x 到 y 的任何变化需要至少 $t + 1$ 个位的改变才能到达另一个有效码字. 因此, 可以检查出 y 必出错.

如果 x 发生的错误个数为 $t + 1$, 则 $d(x, y) = t + 1$. 这是最小汉明距离的情况, 意味着 y 可能是码 C 中的另一个码字, 因此无法仅通过检测 $t + 1$ 个错误来判断 y 是否出错.

\Rightarrow 必要性 (码 C 至多可检查 t 个错误 $\Rightarrow d(C) = t + 1$). 反过来, 如果码 C 至多可检查 t 个错误, 这意味着对于任意 $x \in C$ 和任意 $y \in V(n, q)$ ($V(n, q)$ 表示所有可能的 n 位长、基于 q 元字母表的向量空间), 当 $d(x, y) \leq t$ 时, y 必定不是码字. 这是因为如果 y 是码字, 则可能无法检测出 x 到 y 的 t 或更少的错误. 由此推断, 码 C 的最小距离必须大于 t , 否则不可能保证所有最多 t 个错误都能被检测出来. 因此, 有 $d(C) > t$. 如果 $d(C) = t + 2$ 或更大, 根据充分性, 码 C 将能至多检查 $t + 1$ 或更多个错误, 这与原假设矛盾. 因此, 最小汉明距离不能大于 $t + 1$, 即 $d(C) = t + 1$.

定义 6.2.3

当对码 C 采用最小 Hamming 距离译码时, 如果任意一个码字发生至多 t 个错误时, 都能正确译码, 则称码 C 为可纠正 t 个错误的纠错码. 如果 C 能纠正 t 个错误而不能纠正 $t + 1$ 个错误, 则称码 C 为至多可纠正 t 个错误的纠错码.



定理 6.2.2

码 C 至多可纠正 t 个错误的充分必要条件为 $d(C) = 2t + 1$ 或 $2t + 2$.



证明 \Leftarrow 先证充分性. 先证 $d(C) = 2t + 1$ 或 $2t + 2$ 时 C 至多可纠正 t 个错误, 对于 $x \in C$, 定义 $B_x(t) = \{y \in V(n, q) \mid d(x, y) \leq t\}$, 则对 $\forall x' \in C, x' \neq x, B_x(t) \cap B_{x'}(t) = \emptyset$, 否则若 $\exists z \in B_x(t) \cap B_{x'}(t)$, 则有 $d(z, x) \leq t, d(z, x') \leq t$, 从而

$$d(x, x') \leq d(x, z) + d(z, x') \leq 2t$$

与 $d(C) = 2t + 1$ 或 $2t + 2$ 矛盾. 故对 $y \in V(n, q), d(x, y) \leq t$, 对 $\forall x' \in C, x' \neq x$, 必有 $d(x', y) > t$. 于是将 y 译为 x , 即当 x 发生 $\leq t$ 个错误时可纠正.

再证此时 C 不能纠正 $t + 1$ 个错误. 事实上, 若 $d(C) = 2t + 1$, 设 $x, x' \in C, x = (x_1, \dots, x_n), x' = (x'_1, \dots, x'_n)$. $x_{i_1} \neq x'_{i_1}, x_{i_2} \neq x'_{i_2}, \dots, x_{i_{2t+1}} \neq x'_{i_{2t+1}}$, 其余情况 $x_j = x'_j$

取 $y = y_1 \cdots y_n$, 令

$$y_j = \begin{cases} x_j & \text{如果 } j = i_1, \dots, i_t \\ x'_j & \text{如果 } j = i_{t+1}, \dots, i_{2t+1} \\ x_j = x'_j & \text{否则} \end{cases}$$

则有 $d(x, y) = t + 1, d(x', y) = t$. 若 $d(C) = 2t + 2$. 重复上述过程, $\exists y$ 使得

$$d(x, y) = t + 1, \quad d(x', y) = t + 1$$

于是 $\exists y \in B_x(t+1) \cap B_{x'}(t+1)$, 故当收到 y 时不能确定将其译为那个码字, 因此码 C 不能纠正 $t+1$ 个错误.

\Rightarrow 设码 C 至多可纠正 t 个错误, 则必有 $d(C) > 2t$. 否则 $d(C) \leq 2t = 2(t-1) + 2$. 由充分性可知码 C 可纠正至多 $t-1$ 个错误, 矛盾. 另外 $d(C) \leq 2t+2$, 否则如果 $d(C) > 2t+2$, 则 $d(C) = 2t+3 = 2(t+1) + 1$. 则 C 可纠正 $t+1$ 个错误, 矛盾. 故 $d(C) = 2t+1$ 或 $2t+2$.

推论 6.2.1

$d(C) = d$ 的充分必要条件为码 C 至多可纠正 $\lfloor \frac{d-1}{2} \rfloor$ 个错误.



证明 $\lfloor \frac{d-1}{2} \rfloor$ 为不超过 $\frac{d-1}{2}$ 的最大正整数, 故

$$\left\lfloor \frac{d-1}{2} \right\rfloor = \begin{cases} \frac{d-1}{2} & d \text{ 为奇数} \\ \frac{d}{2} - 1 & d \text{ 为偶数} \end{cases}$$

于是 $\frac{d-1}{2} = t \Rightarrow d = 2t+1$, $\frac{d}{2} - 1 = t \Rightarrow d = 2t+2$

例题 6.2.2 称 $C = \{\underbrace{00 \cdots 0}_n, \underbrace{11 \cdots 1}_n, \dots, \underbrace{(q-1)(q-1) \cdots (q-1)}_n\}$ 为码长为 n 的 q 元重复码, $d(C) = n$, 码 C 是一个至多可纠正 $\lfloor \frac{n-1}{2} \rfloor$ 也是一个至多可检查 $n-1$ 个错误的检错码.

6.3 编码理论的基本问题

对于一个 q 元 (n, M, d) 码, 码率和码字个数以及码的最小距离都是衡量码的重要指标.

- (1) 码率大, 意味着冗余小, 码字的传输效率高.
- (2) 码字个数多, 意味着可以多发送信息.
- (3) 最小距离大, 意味着可以多纠正错误.

但是, 我们做不到同时让码率和码字个数以及最小距离都达到最优. 因此, 我们通常是固定其中的两个参数, 而让另外一个参数达到最优. 通常我们固定码长和最小距离, 而让码字个数达到最优.

编码理论的基本问题: 固定 n, d , 考虑 M 的最大值 (此时码率也大), 记 $A_q(n, d)$ 为所有 q 元 (n, M, d) 码中 M 的最大值, 编码理论的基本问题之一就是求出 $A_q(n, d)$, 并构造相应的 q 元 (n, M, d) 码. 对于简单的情形, 我们有下面的结论.

定理 6.3.1

定理: 对 $\forall n \geq 1, A_q(n, 1) = q^n, A_q(n, n) = q$



证明 令 $C = V(n, q)$, 则 C 是 q 元 $(n, q^n, 1)$ 码. 事实上, 至少可找到两个码字最小距离为 1, 如: $(1, 2, \dots, q-1)$ 和 $(2, 2, \dots, q-1)$.

$$A_q(n, d) \geq q^n (C \subseteq V(n, q)), \text{ 而 } |C| = q^n, \text{ 故 } A_q(n, 1) = q^n.$$

设 C 是一个 q 元 (n, M, n) 码, 则 $\forall x, y \in C, x = (x_1, \dots, x_n), y = (y_1, \dots, y_n), x_i \neq y_i, i =$

$1, \dots, n$, 因此, 所有码字在一个固定分量位置上出现的字符一定互不相同, 于是 $M \leq q$. 由此可知 $A_q(n, n) \leq q$, 又码长为 n 的 q 元重复码是一个 q 元 (n, q, n) 码, 故 $A_q(n, n) = q$.

6.3.1 码的等价变换

为了进一步讨论码的基本问题, 我们先介绍码的等价变换.

设 $A = \{a_1, a_2, \dots, a_n\}$ 为一个有限集合, 称 $A \rightarrow A$ 的一一映射为 A 上的一个置换.

$$\begin{aligned}\sigma: A &\rightarrow A \\ a_i &\mapsto \sigma(a_i) \quad i = 1, \dots, n\end{aligned}$$

$$\text{即 } \sigma = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ \downarrow & \downarrow & \cdots & \downarrow \\ \sigma(a_1) & \sigma(a_2) & \cdots & \sigma(a_n) \end{pmatrix}. \quad A = \{a_1, a_2, \dots, a_n\} = \{\sigma(a_1), \dots, \sigma(a_n)\}.$$

码的换位型置换与换元型置换码

定义 6.3.1 (换位型置换码)

设 C 为 q 元 (n, M, d) 码, 记 $\sigma_1 = \begin{pmatrix} 1 & 2 & \cdots & n \\ \downarrow & \downarrow & \cdots & \downarrow \\ \sigma_1(1) & \sigma_1(2) & \cdots & \sigma_1(n) \end{pmatrix}$ 称其为换位型置换. 对 $\forall x \in C$, 对 x 的分量坐标进行换位型置换, 即对 $\forall x = (x_1, \dots, x_n)$,

$$\sigma_1(x) = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ \downarrow & \downarrow & \cdots & \downarrow \\ x_{\sigma_1(1)} & x_{\sigma_1(2)} & \cdots & x_{\sigma_1(n)} \end{pmatrix} \quad (\text{一个置换})$$

记 $C_1 = \sigma_1(C) = \{\sigma_1(x) \mid x \in C\}$, 称之为码 C 的换位型置换码.



定义 6.3.2 (换元型置换码)

记 $\sigma_2 = \begin{pmatrix} 0 & 1 & \cdots & q-1 \\ \downarrow & \downarrow & \cdots & \downarrow \\ \sigma_2(0) & \sigma_2(1) & \cdots & \sigma_2(q-1) \end{pmatrix}$, 称 σ_2 为换元型置换, 记 $\bar{\sigma}_2 = (\sigma_{21}, \sigma_{22}, \dots, \sigma_{2n})$ 其中 $\sigma_{2j} (1 \leq j \leq n)$ 是换元型置换, 对 $\forall x = (x_1, x_2, \dots, x_n) \in C$,

$$\text{令 } \sigma_2(x) = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ \downarrow & \downarrow & \cdots & \downarrow \\ \sigma_{21}(x_1) & \sigma_{22}(x_2) & \cdots & \sigma_{2n}(x_n) \end{pmatrix}, \quad \sigma_{2j}: F_q \rightarrow F_q \quad (n \text{ 个置换}), \text{ 每个位置}$$

上对应一个置换记 $C_2 = \bar{\sigma}_2(C) = \{\bar{\sigma}_2(x) \mid x \in C\}$, 称之为码 C 的换元型置换码.



例题 6.3.1 三元 $(3, 3, 3)$ 码 $C = \{012, 120, 201\}$, 令 $\sigma_{2j} = \begin{pmatrix} 0 & 1 & 2 \\ \downarrow & \downarrow & \downarrow \\ 1 & 2 & 0 \end{pmatrix}$, 其中 $j = 1, 2, 3$, 则

$\sigma_{21} = \sigma_{22} = \sigma_{23}$, $\bar{\sigma}_2 = (\sigma_{21}, \sigma_{22}, \sigma_{23})$, 则 C 的换元型置换码 $C_2 = \{120, 201, 012\}$, 再

$$\text{令 } \sigma'_{2j} = \begin{pmatrix} 0 & 1 & 2 \\ \downarrow & \downarrow & \downarrow \\ 2 & 1 & 0 \end{pmatrix}, \bar{\sigma}'_2 = (\sigma'_{21}, \sigma'_{22}, \sigma'_{23}), \sigma'_{21} = \sigma'_{22} = \sigma'_{23},$$

则有 C 的换元型置换码为 $C'_2 = \{210, 102, 021\}$. 若令 $C = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}$, 则

$$C_2 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}, C'_2 = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

例题 6.3.2 $C = \{010, 101\}$, $\bar{\sigma}_2 = (\sigma_{21}, \sigma_{22}, \sigma_{23})$

$$\sigma_{21} = \begin{pmatrix} 0 & 1 \\ \downarrow & \downarrow \\ 1 & 0 \end{pmatrix} \quad \sigma_{22} = \begin{pmatrix} 0 & 1 \\ \downarrow & \downarrow \\ 0 & 1 \end{pmatrix} \quad \sigma_{23} = \begin{pmatrix} 0 & 1 \\ \downarrow & \downarrow \\ 1 & 0 \end{pmatrix}$$

于是 C 的换元型置换码 $C_2 = \{111, 000\}$, 可看出 C 与 C_2 的最小距离相同.

例题 6.3.3 设 $C = \{012, 210, 010\}$, $\sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 3 & 1 \end{pmatrix}$, 则码 C 的换位型置换码 $C_1 = \{201, 021, 001\}$.

设 $\sigma'_1 = \begin{pmatrix} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 3 & 1 & 2 \end{pmatrix}$, 则码 C 的换位型置换码 $C'_1 = \{120, 102, 100\}$, 若记 $C = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$,

则

$$C_1 = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}, C'_1 = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 1 & 0 & 0 \end{pmatrix}$$

可以看出 C 与 C_1, C'_1 的最小距离相同, 且它们是列的置换.

注:

- (1) 若设 σ_1^{-1} 与 $\bar{\sigma}_2^{-1}$ 分别是 σ_1 和 $\bar{\sigma}_2$ 的逆置换, 则有 $\sigma_1^{-1}(C_1) = C, \bar{\sigma}_2^{-1}(C_2) = C$.
- (2) 将 q 元 (n, M) 码排成一个 $M \times n$ 矩阵, 每个码字为一行

$$C = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{M1} & x_{M2} & \cdots & x_{Mn} \end{pmatrix}_{M \times n},$$

则码 C 的换位置换等价于矩阵的列的置换. 码 C 的换元置换等价于矩阵每一列中字符的置换.

结论: 码的换位置换和换元置换不会改变码的码长, 码字个数和最小距离.

定义 6.3.3

称换位型置换和换元型置换为码的等价变换. 对于两个 q 元码, 如果其中一个可以经过等价变换转化为另一个, 则称这两个 q 元码是等价的.

**引理 6.3.1**

任意一个 q 元 (n, M, d) 码 C 都等价于一个包含零码字 $\underbrace{00 \cdots 0}_{n \text{ 个}}$ 的 q 元 (n, M, d) 码.



证明 将 C 中的所有码字排成一个 $M \times n$ 矩阵, 每个码字一行, 对矩阵的每列分别做换元型置

$$\text{换, 总可以将矩阵的第一行变为零向量. 事实上, } C = \begin{pmatrix} x_{11} \\ \vdots \\ x_{i-1,1} \\ 0 \\ x_{i+1,1} \\ \vdots \\ x_{M1} \end{pmatrix} \quad \text{令 } \sigma_{21} = \begin{cases} x_{11} \\ \downarrow \\ 0 \end{cases}$$

例题 6.3.4 设 C 和 C' 为两个二元 $(5, 4, 3)$ 码.

$$C = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix} \quad C' = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix}$$

证明: C 与 C' 等价.

$$\text{证明: 对码 } C \text{ 的第 3 个坐标做换元置换, } \sigma_{23} = \begin{pmatrix} 0 & 1 \\ \downarrow & \downarrow \\ 1 & 0 \end{pmatrix}, C_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix}, \text{再}$$

$$\text{交换 3,4 两行 (码字的位置改变, 码字不变) 得到 } C'_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix} \text{再交换 2,4 列, 得}$$

$$C' = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix}$$

6.3.2 $A_q(n, d)$ 的性质

主要讨论关于 $A_q(n, d)$ 的性质.

定义 6.3.4 (码 C 的最小重量:)

设 $C \subset V(n, q)$, 是一个 q 元码, 码 C 的最小重量 (简称为重量) 定义为码 C 中非零码字重量的最小值, 记为 $\omega(C)$, 即 $\omega(C) = \min\{\omega(x) \mid x \in C, x \neq 0\}$.



例题 6.3.5 $C = \{1100, 1000, 0110\}, \omega(C) = 1$.

定义 6.3.5

设 $x = x_1x_2 \cdots x_n \in V(n, 2), y = y_1y_2 \cdots y_n \in V(n, 2), x$ 与 y 的交定义为 $x \cap y = (x_1y_1, x_2y_2, \cdots, x_ny_n)$



例题 6.3.6 $x = 1010, y = 0111, x \cap y = (0, 0, 1, 0) = 0010$.

Hamming 距离和 Hamming 重量的关系:**引理 6.3.2**

关于两个向量之间的 Hamming 距离和 Hamming 重量, 有如下关系:

- (1) 对 $\forall x, y \in V(n, q), d(x, y) = \omega(x - y)$
- (2) 对 $\forall x, y \in V(n, 2), d(x, y) = \omega(x) + \omega(y) - 2\omega(x \cap y)$



证明 (1) 当我们从向量 x 减去向量 y 时, 结果向量 $x - y$ 在任何位置上的元素要么是零 (如果 x 和 y 在该位置上的元素相同), 要么是非零 (如果 x 和 y 在该位置上的元素不同). 因此, $x - y$ 中非零元素的数量正好等于 x 和 y 之间不同元素的数量, 即 $d(x, y)$. 这意味着 $d(x, y)$ 等于 $x - y$ 的 Hamming 重量 $\omega(x - y)$.

(2) 在二进制向量中, Hamming 重量 $\omega(x)$ 和 $\omega(y)$ 分别是 x 和 y 中 1 的数量. 交集 $x \cap y$ 表示 x 和 y 同时为 1 的位置, 其 Hamming 重量 $\omega(x \cap y)$ 表示这些共有的 1 的数量. 因此, $d(x, y)$ 实际上是 x 中独有的 1 加上 y 中独有的 1 的总数. 可以表示为 $\omega(x) + \omega(y)$ 减去两倍的共有 1 的数量 (因为共有的 1 在 $\omega(x)$ 和 $\omega(y)$ 中各被计算了一次), 即 $d(x, y) = \omega(x) + \omega(y) - 2\omega(x \cap y)$

或者 $d(x, y) = \omega(x - y)$, 当 $x_i \neq 0, y_i \neq 0$ 时, $x_iy_i \neq 0$, 故

$$d(x, y) = \omega(x - y) = \omega(x) + \omega(y) - 2\omega(x \cap y).$$

例题 6.3.7 $x = 01100, y = 10110. d(x, y) = \omega(x) + \omega(y) - 2 = 2 + 3 - 2 = 3$

定理 6.3.2

设 d 为奇数, 二元 (n, M, d) 码存在的充分必要条件为存在二元 $(n + 1, M, d + 1)$ 码.



证明 \Rightarrow : 设 C 为一个二元 (n, M, d) 码, 对于 $\forall x = x_1x_2 \cdots x_n \in C$, 定义

$$\hat{x} = \begin{cases} x_1x_2 \cdots x_n0, & \text{如果 } \omega(x) \text{ 是偶数} \\ x_1x_2 \cdots x_n1, & \text{如果 } \omega(x) \text{ 是奇数} \end{cases}$$

于是有 $\hat{x} = x_1x_2 \cdots x_nx_{n+1}$, 其中 $x_{n+1} = \sum_{i=1}^n x_i \pmod{2}$ 或 $\sum_{i=1}^{n+1} x_i \equiv 0 \pmod{2}$. 令 $\hat{C} = \{\hat{x} \mid x \in C\}$, 对于 $\forall \hat{x}, \hat{y} \in \hat{C}$, 由于 $\omega(\hat{x})$ 和 $\omega(\hat{y})$ 都是偶数, 故 $d(\hat{x}, \hat{y}) = \omega(\hat{x}) + \omega(\hat{y}) - 2\omega(\hat{x} \cap \hat{y})$ 也是偶数, 因此 $d(\hat{C})$ 是偶数.

因为 d 是奇数, 又 $d(C) \leq d(\hat{C})$, 故有 $d < d(\hat{C}) \leq d+1$ (由构造知最小距离至多多 1), 于是有 $d(\hat{C}) = d+1$, 于是 \hat{C} 是一个二元 $(n+1, M, d+1)$ 码. 称由 C 到 \hat{C} 的构造方法为对码 C 增加一个奇偶校验位.

\Leftarrow : 设 D 是一个二元 $(n+1, M, d+1)$ 码, 则存在 $x, y \in D$, 使得 $d(x, y) = d+1$, 令 $x = x_1 \cdots x_{i-1} 1 x_{i+1} \cdots x_{n+1}, y = y_1 \cdots y_{i-1} 0 y_{i+1} \cdots y_{n+1}$. D 中所有码字第 i 个位置的元素去掉则得到 M 个长度为 n 的码字, 且 $\hat{x} = x_1 \cdots x_{i-1} x_{i+1} \cdots x_{n+1}, \hat{y} = y_1 \cdots y_{i-1} y_{i+1} \cdots y_{n+1}$, $d(\hat{x}, \hat{y}) = d$, 且对其它得到的新的长度为 n 的码字, 其距离必 $\geq d$. 故得到的新码为 (n, M, d) 码.

推论 6.3.1

如果 d 是奇数, 则 $A_2(n+1, d+1) = A_2(n, d)$. 它等价于如果 d 是偶数, 则 $A_2(n, d) = A_2(n-1, d-1)$.



例题 6.3.8 确定 $A_2(5, 3)$. 设 C 是一个二元 $(5, M, 3)$ 码, 我们来求最大的 M .

假设 $00 \cdots 0 \in C$, 由于 $d(C) = 3$, 故 C 中任意码字 $x, x \neq 0$, 必有 $\omega(x) \geq 3$, 另外, C 中重量为 4 或 5 的码字至多有 1 个, 否则, 若 C 中存在两个重量为 4 或 5 的码字 x 和 y , 则 x 和 y 中为零的分量至多各有 1 个. 因此 $d(x, y) \leq 2$, 这与 $d(C) = 3$ 矛盾.

故 C 中至少含有 $M-2$ 个重量为 3 的码字, 不妨设 $11100 \in C$, 因 $d(C) = 3$, 不难验证另外重量为 3 的码字只能为 $10011, 01011, 00111$ 中的一个. 事实上

$$\begin{aligned} 3 \leq d(x, y) &= \omega(x) + \omega(y) - 2\omega(x \cap y) \\ &= 3 + 3 - 2\omega(x, y) = 6 - 2\omega(x \cap y) \\ &\Rightarrow \omega(x \cap y) = 1 \end{aligned}$$

因 $11100 \in C$, 故对 $y, d(x, y) = 4, \omega(x \cap y) = 1$. y 可能为 $10011, 01011, 00111$, 又任两个的距离均 ≤ 2 , 故由 $d(C) = 3$ 知, y 只可能为上述 3 个码字中的一个. 不妨取 $00111 \in C$, 于是有 $A(3, 5) \leq 1 + 1 + 2 = 4$, 最后对于重量是 4 或 5 的码字进行讨论知 11011 是 C 中的一个码字.

$$\text{因此 } A_2(5, 3) = 4, \quad C = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix} \quad \text{故由推论知 } A_2(6, 4) = 4$$

$$(5, 4, 3) \text{ 码 } \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix} \xrightarrow[\text{奇偶校验位}]{\text{增加}} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \end{pmatrix} (6, 4, 4) \text{ 码}$$

6.3.3 $A_q(n, d)$ 的界

定义 6.3.6

设 $x \in V(n, q)$, r 是非负整数, 以 x 为中心, 以 r 为半径的球定义为

$$S_q(x, r) = \{y \in V(n, q) \mid d(x, y) \leq r\}$$



引理 6.3.3

对于 $\forall x \in V(n, q)$, 球 $S_q(x, r)$ 中所含 $V(n, q)$ 中向量的个数为

$$\binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{r}(q-1)^r.$$



证明 因为与 x 距离为 i 的向量共有 $\binom{n}{i}(q-1)^i$ (共 n 个位置, $n-i$ 个位置对应分量相同, 故固定, 其余 i 个位置 x 中一个位置的分量取定, 则 $y \in V(n, q)$, $d(x, y) \leq i$, y 在这个位置有 $q-1$ 种取法, 与 x 取定的不同, 故对于 x 的 i 个位置 y 有 $(q-1)^i$ 种取法, 故共有 $\binom{n}{i}(q-1)^i$ 个), 从而引理得证. 即 $\sum_{i=0}^r \binom{n}{i}(q-1)^i$.

定理 6.3.3 (Hamming 界)

对任意 q 元 $(n, M, 2t+1)$ 码, 我们有

$$M \left\{ \binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{t}(q-1)^t \right\} \leq q^n.$$



证明 设 C 是一个 q 元 $(n, M, 2t+1)$ 码, 以 C 中的码字为中心, 以 t 为半径的球必定互不相交 (事实上, $\exists x_1, x_2 \in C$ 使得 $y \in S_q(x_1, t) \cap S_q(x_2, t)$, 则 $d(x_1, y) \leq t, d(x_2, y) \leq t$, 从而 $d(x_1, x_2) \leq d(x_1, t) + d(x_2, t) \leq t + t = 2t$ 与码 C 的最小距离为 $2t+1$ 矛盾) 由于每个球中含有 $\sum_{i=0}^t \binom{n}{i}(q-1)^i$ 个 $V(n, q)$ 中的向量且 $|V(n, q)| = q^n$, 故 $M \sum_{i=0}^t \binom{n}{i}(q-1)^i \leq q^n$.

注:

(1) 对于任意二元 $(n, M, 2t+1)$ 码, 我们有

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

(2) Hamming 界给出了 $A_q(n, d)$ 的一个上界, 如二元 $(5, M, 3)$ 码

$$M(1+5) \leq 2^5 = 32 \Rightarrow M \leq 5$$

故 $A_2(5, 3) \leq 5$. 当然, 对于满足 Hamming 界中的不等式的 n, M, d 并不意味着一定存在具有此参数的码. 事实上, $A_2(5, 3) = 4$. 因此, 不存在二元 $(5, 5, 3)$ 码.

定义 6.3.7

设 C 是一个 q 元 $(n, M, 2t + 1)$ 码, 如果

$$M \left\{ \binom{n}{0} + \binom{n}{1} (q-1) + \binom{n}{2} (q-1)^2 + \cdots + \binom{n}{t} (q-1)^t \right\} = q^n$$

则称 C 为完备码.



对于二元重复码 $C = \{00 \cdots 0, 11 \cdots 1\}$, 它是一个二元 $(n, 2, n)$ 码, 当 n 为奇数时, C 是完备码. 另外, 不难验证, 只含一个码字的码以及由 $V(n, q)$ 构成的 q 元 $(n, q^n, 1)$ 码都是完备码. 这三种完备码称为平凡的完备码.

定理 6.3.4 (Gilbert-Varshamov 界)

$$A_q(n, d) \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i}.$$



证明 设 $M = A_q(n, d)$, C 是一个 q 元 (n, M, d) 码, 则

$$V(n, q) \subseteq \bigcup_{x \in C} S_q(x, d-1)$$

用反证法, 若 $\exists y \in V(n, q), y \notin \bigcup_{x \in C} S_q(x, d-1)$, 则对 $\forall x \in C, d(x, y) \geq d$, 则可得到一个 q 元 (n, M', d) 码, $C' = C \cup \{y\}$, $M' = M + 1 > M$. 这与 $M = A_q(n, d)$ 矛盾, 因此

$$M \left\{ \binom{n}{0} + \binom{n}{1} (q-1) + \cdots + \binom{n}{d-1} (q-1)^{d-1} \right\} \geq q^n.$$

定理 6.3.5 (Singleton 界)

$$A_q(n, d) \leq q^{n-d+1}.$$



证明 设 $M = A_q(n, d)$, C 是一个 q 元 (n, M, d) 码. 将码 C 中所有码字都去掉最后 $d-1$ 个分量, 则得到的 $V(n-d+1, q)$ 中的 M 个向量一定互不相同. 否则就会有 $d(C) \leq d-1$, 这与 C 的最小距离为 d 矛盾, 而 $V(n-d+1, q)$ 中有 q^{n-d+1} 个元素, 故 $M \leq q^{n-d+1}$.

6.4 习题课

6.4.1 基本概念

(1) 码: (n, M) 码, 码率 $R = \frac{\log_q M}{n}$.

(2) Hamming 距离

$$d(x, y) = \sum_{j=1}^n d(x_j, y_j) \quad d(x_j, y_j) = \begin{cases} 1 & x_j \neq y_j \\ 0 & x_j = y_j \end{cases}$$

- (3) 最小距离译码 (极大似然译码) .
- (4) 码的最小距离: $d(C) = \min\{d(x, y) \mid x, y \in C \quad x \neq y\}$.
- (5) (n, M, d) 码检错码, 纠错码, 完备码, 系统码.

6.4.2 基本结论

1. 码 C 至多可检查 t 个错误 $\Leftrightarrow d(C) = t + 1$.
2. 码 C 至多可纠正 t 个错误 $\Leftrightarrow d(C) = 2t + 1$ 或 $2t + 2$.
3. Hamming 界若 C 是 q 元 $(n, M, 2t + 1)$ 码, 则


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

4. (G-V) 界: 若存在 q 元 (n, M, d) 码, 则 $A_q(n, d) \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i} (q-i)^i}$

5. (Singleton 界) $A_q(n, d) \leq q^{n-d+1}$.

6. $A_q(n, d)$ 的性质: 设 d 为奇数, 则存在二元 (n, M, d) 码 \Leftrightarrow 存在二元 $(n+1, M, d+1)$ 码, 从而 $A_2(n, d) = A_2(n+1, d+1)$.

6.4.3 课后习题

 **练习 6.4.1** 设 $C = \{11100, 01001, 10010, 00111\}$ 是一个二元 $(5, 4)$ 码.

- (1) 求码 C 的最小距离.
- (2) 根据最小距离译码原则, 对接收到的字 10000, 01100, 00100 分别进行译码.
- (3) 计算码 C 的码率.

解: (1) 设 C 中的四个码字分别为 x_1, x_2, x_3, x_4 , 则

$$\begin{aligned} d(x_1, x_2) &= \begin{pmatrix} \underline{1} & 1 & \underline{1} & 0 & \underline{0} \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} = 3 \\ d(x_1, x_3) &= \begin{pmatrix} 1 & \underline{1} & \underline{1} & \underline{0} & 0 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix} = 3 \\ d(x_1, x_4) &= \begin{pmatrix} \underline{1} & \underline{1} & 1 & \underline{0} & \underline{0} \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} = 4 \\ d(x_2, x_3) &= \begin{pmatrix} \underline{0} & \underline{1} & 0 & \underline{0} & \underline{1} \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix} = 4 \\ d(x_2, x_4) &= \begin{pmatrix} 0 & \underline{1} & \underline{0} & \underline{0} & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} = 3 \\ d(x_3, x_4) &= \begin{pmatrix} \underline{1} & 0 & \underline{0} & 1 & \underline{0} \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} = 3 \end{aligned}$$


于是 $d(C) = 3$

(2) 设 $x = 10000$,

$$\begin{aligned} d(x, x_1) &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix} = 2 \\ d(x, x_2) &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} = 3 \\ d(x, x_3) &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix} = 1 \\ d(x, x_4) &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} = 4 \end{aligned}$$

故将 x 译为 10010, 同理可将 01100 译为 11100, 将 00100 译为 11100 或 00111.

$$(3) \text{ 码率 } R(C) = \frac{\log_q M}{n} = \frac{\log_2 4}{5} = \frac{2}{5}.$$

 **练习 6.4.2** 设 $C = \{00000000, 00001111, 00110011, 00111100\}$ 是一个二元 $(8, 4)$ 码.

(1) 计算码 C 中不同码字的 Hamming 距离和码 C 的最小距离.

(2) 在一个二进制中, 如果把某一个码字中的 0 和 1 互换, 即将 0 换为 1, 1 换为 0, 则我们将所得的字称为原码字的补. 一个二进制的所有码字的补构成的集合称为原码的补码. 求码 C 的补码, 并求补码中所有不同码字之间的 Hamming 距离和补码的最小距离. 它们与 (1) 中的结果有什么关系?

(3) 将 (2) 中的结果推广到一般的二进制.

解: (1) 仍记码 C 中的四个码字为 x_1, x_2, x_3, x_4 , 则有

$$\begin{aligned} d(x_1, x_2) &= d(x_1, x_3) = d(x_1, x_4) = 4 \\ d(x_2, x_3) &= 4 \quad d(x_2, x_4) = 4 \\ d(x_3, x_4) &= 4 \end{aligned}$$

于是 $d(C) = 4$.

(2) 设码 C 为 (n, M, d) 码, 设 C_α 为 C 的补码, 则

$$C_\alpha = \{11111111, 11110000, 11001100, 11000011\}$$

C 的补码中码字间的 Hamming 距离与 C 的相应码字之间 Hamming 距离是相等的, C 的补码和 C 的最小距离相等.

(3) 设码 C 为 (n, M, d) 码, 码 C_α 为 (n, M, d') 码, 则 $x + x_\alpha = (1, 1, \dots, 1), x' + x'_\alpha = (1, 1, \dots, 1), \forall x \in C, x_\alpha \in C_\alpha$,

$$\begin{aligned} d(x_\alpha, x'_\alpha) &= \omega(x_\alpha + x'_\alpha) \\ &= \omega((1, 1, \dots, 1) + x + (1, 1, \dots, 1) + x') \\ &= \omega(x + x') = d(x, x'). \end{aligned}$$


第一步: 计算补码之间的汉明距离. $d(x_\alpha, x'_\alpha) = \omega(x_\alpha + x'_\alpha)$. 这一步说明补码 C_α 中两个码字 x_α 和 x'_α 之间的汉明距离等于它们相加后 (在二进制中相加等同于按位异或) 的汉明重量 (即非零位的数量).

第二步: 考虑补码的定义. x_α 和 x'_α 是原码 x 和 x' 的补码, 即它们是通过将 x 和 x' 中的

每一位取反得到的.

第三步: 利用全 1 向量加法的性质. $\omega((1, 1, \dots, 1) + x + (1, 1, \dots, 1) + x')$. 这一步展示了如何通过向量加法将补码转换回原码的操作. 由于 x_α 和 x'_α 是通过将 x 和 x' 的每一位取反得到的, 所以 $x + (1, 1, \dots, 1)$ 等同于 x 的补码, 即 x_α . 这里, $(1, 1, \dots, 1)$ 表示全 1 向量, 与任何码字相加 (按位异或) 都会得到该码字的补码. 因此, 通过两次加全 1 向量, 我们实际上将 x_α 和 x'_α 转换回了原码 x 和 x' .

第四步: 汉明距离的等价性. $\omega(x + x') = d(x, x')$. 通过上述转换后, 我们实际上计算的是原码 x 和 x' 相加 (按位异或) 的结果的汉明重量, 这正是 x 和 x' 之间的汉明距离.

 **练习 6.4.3** (1) 证明: 对任意三元 $(3, M, 2)$ 码, 一定有 $M \leq 9$.

(2) 证明: 三元 $(3, 9, 2)$ 码一定存在. 于是, $A_2(3, 2) = 3^2$.

(3) 证明: $A_q(3, 2) = q^2$, 其中 $q \geq 2, q$ 是素数的幂次方.

解: (1) 的证明. 由 Singleton 界

$$A_3(3, 2) \leq 3^{n-d+1} = 3^{3-2+1} = 3^2 = 9,$$


故 $M \leq 9$.

(2) 的证明. 可构造出码 $C = \{000, 110, 011, 220, 022, 121, 201, 102, 212\}$ 为 $(3, 9, 2)$ 码, 故 $A_3(3, 2) = 3^2 = 9$.

(3) 的证明. 由 Singleton 界可知

$$A_q(3, 2) \leq q^{n-d+1} = q^{3-2+1} = q^2$$

另一方面, 令 $C = \{(a, b, a+b) \mid a, b \in F_q\}$, 则 $|C| = q^2, d(C) = 2$, 即 C 是一个 q 元 $(3, q^2, 2)$ 码. 因此 $A_q(3, 2) = q^2$.

 **练习 6.4.4** 证明: 如果存在一个二元 (n, M, d) 码, 则一定存在一个二元 $(n-1, M', d)$ 码, 其中 $M' \geq M/2$. 于是, $A_2(n, d) \leq 2A_2(n-1, d)$.

解: 证明: 设 C 是一个二元 (n, M, d) 码, 记

$$C_0 = \{x \in C \mid x = x_1 x_2 \cdots x_{n-1} 0\},$$

$$C_1 = \{x \in C \mid x = x_1 x_2 \cdots x_{n-1} 1\}.$$


则一定有 $|C_0| \geq \frac{M}{2}$ 或 $|C_1| \geq \frac{M}{2}$ (否则 $|C_1| + |C_2| = |C| < M$). 不妨设 $|C_0| \geq \frac{M}{2}$,

$$\text{令 } C' = \{x' = x_1 x_2 \cdots x_{n-1} \mid x = x_1 \cdots x_{n-1} 0 \in C_0\},$$

则 C' 是 $(n-1, M', d')$ 码, 其中 $M' = |C_0|, d' \geq d$, 则必存在 $(n-1, M', d)$ 码, 事实上, 可适当改变 $(n-1, M', d')$ 中的码字, 使其最小距离为 d . 令 $|C| = A_2(n, d)$, 则有 $M' \geq \frac{1}{2} A_2(n, d)$,

$$\text{即 } A_2(n-1, d) \geq M' \geq \frac{1}{2} A_2(n, d)$$

$$\Rightarrow A_2(n, d) \leq 2A_2(n-1, d).$$

 **练习 6.4.5** 设 E_n 是 $V(n, 2)$ 中所有具有偶数重量的向量的集合. 证明: E_n 是一个由 $V(n-1, 2)$ 中的向量增加一个奇偶校验位所得到的码. 于是, E_n 是一个二元 $(n, 2^{n-1}, 2)$ 码.

解: 设 $C_1 = \left\{ x = x_1x_2 \cdots x_{n-1}x_n \in V(n, 2) \mid x_n = \left(\sum_{i=1}^{n-1} x_i \right) (\bmod 2) \right\}$,

$$C_2 = \left\{ x = x_1x_2 \cdots x_{n-1}x_n \in V(n, 2) \mid x_n = \left[\left(\sum_{i=1}^{n-1} x_i \right) + 1 \right] (\bmod 2) \right\}$$

则 $C_1 \cap C_2 = \emptyset$, $|C_1| = 2^{n-1}$, $|C_2| = 2^{n-1}$.

事实上 C_1 中的码字 x 满足 $(x_1 + \cdots + x_{n-1} + x_n) \equiv 0 (\bmod 2)$, 共有 2^{n-1} 个码字. 即 $|C_1| = 2^{n-1}$. 同理 $|C_2| = 2^{n-1}$, 故 $V(n, 2) = C_1 \cup C_2$, $E_n = C_1$. 因此 E_n 是由 $V(n-1, 2)$ 增加 1 个奇偶校验位得到的码, 从而为二元 $(n, 2^{n-1}, 2)$ 码.

练习 6.4.6 证明: 如果存在一个二元 (n, M, d) 码, 并且 d 是偶数, 则一定存在一个二元 (n, M, d) , 其中每个码字都具有偶数重量.

解: 证明: 设 C 是一个二元 (n, M, d) 码, d 是偶数, 则 $d \geq 2$. 存在 $x = x_1x_2 \cdots x_{n-1}x_n, y = y_1y_2 \cdots y_{n-1}y_n$, 使得 $d(x, y) = d$, 不妨设 $x_n \neq y_n$, 将码 C 中每个码字的最后一个分量去掉, 得到

$$C' = \{x' = x_1x_2 \cdots x_{n-1} \in V(n-1, 2) \mid x_1x_2 \cdots x_{n-1}x_n \in C\},$$

则 C' 是一个 $(n-1, M, d-1)$ 码, $d-1$ 为奇数. 令 C'' 是 C' 通过增加奇偶校验位得到的码, 则由定理可知 C'' 是 (n, M, d) 码, 且 C'' 的每个码字的重量为偶数.

练习 6.4.7 证明: 码长为 n 且只含两个码字的不等价的二源码的个数为 n .

解: 证明: C 是只含两个码字的 n 长码, 不妨设 $00 \cdots 0 \in C$, 可设另一个码字 x 的重量为 $\omega(x) = i$ ($1 \leq i \leq n$), 显然 C 等价于 $\{00 \cdots 0 \overbrace{11 \cdots 1}^{i \text{ 个}}, 0 \cdots 000 \cdots 0\}$, 若 $C_1 = \{00 \cdots 0, 00 \cdots 0 \overbrace{11 \cdots 1}^{j \text{ 个}}\}$, 若 $i \neq j$, 则 C 与 C_1 不等价, i 的个数决定了码 C 的类, $1 \leq i \leq n$, 共 n 类.

练习 6.4.8 证明: 任何一个 q 元 (n, q, n) 码都等价于 q 元重复码.

解: 证明: 设 C 是一个 q 元 (n, q, n) 码, 将 C 中码字构成一个 $q \times n$ 矩阵 G , 使得 C 中每个码字是 G 的行向量, 由于 $d(C) = n$, C 中任何两个码字同一位置分量不相同, 因此 G 中任意一列元素均是 $0, 1, \cdots, q-1$ 的排列, 对 G 进行换元置换, 使得置换后 G 的每一列的顺序均为

$$0, 1, \cdots, q-1, \text{ 此时 } C \text{ 等价于 } \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ q-1 & q-1 & \cdots & q-1 \end{pmatrix}.$$

第7章 线性码

线性码是一类很重要的分组码,是讨论各种码的基础. 线性码的编码方法和译码方法都很简单. 许多特殊的线性码都有很好的性质. 绝大多数已知的好码都是线性码. 本章主要介绍有关线性码的基本概念,基本性质以及线性码的编码和译码方法等.

7.1 线性码的定义

7.1.1 线性码的概念及性质

定义 7.1.1

设 $L \subseteq V(n, q)$, 若 L 是 $V(n, q)$ 的子空间, 则称 L 为 q 元线性码. 若 $\dim L = k$, 则称 L 为一个 q 元 $[n, k]$ 线性码, 若 $d(L) = d$, 则称 L 是一个 q 元 $[n, k, d]$ 码.

定理 7.1.1

设 $L \subseteq V(n, q)$ 是线性码, 则 $d(L) = \omega(L)$.

证明 首先设 $L' = \{x - y \mid \forall x, y \in L\}$, 则有 $L = L'$, 因为 L 是子空间, 故对 $\forall x, y \in L$, 有 $x - y \in L$, 即 $L' \subseteq L$.

反之, 对 $\forall x \in L, x = \underbrace{(x + y)}_{\in L} - y$, 故 $x \in L'$, 从而 $L = L'$. 于是

$$d(L) = \min_{x \neq y \in L} \{d(x, y)\} = \min_{x \neq y \in L} \{\omega(x - y)\} = \min_{0 \neq x \in L} \{\omega(x)\} = \omega(L)$$

例题 7.1.1 $n = 3, q = 2, V(3, 2), L = \{100, 010, 110, 000\}$ 是一个 2 元 $[3, 2]$ 线性码, $d(L) = 1$, 即 L 是一个 2 元 $[3, 2, 1]$ 码.

7.1.2 线性码的表示方法

定义 7.1.2

设 L 是一个 q 元 $[n, k]$ 线性码. 由 L 的一组基作为行向量所组成的 $k \times n$ 矩阵 G , 称为线性码 L 的生成矩阵.

注: (1) L 是 q 元 $[n, k]$ 线性码, $\dim L = k$. 设 $\alpha_i = (a_{i1}, \dots, a_{in})$, $(i = 1, \dots, k)$ 为 L 的一组基, 则 L 的生成矩阵 G 为

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{pmatrix}$$

上例中线性码 L 的生成阵 $G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

(2) L 的生成阵 G 若具有形式 $G = (I_k | A)$, 则称 G 为 L 的标准型生成矩阵, 其中 I_k 为 $k \times k$ 阶单位矩阵, A 为 $k \times (n - k)$ 阶矩阵.

(3) 设 q 元 $[n, k]$ 线性码 L 的生成矩阵为 G , 则 $L = \{xG \mid x \in V(k, q)\}$. 事实上, $\forall y \in L, y = \underbrace{x_1\alpha_1 + x_2\alpha_2 + \cdots + x_k\alpha_k}_{\text{行变列}},$

$$x_{1 \times k} G_{k \times n} = \left[\begin{array}{c} (x_1, x_2, \dots, x_k) \end{array} \left(\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{array} \right) \right]_{1 \times n}$$

xG 是 L 中的码字, x 取遍 $V(k, q)$ 时得到 L 中的所有码字.

$$\begin{aligned} & (x_1\alpha_1 + x_2\alpha_2 + \cdots + x_k\alpha_k) \\ &= x_1(a_{11}, a_{12}, \dots, a_{1n}) + x_2(a_{21}, a_{22}, \dots, a_{2n}) + \cdots + x_k(a_{k1}, a_{k2}, \dots, a_{kn}) \\ &= (x_1a_{11} + x_2a_{21} + \cdots + x_ka_{k1}, \dots, x_1a_{1n} + x_2a_{2n} + \cdots + x_ka_{kn}) \end{aligned}$$

故每个码字具有 xG 的形式, 从而 $|L| = q^k$.

$$(x \in V(k, q), \text{ 而 } |V(k, q)| = q^k, x \neq x', xG \neq x'G)$$

若 $x_1\alpha_1 + \cdots + x_k\alpha_k = x'_1\alpha_1 + \cdots + x'_k\alpha_k$ 则 $(x_1 - x'_1)\alpha_1 + \cdots + (x_k - x'_k)\alpha_k = 0 \Rightarrow x_i = x'_i$ 矛盾.

(4) $[n, k]$ 线性码 L 完全由它的生成矩阵来决定.

(5) $[n, k]$ 线性码 L 的码率 $R(L) = \frac{\log_q q^k}{n} = \frac{k}{n}$

例题 7.1.2 设码 L 的生成矩阵为 $G = \left(\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right)$. G 为标准型生成矩阵, $\forall x \in$

$V(4, 2), x = (x_1, x_2, x_3, x_4)$


$$\begin{aligned} xG &= (x_1, x_2, x_3, x_4) \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \\ &= (x_1, x_2, x_3, x_4, x_2 + x_3 + x_4, x_1 + x_3 + x_4, x_1 + x_2 + x_4)_{1 \times 7} \end{aligned}$$

L 是 $[7, 4]$ 码.

7.2 线性码的对偶码

7.2.1 对偶码的定义

定义 7.2.1 (内积)


设 $x = x_1x_2 \cdots x_n, y = y_1y_2 \cdots y_n \in V(n, q)$, 称 $x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n$ 为 x 与 y 的内积, 记为 $x \cdot y$ 或 $\langle x, y \rangle$, 当 $x \cdot y = 0$ 时, 称 x 与 y 是正交的. 

例题 7.2.1 $x = 10111, y = 11101 \in V(5, 2)$, 则 $x \cdot y = 1 \times 1 + 0 \times 1 + 1 \times 1 + 1 \times 0 + 1 \times 1 \equiv 1 \pmod{2}$

定义 7.2.2 (对偶码)

设 L 是一个 q 元 $[n, k]$ 线性码, 称

$$L^\perp = \{x \in V(n, q) \mid \text{对 } \forall c \in L, x \cdot c = 0\}$$

为 L 的对偶码, 即 $\forall x \in L^\perp, x$ 与 L 中的任一个码字都正交. 

7.2.2 对偶码的性质


定理 7.2.1

设 L 是一个 q 元 $[n, k]$ 线性码, 下面性质成立.

(1) 若 G 是线性码 L 的生成矩阵, 则

$$L^\perp = \{x \in V(n, q) \mid xG^T = 0\}.$$

(2) L 的对偶码 L^\perp 是一个 q 元 $[n, n-k]$ 线性码.

(3) $(L^\perp)^\perp = L$. 

证明 (1) 设 G 为线性码 L 的生成阵. 记 $G = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix}$, $\alpha_1, \cdots, \alpha_k$ 为 L 的基.

$\forall x \in L^\perp, x$ 与 L 的每一个码字都正交 $\Leftrightarrow x$ 与 G 的每一行都正交即可即有 $x\alpha_1^T = 0, x\alpha_2^T = 0, \cdots, x\alpha_k^T = 0$ 即 $x(\alpha_1^T, \alpha_2^T, \cdots, \alpha_k^T) = 0 \Rightarrow xG^T = 0$

(2) 由 (1) 可知 $L^\perp = \{x \in V(n, q) \mid xG^T = 0\}$, 而秩 $G = k$, 则 $\dim L^\perp = n - k$, 故 L^\perp 是 q 元 $[n, n-k]$ 码.

(3) 因为对 $\forall x \in L, \forall y \in L^\perp$ 有 $x \cdot y = 0$, 故 $x \in (L^\perp)^\perp$, 从而 $L \subseteq (L^\perp)^\perp$, 又

$$\dim (L^\perp)^\perp = n - \dim (L^\perp) = n - (n - k) = k = \dim(L)$$

所以有 $(L^\perp)^\perp = L$.

例题 7.2.2 对于二元 $[4, 2]$ 线性码 $L = \{0000, 1100, 0011, 1111\}$.

(1) L 是线性码, 易证码中任意两个码字的线性组合 (在二元情况下, 等同于按位异或) 仍然在

码中.(封闭性成立).

(2) L 的生成矩阵为 $G = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$.

(3) $L^\perp = \{xG^T = 0 \mid x \in V(n, q)\}$. 即 $\begin{cases} x_1 + x_2 = 0 \\ x_3 + x_4 = 0 \end{cases}$ 解得 $\xi_1 = (1, 1, 0, 0), \xi_2 = (0, 0, 1, 1)$.

$L^\perp = L, L^\perp$ 也是 $[4, 2]$ 线性码.

(4) 若 $L \subseteq L^\perp$, 则称 L 是自正交的. 若 $L = L^\perp$, 则称 L 是自对偶的.

注: 实数域 R 上的线性空间 $W \cap W^\perp = \{0\}$, 而有限域上的线性空间性质不同, 如上例中 $L = L^\perp$.

例题 7.2.3 设 $L = \{000, 110, 011, 101\}$, L 是一个 $[3, 2]$ 码.

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

生成阵为 $G = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, $L^\perp = \{000, 111\}$, L^\perp 是二元 $[3, 1]$ 线性码.

7.2.3 线性码的校验矩阵

定义 7.2.3

设 L 是一个 q 元 $[n, k]$ 线性码, L^\perp 的生成矩阵 H 称为线性码 L 的校验矩阵.



注:

(1) 线性码 L 的校验阵不唯一, 但 $\text{rank}(H) = n - k$

(2) 设 L 是一个 q 元 $[n, k]$ 线性码, G 为其生成矩阵, H 为其校验矩阵, 则 $x \in L \Leftrightarrow xH^T = 0$

证明 $\Rightarrow: \forall x \in L, x = (\lambda_1, \lambda_2, \dots, \lambda_k)G$,

$$xH^T = (\lambda_1, \lambda_2, \dots, \lambda_k)GH^T = (\lambda_1, \lambda_2, \dots, \lambda_k)(GH^T) = 0$$

故 $\forall x \in L, xH^T = 0$

\Leftarrow : 当 $xH^T = 0$ 时, 则 $x \in (L^\perp)^\perp$, 由对偶码的性质知 $(L^\perp)^\perp = L$, 故有 $x \in L$.

(3) 若 $G = (I_k \mid A)$ 为线性码 L 的标准生成矩阵, 则 L 的校验阵为 $H = (-A^T \mid I_{n-k})$.

证明 $GH^T = (I_k \mid A) \begin{pmatrix} -A \\ I_{n-k} \end{pmatrix} = -A + A = 0$, 因此 H 的每一行与 G 的每一行都正交, 又 $\text{rank}(H) = n - k = \dim(L^\perp)$, 故 H 是 L 的校验矩阵.

(4) 对于二元线性码 L , 如果其标准型生成矩阵为 $G = (I_k \mid A)$, 则其校验矩阵为 $H = (A^T \mid I_{n-k})$.

定理 7.2.2

设 L 是一个 q 元 $[n, k]$ 线性码, 其校验阵为 H , 则 $d(L) = d$ 的充要条件为 H 的任意 $d-1$ 列线性无关, 存在 d 列线性相关.



证明 设 $H = (H_1, H_2, \dots, H_n)$, 其中 H_1, H_2, \dots, H_n 为 H 的 n 个列向量, $x = x_1x_2 \cdots x_n \in V(n, q)$

$$x \in L \Leftrightarrow xH^T = 0 \Leftrightarrow x_1H_1 + \cdots + x_nH_n = 0$$

\Rightarrow : 因为 $d(L) = \omega(L)$, 存在 $x = x_1x_2 \cdots x_n, \omega(x) = d$. 不妨设 x_1, x_2, \dots, x_d 不为 0, 则有

$$xH^T = x_1x_2 \cdots x_dx_{d+1} \cdots x_n \begin{pmatrix} H_1 \\ H_2 \\ \vdots \\ H_n \end{pmatrix} = x_1H_1 + \cdots + x_dH_d = 0$$

从而 H_1, H_2, \dots, H_d 线性相关.

下面证明 H 的任意 $d-1$ 列线性无关, 假设 H 中存在 $d-1$ 列线性相关. 设为 $H_{i1}, H_{i2}, \dots, H_{id-1}$, 则存在不全为 0 的 $x_{i1}, x_{i2}, \dots, x_{id} \in F_q$, 使得 $x_{i1}H_{i1} + x_{i2}H_{i2} + \cdots + x_{id-1}H_{id-1} = 0$.

令 $x = (0, \dots, 0, x_{i1}, 0, \dots, 0, x_{i2}, 0, \dots, 0, x_{id-1}, 0, \dots, 0)$, 则有 $xH^T = x_{i1}H_{i1} + x_{i2}H_{i2} + \cdots + x_{id-1}H_{id-1} = 0$, 即 $x \in L$. 而 $\omega(x) = d-1 < d$ 与 $d(L) = d$ 矛盾. 因此 H 任意 $d-1$ 列线性无关.

\Leftarrow : 设 H 中存在 d 列线性相关, 而任意 $d-1$ 列线性无关. 设 $H_{i1}, H_{i2}, \dots, H_{id}$ 线性相关, 则存在不全为 0 的 $x_{i1}, x_{i2}, \dots, x_{id}$, 使得 $x_{i1}H_{i1} + x_{i2}H_{i2} + \cdots + x_{id}H_{id} = 0$. 因为 H 的任意 $d-1$ 列线性无关, 则 $x_{ij} \neq 0, (j = 1, \dots, d)$ (否则若某一 $x_{ij} = 0$, 不妨设 $x_{i1} = 0$, 则 $x_{i2}H_{i2} + x_{i3}H_{i3} + \cdots + x_{id}H_{id} = 0$, 而 $H_{i2}, H_{i3}, \dots, H_{id}$ 线性无关, 故 $x_{i2} = \cdots = x_{id} = 0$ 矛盾). 令 $x = (0, \dots, 0, x_{i1}, 0, \dots, 0, x_{i2}, 0, \dots, 0, x_{id}, 0, \dots, 0)$, 则有 $xH^T = 0$, 故 $x \in L, \omega(x) = d$. 又 H 的任意 $d-1$ 列线性无关, 故 L 中不存在重量小于 d 的非零码字.

(事实上令 $x = (x_1, \dots, x_{d-1}, 0, \dots, 0)$, 则 $xH^T = x_{i1}H_{i1} + x_{i2}H_{i2} + \cdots + x_{id-1}H_{id-1} = 0 \Rightarrow x_1 = \cdots = x_{d-1} = 0 \Rightarrow x = 0$). 因此 $d(L) = d$.

例题 7.2.4 考虑线性码 Hamming 码 $\text{Ham}(3,2)$ 的生成阵 $G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$ 从而 Ham-

ming 码 $\text{Ham}(3,2)$ 的校验阵为 $H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix}$. 容易看出 H 的任意两列线

性无关. $\alpha_3 = \alpha_1 + \alpha_2$, 故前 3 列 $\alpha_1, \alpha_2, \alpha_3$ 线性相关. 从而 $d(C) = 3, C$ 是一个二元 $[7, 4, 3]$ 码.

7.2.4 线性码的界

定理 7.2.3 ($G-V$ 界)

设 q 是一个素数的方幂, 如果 n, k, d 满足

$$\sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i < q^{n-k} \quad (7.2.1)$$

则一定存在一个最小距离是 d 的 q 元 $[n, k]$ 线性码. 因此

$$A_q(n, d) \geq q^k$$

其中 k 为使不等式 (7.2.1) 成立的最大整数.



证明 假设参数 n, k, d 满足不等式 (7.2.1), 由最小距离 d 满足的条件, 若我们能构造出一个 $(n-k) \times n$ 阶校验矩阵 H , 使得 H 的任意 $d-1$ 列线性无关, 则定理成立.

首先在 $V(n-k, q)$ 中任取一个非零向量 h_1 作为 H 的第 1 列, 然后在 $V(n-k, q)$ 中任取 $h_2, h_2 \neq \lambda h_1 (\lambda \in F_q)$ 作为 H 的第 2 列. 一般地, 取 H 的第 i 列 h_i 为 $V(n-k, q)$ 中的一个非零向量, 并且 h_i 不是前面已选取的 $i-1$ 个向量 h_1, h_2, \dots, h_{i-1} 中的任意不多于 $d-2$ 个向量的线性组合 (保证任意 $d-1$ 列线性无关). 我们用 N_i 表示 h_1, h_2, \dots, h_{i-1} 中的任意不多于 $d-2$ 个向量的不同的线性组合的个数, 则

$$N_i \leq 1 + \binom{i-1}{1} (q-1) + \binom{i-1}{2} (q-1)^2 + \dots + \binom{i-1}{d-2} (q-1)^{d-2}$$

于是, h_i 有 $q^{n-k} - N_i$ 种取法, 因此只要不等式 (7.2.1) 成立, 一定可以构造出校验矩阵 H , 从而得到一个最小距离至少是 d 的 q 元 $[n, k]$ 码.

定理 7.2.4

设 L 是一个 q 元 $[n, k, d]$ 线性码, 则 $d \leq n - k + 1$.



证明 L 是一个 q 元 $[n, k, d]$ 线性码, 其对偶码 L^\perp 是一个 q 元 $[n, n-k]$ 线性码, 线性码 L 的校验阵 H 是一个 $(n-k) \times n$ 阶矩阵, 并且它的 $n-k$ 行线性无关. H 的任意 $n-k+1$ 列一定线性相关 (行秩与列秩相等), 故 $d \leq n - k + 1$

注: 对于一个 q 元 $[n, k, d]$ 线性码 L , 如果 $d = n - k + 1$, 则称 L 为最大距离可分码, 简称为 MDS 码.

7.3 线性码的译码方法

定义 7.3.1 (伴随式)

设 L 是一个 q 元 $[n, k]$ 线性码, H 为它的校验阵, 对 $\forall x \in V(n, q)$, 称 xH^T 为 x 的伴随式, 记为 $S(x)$.



注:

$$(1) S(x) = 0 \Leftrightarrow x \in L$$

$$(2) \text{ 设 } H = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_{n-k} \end{pmatrix}_{(n-k) \times n}, \text{ 则}$$

$$\begin{aligned} S(x) &= xH^T = x_{1 \times n} (h_1^T, h_2^T, \dots, h_{n-k}^T)_{n \times (n-k)} \\ &= (xh_1^T, xh_2^T, \dots, xh_{n-k}^T) \\ &= (x \cdot h_1, x \cdot h_2, \dots, x \cdot h_{n-k}) \end{aligned}$$

考虑商空间 $\frac{V(n,q)}{L} = \{x + L \mid x \in V(n,q)\}$. 集合 $x + L = \{x + c \mid \forall c \in L\}$ 称为 L 的陪集.

对 $\forall a \in F_q, x + L, y + L \in \frac{V(n,q)}{L}$. 定义 $a(x + L) = ax + L, (x + L) + (y + L) = (x + y) + L$, 则 $x + L = y + L \Leftrightarrow x - y \in L$.

定理 7.3.1

设 L 是一个 q 元 $[n, k]$ 线性码, H 是它的校验阵, 则对于 $x, y \in V(n, q), x + L = y + L \Leftrightarrow xH^T = yH^T$.



证明

$$\begin{aligned} x + L = y + L &\Leftrightarrow x - y \in L \\ &\Leftrightarrow (x - y)H^T = 0 \\ &\Leftrightarrow xH^T = yH^T. \end{aligned}$$

定理 7.3.2

设 L 是一个 $[n, k]$ 线性码, H 是它的校验阵, 最小距离译码等价于将收到的字 x 译成码字 $c = x - a$, 其中 a 是陪集 $x + L$ 中具有最小重量, 且与 x 具有相同伴随式.



注: (1) 接收端接收到字为 x , x 译为 c , 即 $a = x - c$ 为最小. 当 c 取遍 L 中元素时, a 取遍陪集 $x + L$. 于是根据最小距离译码, 将 x 译为 $c = x - a$, 其中 a 是 $x + L$ 中具有最小重量的字.

(2) 标准阵译码

$$\begin{array}{cccccc} 0 & c_1 & c_2 & \cdots & c_{m-1} & \\ a_1 & a_1 + c_1 & a_1 + c_2 & \cdots & a_1 + c_{m-1} & a_1 + L \\ a_2 & a_2 + c_1 & a_2 + c_2 & \cdots & a_2 + c_{m-1} & a_2 + L \\ \vdots & \vdots & \vdots & & \vdots & \\ a_{s-1} & a_{s-1} + c_1 & a_{s-1} + c_2 & \cdots & a_{s-1} + c_{m-1} & a_{s-1} + L \end{array}$$

其中 $m = q^k, s = q^{n-k}, a_i \notin L, a_i$ 是 $a_i + L$ 中的重量最小的字, $i = 1, \dots, s-1$, 选取不在前 i 行出现且重量最小的字 a_i 与第一行的每个码字相加得到第 $i+1$ 行, 得到 $a_i + L$. 此过程一直进行到表中包含 $V(n, q)$ 中的所有字, 形成标准阵. 若 x 出现在第 $i+1$ 行第 $j+1$ 列, $i \geq 0, j \geq 0$,

则 $x = c_j + a_i$, 将 x 译为 $c_j = x - a_i$, 即译为包含 x 的那一列中最上面的码字.

例题 7.3.1 设 L 是一个二元 $[4, 2]$ 线性码, 其生成矩阵为

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

则 $L = \{0000, 1011, 0101, 1110\}$ (事实上, $|L| = q^k = 2^2 = 4$. $L = xG = (x_1, x_2) \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$,

而 $(x_1, x_2) = (0, 0), (1, 0), (0, 1), (1, 1)$, 计算可得 L 的 4 个码字) .

标准阵

0000	1011	0101	1110	
1000	0011	1101	0110	$a_1 + L$
0100	1111	0001	1010	$a_2 + L$
0010	1001	0111	1100	$a_3 + L$

($|V(4, 2)| = 2^4 = 16$, 故上述为标准阵). 收到字 $x = 1100$, 则由上表可知将 1100 译为 1110 .

(3) 伴随式译码. 由于标准阵中每一行元素的伴随式相同, 故只需计算陪集头和对应的伴随式. 如果接收到 x , 计算伴随式 xH^T , 确定 xH^T 被译为 $c = x - a_i$, ($x = c + a_i$) 称为伴随式译码.

例题 7.3.2 如上例 $G = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$, 标准型生成矩阵 $G' = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$, 则校验阵

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

陪集头	伴随式 (xH^T)	$x_{1 \times 4} H_{4 \times 2}^T$
0 0 0 0	0 0	
1 0 0 0	1 1	
0 1 0 0	0 1	
0 0 1 0	1 0	

($V(2, 2)$ 中的所有元已算出, 故不再计算 $(0001)H^T$)

设在信道中接收到的字为 $x = 0001$, 计算伴随式

$$xH^T = (0001) \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

故将 x 译为 $c = x - a_2 = 0001 - 0100 = 0101$.

7.4 线性码的重量分布

定义 7.4.1 (线性码 L 的重量分布多项式)

设 L 是一个 q 元 $[n, k]$ 线性码, A_i 表示 L 中重量等于 i 的码字的个数, $0 \leq i \leq n$. 称 A_0, A_1, \dots, A_n 为 L 的重量分布, 称多项式

$$W_L(z) = \sum_{i=0}^n A_i z^i = A_0 + A_1 z + \dots + A_n z^n$$

为 L 重量分布多项式. 显然, $W_L(z) = \sum_{x \in L} z^{w(x)}$



例题 7.4.1 设 L 是一个二元 $[3, 2]$ 线性码, $L = \{000, 011, 101, 110\}$, 其对偶码为 $L^\perp = \{000, 111\}$. L 和 L^\perp 的重量分布多项式分别为

$$W_L(z) = 1 + 3z^2$$

$$W_{L^\perp}(z) = 1 + z^3.$$

线性码的 Mac Williams 恒等式:

一般而言, 确定码的重量分布是一件困难的事情. 对于线性码, MacWilliams 恒等式给出了线性码 L 的重量分布多项式与其对偶码 L^\perp 的重量分布多项式之间的一种关系. 我们下面主要介绍二元线性码的 MacWilliams 恒等式, 并给出它的证明. 对于 q 元情形, 只给出结论.

引理 7.4.1

设 L 是一个二元 $[n, k]$ 线性码, $y \in V(n, 2)$, 并且 $y \notin L^\perp$, 则 L 中使 $x \cdot y$ 等于 0 和 1 的码字 x 的个数相等.



证明 设

$$A = \{x \in L \mid x \cdot y = 0\},$$

$$B = \{x \in L \mid x \cdot y = 1\}.$$

因为 $y \notin L^\perp$, 所以存在 $u \in L$, 使得 $u \cdot y = 1$. 记

$$u + A = \{u + x \mid x \in A\},$$

$$u + B = \{u + x \mid x \in B\}.$$

我们有

$$u + A \subseteq B,$$

$$u + B \subseteq A.$$

于是 $|A| = |B|$.

引理 7.4.2

设 L 是二元 $[n, k]$ 线性码, $y \in V(n, 2)$, 则

$$\sum_{x \in L} (-1)^{x \cdot y} = \begin{cases} 2^k, & \text{如果 } y \in L^\perp; \\ 0, & \text{如果 } y \notin L^\perp. \end{cases}$$



证明 如果 $y \in L^\perp$, 则对于任意 $x \in L, x \cdot y = 0$. 因此,

$$\sum_{x \in L} (-1)^{x \cdot y} = |L| = 2^k.$$

如果 $y \notin L^\perp$, 由引理 7.4.1 知, 当 x 取遍 L 中的所有码字时, 有 2^{k-1} 个码字使得 $x \cdot y = 0$, 同样有 2^{k-1} 个码字使得 $x \cdot y = 1$. 因此, 我们有

$$\sum_{x \in L} (-1)^{x \cdot y} = 0$$

引理 7.4.3

设 $x \in V(n, 2)$, 则

$$\sum_{y \in V(n, 2)} z^{\omega(y)} (-1)^{x \cdot y} = (1 - z)^{\omega(x)} (1 + z)^{n - \omega(x)}$$



证明

因为

$$\sum_{j=0}^1 z^j (-1)^{x_i j} = \begin{cases} 1 + z, & \text{如果 } x_i = 0; \\ 1 - z, & \text{如果 } x_i = 1. \end{cases}$$

所以

$$\begin{aligned} \sum_{y \in V(n, 2)} z^{\omega(y)} (-1)^{x \cdot y} &= \sum_{y_1=0}^1 \sum_{y_2=0}^1 \cdots \sum_{y_n=0}^1 z^{\sum_{i=1}^n y_i} (-1)^{\sum_{i=1}^n x_i y_i} \\ &= \sum_{y_1=0}^1 \sum_{y_2=0}^1 \cdots \sum_{y_n=0}^1 \left(\prod_{i=1}^n z^{y_i} (-1)^{x_i y_i} \right) \\ &= \prod_{i=1}^n \left(\sum_{j=0}^1 z^j (-1)^{x_i j} \right) \\ &= (1 - z)^{\omega(x)} (1 + z)^{n - \omega(x)}. \end{aligned}$$

定理 7.4.1 (二元线性码的 Mac Williams 恒等式)

设 L 是一个二元 $[n, k]$ 线性码, L^\perp 为其对偶码, 则有

$$W_{L^\perp}(z) = \frac{1}{2^k} (1 + z)^n W_L \left(\frac{1 - z}{1 + z} \right).$$

由于 L^\perp 是一个二元 $[n, n - k]$ 线性码, 并且 $(L^\perp)^\perp = L$, 所以 Mac Williams 恒等式可以写成

$$W_L(z) = \frac{1}{2^{n-k}} (1 + z)^n W_{L^\perp} \left(\frac{1 - z}{1 + z} \right).$$



证明 设

$$f(z) = \sum_{x \in L} \left(\sum_{y \in V(n, 2)} z^{\omega(y)} (-1)^{x \cdot y} \right).$$

一方面, 根据引理 7.4.3,

$$\begin{aligned} f(z) &= \sum_{x \in L} (1-z)^{\omega(x)} (1+z)^{n-\omega(x)} \\ &= (1+z)^n \sum_{x \in L} \left(\frac{1-z}{1+z} \right)^{\omega(x)} \\ &= (1+z)^n W_L \left(\frac{1-z}{1+z} \right). \end{aligned}$$

另一方面, 根据引理 7.4.2,

$$f(z) = \sum_{y \in V(n,2)} z^{\omega(y)} \sum_{x \in L} (-1)^{x \cdot y} = \sum_{y \in L^\perp} z^{\omega(y)} 2^k = 2^k W_{L^\perp}(z).$$

因此, 我们有

$$W_{L^\perp}(z) = \frac{1}{2^k} (1+z)^n W_L \left(\frac{1-z}{1+z} \right)$$

定理 7.4.2 (q 元线性码的 Mac Williams 恒等式)

设 L 是一个 q 元 $[n, k]$ 线性码, L^\perp 为其对偶码, 则有

$$W_{L^\perp}(z) = \frac{1}{q^k} (1 + (q-1)z)^n W_L \left(\frac{1-z}{1+(q-1)z} \right).$$



该定理不证明.

例题 7.4.2 在第一个例子中, $W_L(z) = 1 + 3z^2$, $[3, 2]$ 码, 则

$$\begin{aligned} W_{L^\perp}(z) &= \frac{1}{2^2} (1+z)^3 W_L \left(\frac{1-z}{1+z} \right) \\ &= \frac{1}{4} (1+z)^3 \times \left[1 + 3 \left(\frac{1-z}{1+z} \right)^2 \right] \\ &= \frac{1}{4} [(1+z)^3 + 3(1-z)^2(1+z)] \\ &= 1 + z^3. \end{aligned}$$

同样

$$\begin{aligned} W_L(z) &= \frac{1}{2^k} (1+z)^n W_{L^\perp} \left(\frac{1-z}{1+z} \right) \quad L^\perp[3, 1] \text{ 码} \\ &= \frac{1}{2} (1+z)^3 W_{L^\perp} \left(\frac{1-z}{1+z} \right) \\ &= \frac{1}{2} (1+z)^3 \left[1 + \left(\frac{1-z}{1+z} \right)^3 \right] \\ &= \frac{1}{2} [(1+z)^3 + (1-z)^3] \\ &= 1 + 3z^2. \end{aligned}$$

7.5 习题课

7.5.1 本章小结


1. 基本概念

线性码, 生成矩阵, 校验阵, 标准阵 (列) 译码, 伴随式, 对偶码, 重量分布多项式, MDS 码.


2. 基本结论

- (1) 校验阵与最小距离的关系
- (2) 线性码的 $G-V$ 界. $d \leq n - k + 1$
- (3) 标准阵译码方法
- (4) Mac Williams 恒等式 (二元情形)

7.5.2 课后习题

 **练习 7.5.1** 一个二元 $(11, 24, 5)$ 码会是线性码吗? 为什么?

解: 二元 $(11, 24, 5)$ 码不会是线性码. 二元线性码的码字个数应该是 2^k , 而 24 不是 2 的幂.


 **练习 7.5.2** 设 E_n 是 $V(n, 2)$ 中所有具有偶数重量的向量的集合. 证明: E_n 是线性码, 确定 E_n 的参数 $[n, k, d]$ 以及其标准型的生成矩阵.

解: 证明: $\forall x, y \in E_n, d(x, y) = \omega(x - y), \omega(x - y) = \omega(x) + \omega(y) - 2\omega(x \cap y)$. 故 $\omega(x + y)$ 为偶数, $x + y = x - y, x + y \in E_n, E_n$ 是线性码. 由上一章课后题知 E_n 是一个 $[n, n - 1, 2]$ 线性码, 标准生成阵为

$$G = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{pmatrix}$$

设 $L = \{xG \mid \forall x \in V(n - 1, q)\}$, 对 n 用归纳法证明, 则 L 中的元均具有偶重量, 从而 $L \subseteq E_n$. 又 $\dim L = \dim E_n = n - 1$, 故 $L = E_n$.

($\forall x \in L, x = v_{j1} + v_{j2} + \cdots + v_{jk}$, 系数取在 F_2 中 $\{0, 1\}, k = 2$ 时, $x = v_{j1} + v_{j2}, \omega(x) = \omega(v_{j1}) + \omega(v_{j2}) - 2\omega(v_{j1} \cap v_{j2})$ 为偶数易用归纳法证得)

 **练习 7.5.3** 证明: 对于任意一个二元线性码 L , 一定满足下列条件之一.

- (1) L 中所有码字都具有偶数重量;
- (2) L 中一半码字具有偶数重量, 另一半码字具有奇数重量.

解: 证明: 设 L 是一个二元 $[n, k]$ 线性码, 且 L 中存在码字 x_0 , 使得 $\omega(x_0)$ 为奇数. 令 $L_1 = \{x \in L \mid \omega(x) \text{ 是偶数}\}, L_2 = \{x \in L \mid \omega(x) \text{ 是奇数}\}, x_0 + L_1 = \{x_0 + x \mid \forall x \in L_1\}, x_0 + L_1$ 中元素重量为奇数, $x_0 + L_1 \subseteq L_2$ 同理 $x_0 + L_2 \subseteq L_1$.

$$|L_2| \geq |x_0 + L_1| = |L_1| \geq |x_0 + L_2| = |L_2| \Rightarrow |L_1| = |L_2|.$$

(否定一个证另一个)

(方法二) 令

$$\begin{aligned} E &= \{x \in L \mid W(x) \text{ 为偶数}\}, \\ O &= \{y \in L \mid W(y) \text{ 为奇数}\}. \end{aligned}$$

因为零向量属于 E , 所以 $E \neq \emptyset$. 假设 $E \neq L$, 则 $O \neq \emptyset$. 于是存在 $y \in O$. 令

$$y + E = \{y + x \mid x \in E\}.$$


显然, $y + E \subseteq C$, 并且 $|y + E| = |E|$. 因为

$$W(y + x) = W(y) + W(x) - 2W(y \cap x),$$

所以 $y + E$ 中的每个码字都具有奇数重量, 即 $y + E \subseteq O$. 因此, $|E| = |y + E| \leq |O|$. 今

$$y + O = \{y + u \mid u \in O\}.$$

显然, $y + O \subseteq E$. 于是, $|O| = |y + O| \leq |E|$. 由 $|E| \leq |O|$ 和 $|O| \leq |E|$ 立即可得 $|E| = |O| = \frac{1}{2}|L|$.

 **练习 7.5.4** 设 L 是一个二元 $[n, k]$ 线性码, 并且对于所有码字, 任一位置的分量不全为零, L_1 是 L 在某一固定位置上取 0 的码字构成的集合. 证明: L_1 是一个二元 $[n, k-1]$ 线性码.

解: 设 L 是一个二元 $[n, k]$ 线性码.

$$\begin{aligned} L_1 &= \{x_1 x_2 \cdots x_i \cdots x_n \in L \mid x_i = 0\}, \\ L_2 &= \{x_1 x_2 \cdots x_i \cdots x_n \in L \mid x_i = 1\}. \end{aligned}$$

$L_1 \cup L_2 = L$, $L_1 \cap L_2 = \emptyset$, 由条件 $L_2 \neq \emptyset$. 设 $x_0 \in L_2$, 当 $x \in L_1$ 时, $x_0 + x \in L$. $x_0 + x$ 的第 i 个分量为 1. $x_0 + x \in L_2$, $x_0 + L_1 \subseteq L_2$. 同理 $x_0 + L_2 \subseteq L_1$. 故 $|L_1| = |L_2| = 2^{k-1}$. 故 L_1 是一个 $[n, k-1]$ 线性码.

(方法二) 设 C_0 是 C 中第 i 个分量为 0 的码字所构成的集合, C_1 是 C 中第 i 个分量为 1 的码字所构成的集合. 由前提假设可知, $C_0 \neq \emptyset$, $C_1 \neq \emptyset$. 因为对任意 $x, y \in C_0$, $x + y$ 的第 i 个分量还是 0, 所以 $x + y \in C_0$. 因此, C_0 是 $V(n, 2)$ 的一个子空间, 即 C_0 是一个二元线性码. 设 $u \in C_1$, 令

$$\begin{aligned} u + C_0 &= \{u + x \mid x \in C_0\}, \\ u + C_1 &= \{u + v \mid v \in C_1\}. \end{aligned}$$


不难看出, $u + C_0 \subseteq C_1$, $u + C_1 \subseteq C_0$. 于是,

$$\begin{aligned} |C_0| &= |u + C_0| \leq |C_1|, \\ |C_1| &= |u + C_1| \leq |C_0|. \end{aligned}$$

由 $|C_0| \leq |C_1|$ 和 $|C_1| \leq |C_0|$ 立即得到

$$|C_0| = |C_1| = \frac{1}{2}|C| = 2^{k-1}.$$

因此, C_0 是一个二元 $[n, k-1]$ 线性码.

 **练习 7.5.5** 设 L 是一个二元 $[n, k]$ 线性码, 并且其生成矩阵的列向量中不包含零向量. 证明: L 中所有码字的重量之和为 $n2^{k-1}$.

解: 证明: 设 $C_{i,1}$ 是 L 中第 i 个分量为 1 的码字所构成的集合, $1 \leq i \leq n$. 由上题的解题过程

可知

$$|C_{i,1}| = \frac{1}{2}|L| = 2^{k-1}.$$

将 L 中所有码字排成一个 $2^k \times n$ 阶矩阵, 其中每一行是一个码字, 则每一列中 1 的个数为 2^{k-1} . 因此, L 中所有码字的重量之和为 $n2^{k-1}$.

 **练习 7.5.6** 设二元线性码 L 的生成矩阵为

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

试求 L 的最小距离.


解: 对矩阵 G 进行行初等变换

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} = G'$$

G' 是 L 的标准型生成矩阵, 那么 L 的标准型校验矩阵是

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

H 的任意一列线性无关, 存在两列线性相关. 因此, L 的最小距离是 2.

 **练习 7.5.7** 设三元线性码 L 的生成矩阵为

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$$

试求 L 的最小距离, 并证明 L 是完备码.

解: $G = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{array} \right) = (I_2|A)$, 故 L 的校验阵为 $(-A^T|I_2) = \begin{pmatrix} 2 & 2 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix} = H$. (三元码中加法和乘法遵循模 3 的规则) H 中任意两列线性无关, 第 1, 2, 4 列线性相关 ($2H_1 + H_2 = 2H_4$), 故 $d(L) = 3$. 由于

$$3^2 \left(\binom{4}{0} + \binom{4}{1} (3-1) \right) = 3^4$$

因此, L 是完备码.

(方法二) 通过生成矩阵 G 生成的三元线性码 L 包含以下码字:

$$L = \{(0, 0, 0, 0), (0, 1, 1, 2), (0, 2, 2, 1), (1, 0, 1, 1), (1, 1, 2, 0), (1, 2, 0, 2), (2, 0, 2, 2), (2, 1, 0, 1), (2, 2, 1, 0)\}$$

这些码字是通过 G 的行的所有可能的线性组合生成的, 其中的加法和乘法都是在 $GF(3)$ (即模 3 的运算) 下进行的. 这个线性码 L 完全展示了由 G 定义的线性空间的结构. $d(L) = w(L) = 3$. L 是一个三元 $[4, 2, 3]$ 线性码.

 **练习 7.5.8** 设二元线性码 L 的生成矩阵为

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

试求 L 的标准阵, 并对信道接收端接收到的字 11111 和 10000 分别进行译码.

解: 我们首先要从给定的生成矩阵 G 生成所有可能的码字. 然后, 我们将使用这些码字来确定接收到的字的最近邻码字, 即通过汉明距离最小的码字来译码. 对于二元情况, 我们只有 0 和 1 作为系数:

- 使用 0 倍的每行 (不使用基向量) 生成码字 00000.
- 使用 1 倍的第一行生成码字 11010.
- 使用 1 倍的第二行生成码字 01010.
- 使用 1 倍的第二行生成码字 01010.


因此, 由 G 生成的码字集合为 $L = \{00000, 11010, 01010, 10000\}$. 为了译码接收到的字, 我们需要计算它们与 L 中所有码字的汉明距离, 并找到距离最小的码字.

译码 11111 的最小汉明距离为 2, 因此接收到的字 11111 最接近的码字是 11010. 译码 10000 的最小汉明距离为 0, 意味着接收到的字 10000 已经是码 L 中的一个码字. 这表明 10000 是正确传输的, 没有发生错误.

L 的标准阵为

00000	<u>10000</u>	01010	11010
01000	11000	00010	10010
00100	10100	01110	11110
00001	10001	01011	11011
01100	11100	00110	10110
01001	11001	00011	10011
00101	10101	01111	<u>11111</u>
01101	11101	00111	10111

因此把 11111 译为 11010, 10000 译为 10000.

 **练习 7.5.9** 设 L 是一个二元线性码, x 是在信道接收端收到的字, H 为 L 的校验矩阵. 证明: Hx^T 恰好是 x 发生错误位置所对应的 H 的列向量的和.

解: 证明: 设 c 是发送的码字, $e = x - c, e = (e_1, e_2, \dots, e_n), e$ 中不为零的位置恰好是发生错误的位置.

$$H(x - c)^T = Hx^T - Hc^T = Hx^T,$$

$$\text{即 } Hx^T = He^T = \sum_{i=1}^n e_i H_i.$$

 **练习 7.5.10** 设三元线性码 L 的生成矩阵为

$$G = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 \end{pmatrix}.$$

(1) 试求 L 的标准型的生成矩阵.

(2) 试求 L 的标准型的校验矩阵.

(3) 试利用伴随式译码方法对信道接收端接收到的字 2121、1201、2222 分别进行译码.

解:

$$(1) G = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 2 & 1 \end{pmatrix} = \left(\begin{array}{cc|cc} 1 & 0 & 2 & 2 \\ 0 & 1 & 2 & 1 \end{array} \right) = G'$$

 G' 为 L 的标准型的生成矩阵.(2) L 的标准型的校验矩阵为

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}$$

(3) $C = \{0000, 2011, 1022, 1110, 0121, 2102, 2220, 1201, 0212\}$. 要对接收到的字 2121、1201、2222 进行译码, 我们需要计算它们与 H 的乘积 Hx^T , 以找到对应的综合校验向量.

$$x_1 H^T = 22, \quad a_1 = 2000, \quad a_1 H^T = 22, \quad x_1 - a_1 = 0121$$


$$x_2 H^T = 00, \quad a_2 = 0000, \quad a_2 H^T = 00, \quad x_2 - a_2 = 1201.$$

$$x_3 H^T = 02, \quad a_3 = 0002, \quad a_3 H^T = 02, \quad x_3 - a_3 = 2220.$$

因为 $d(C) = W(C) = 3$, 码 C 至多可以纠正一个错误, 所以 $V(4, 3)$ 中的 9 个重量不大于 1 的向量都是陪集代表元. 由于共有 $\frac{3^4}{3^2} = 9$ 个陪集, 所以 $V(4, 3)$ 中的 9 个重量不大于 1 的向量恰好是所有陪集的代表元. 码 C 的伴随式列表为

陪集代表元	伴随
0000	00
1000	11
2000	22
0100	12
0200	21
0010	10
0020	20
0001	01
0002	02

$S(2121) = 22$, 2121 译码为 $2121 - 2000 = 0121$. $S(1201) = 00$, 1201 译码为 $1201 - 0000 = 1201$. $S(2222) = 02$, 2222 译码为 $2222 - 0002 = 2220$.

 **练习 7.5.11** 设二元线性码 L 的校验矩阵为 H , \hat{L} 是由 L 增加一个奇偶校验位得到的扩充码. 证明: \hat{L} 是线性码, 并且其校验矩阵为

$$\hat{H} = \left(\begin{array}{cccc|c} & & & & 0 \\ & & & & 0 \\ & & & & \vdots \\ & & & & 0 \\ \hline 1 & 1 & \cdots & 1 & 1 \end{array} \right).$$

解: 证明: 对任意 $x = x_1 x_2 \cdots x_n x_{n+1} \in \hat{L}$, $y = y_1 y_2 \cdots y_n y_{n+1} \in \hat{L}$, 设 $z = x + y =$

$z_1 z_2 \cdots z_n z_{n+1}$. 因为 $x_1 x_2 \cdots x_n + y_1 y_2 \cdots y_n \in L$, 所以 $z_1 z_2 \cdots z_n \in L$, 而 $\sum_{i=1}^{n+1} x_i \equiv \sum_{i=1}^{n+1} y_i \equiv 0(\text{mod } 2)$, 因此有

$$\sum_{i=1}^{n+1} z_i = \sum_{i=1}^{n+1} (x_i + y_i) \equiv 0(\text{mod } 2)$$

从而 $z \in \hat{L}$, 即 \hat{L} 是线性码. 对于任意 $\hat{x} = x x_{n+1} \in \hat{L}$, 其中 $x = x_1 x_2 \cdots x_n \in L$

$$\overline{H}\hat{x}^T = \begin{pmatrix} Hx^T \\ \sum_{i=1}^{n+1} x_i \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

若 L 是 $[n, k]$ 线性码, 则 \hat{L} 是 $[n+1, k]$ 线性码, $\text{rank}(H) = n - k$, 从而 $\text{rank}(\overline{H}) = n - k + 1 = (n+1) - k$, 因此 \overline{H} 是 \hat{L} 的校验矩阵.

 **练习 7.5.12** 设二元线性码 L 的生成矩阵为

$$G = \left(I_7 \left| \begin{array}{cccc} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{array} \right. \right).$$

试求 L 的校验矩阵, 并求其最小距离.

解: L 的校验矩阵为

$$H = \left(\begin{array}{cccccc|cccc} 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array} \right)$$

H 的任意两列线性无关, 前三列线性相关 (第一列是第二、三列的加和). 因此 L 的最小距离为 3.

 **练习 7.5.13** 设 5 元线性码 L 的生成矩阵为

$$G = \begin{pmatrix} 1 & 2 & 4 & 0 & 3 \\ 0 & 2 & 1 & 4 & 1 \\ 2 & 0 & 3 & 1 & 4 \end{pmatrix}$$

- (1) 试求 L 的标准型的生成矩阵.
- (2) 试求 L 的标准型的校验矩阵.
- (3) 试求 L 的最小距离.

解: (1)


$$G = \begin{pmatrix} 1 & 2 & 4 & 0 & 3 \\ 0 & 2 & 1 & 4 & 1 \\ 2 & 0 & 3 & 1 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 4 & 0 & 3 \\ 0 & 2 & 1 & 4 & 1 \\ 0 & 1 & 0 & 1 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 4 & 0 & 3 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 2 & 1 & 4 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 4 & 0 & 3 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 2 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & 2 & 3 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 2 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 2 & 0 \end{pmatrix} = G'$$

G' 为 L 的标准型的生成矩阵.

(2) L 的标准型的校验矩阵为

$$H = \begin{pmatrix} 0 & 4 & 3 & 1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{pmatrix}$$

(3) L 的校验矩阵 H 中存在两列线性相关 (第一列是第五列的三倍, 注意是 5 元), 任意一列线性无关. 因此 L 的最小距离为 2.


 **练习 7.5.14** 设 H 是一个 q 元 $[n, k]$ 线性码 L 的校验矩阵, K 是 F_q 上的一个 $n - k$ 阶可逆方阵. 证明: KH 仍然是 L 的校验矩阵.

解: 我们知道, $L = \{x \in V(n, q) \mid xH^T = 0\}$. 显然, 对任意 $x \in L$,

$$x(KH)^T = xH^T K^T = 0K^T = 0.$$

反过来, 对任意 $y \in V(n, q)$, 如果 $y(KH)^T = 0$, 则由于 K 可逆, 所以 $xH^T = 0$. 于是, $y \in L$. 因此, $L = \{x \in V(n, q) \mid x(KH)^T = 0\}$, KH 还是线性码 L 的校验矩阵.


另一种证明方法: 因为 K 可逆, 所以 $\text{Rank}(KH) = \text{Rank}(H) = n - k = \dim(L^\perp)$. 由于 KH 的每个行向量都是 H 的行向量的一个线性组合, 所以 KH 的每个行向量都是 L^\perp 的一个码字. 因此, KH 还是 L^\perp 的一个生成矩阵, 即 KH 为 L 的校验矩阵.

 **练习 7.5.15** 对于一个 q 元线性码 L , 如果对任意 $x, y \in L$, 都有 $x \cdot y = 0$, 则称 L 是自正交的. 假设 L 是一个自正交的二元线性码. 证明: L 中的所有码字都具有偶数重量, 并且分量全为 1 的向量 $11 \cdots 1 \in L^\perp$.

解: 证明: 因为 L 是二元自正交的, 所以对 $\forall x = x_1 x_2 \cdots x_n \in L$ 有

$$x \cdot x = \sum_{i=1}^n x_i^2 = (x_1 + x_2 + \cdots + x_n)^2 = 0.$$

因此 $x_1 + x_2 + \cdots + x_n = 0$, 且 $1 \cdot x = x_1 + x_2 + \cdots + x_n = 0$. 所以 L 中的所有码字都具有偶数重量, 并且分量全为 1 的向量 $11 \cdots 1 \in L^\perp$.

 **练习 7.5.16** 证明: q 元 $[n, k]$ 线性码 L 是自对偶的充分必要条件为 L 是自正交的, 并且 $k = n/2$.

解: 证明: 必要性. 由 $L = L^\perp$ 可知 L 是自正交的, $k = \dim L = \dim L^\perp = n - k$, $k = n/2$, n 为偶数.

充分性. 由 L 是自正交的可知 $L \subseteq L^\perp$, 再由 $k = n/2 = \dim L$ 可得 $\dim L^\perp = n - k = n/2$, 因此 $L = L^\perp$.

 **练习 7.5.17** 设 L 是一个二元 $[n, k]$ 线性码, 并且分量全为 1 的向量 $11 \cdots 1 \in L$. 证明: 对于

$i = 0, 1, \dots, n$, 我们有

$$A_i = A_{n-i}$$

其中 A_i 表示 L 中重量为 i 的码字的数目.

解: 证明: 令 $L_i = \{x \in L \mid \omega(x) = i\}$, 由于当 $x \in L$ 时, $x + 11 \cdots 1 \in L$, 所以 $x \mapsto x + 11 \cdots 1$ 是 L_i 到 L_{n-i} 间的一个一一对应. 于是 $|L_i| = |L_{n-i}|$, 从而 $A_i = A_{n-i}$.

 **练习 7.5.18** 设二元线性码 L 的生成矩阵为

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

试求 L 的重量分布.

解: L 的校验矩阵为

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

L 的对偶码 $L^\perp = \{00000, 10010, 11101, 01111\}$


$$W_{L^\perp}(z) = 1 + z^2 + 2z^4$$

由二元线性码的 Mac Williams 恒等式知

$$\begin{aligned} W_L(z) &= \frac{1}{2^2} (1+z)^5 W_{L^\perp} \left(\frac{1-z}{1+z} \right) \\ &= \frac{1}{4} (1+z)^5 \left[1 + \left(\frac{1-z}{1+z} \right)^2 + 2 \left(\frac{1-z}{1+z} \right)^4 \right] \\ &= 1 + 3z^2 + 3z^3 + z^5. \end{aligned}$$

因此线性码 L 的重量分布为

$$A_0 = 1, A_1 = 0, A_2 = 3, A_3 = 3, A_4 = 0, A_5 = 1.$$

 **练习 7.5.19** 设二元 $[9, 7]$ 线性码 L 的生成矩阵为

$$G = \left[I_7 \mid \begin{array}{cc} 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{array} \right].$$

试利用二元线性码的 Mac Williams 恒等式确定 A_0, A_1, A_2, A_3 的值, 并证明分量全为 1 的向量 $11 \cdots 1 \in L$. 进一步, 确定 L 的重量分布多项式

$$W_L(z) = \sum_{i=0}^9 A_i z^i$$

解: L 的校验矩阵为

$$H = \left(\begin{array}{ccccccc|cc} 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{array} \right).$$

L 的对偶码 $L^\perp = \{000000000, 001111110, 110011101, 111100011\}$,

$$W_{L^\perp}(z) = 1 + 3z^6.$$

由二元线性码的 Mac Williams 恒等式知

$$\begin{aligned} W_L(z) &= \frac{1}{2^2}(1+z)^9 W_{L^\perp}\left(\frac{1-z}{1+z}\right) \\ &= \frac{1}{2^2}(1+z)^9 \left[1 + 3\left(\frac{1-z}{1+z}\right)^6\right] \\ &= \frac{1}{2^2}[(1+z)^9 + 3(1-z)^6(1+z)^3] \\ &= 1 + 9z^2 + 27z^3 + 27z^4 + 27z^5 + 27z^6 + 9z^7 + z^9. \end{aligned}$$

由此可知 $A_0 = 1, A_1 = 0, A_2 = 9, A_3 = 27$, 再由 $A_9 = 1$ 知向量 $11 \cdots 1 \in L$.

练习 7.5.20 设 L 是一个二元线性码, L_0 是 L 中所有具有偶数重量的码字构成的码. 证明:

$$W_{L_0}(z) = \frac{1}{2}(W_L(z) + W_L(-z)).$$

解: 证明:

$$\begin{aligned} &\frac{1}{2}(W_L(z) + W_L(-z)) \\ &= \frac{1}{2}[(A_0 + A_1z + \cdots + A_nz^n) + (A_0 + A_1(-)z + \cdots + A_n(-)z^n)] \\ &= A_0 + A_2z^2 + \cdots + A_{\lfloor \frac{n}{2} \rfloor} z^{\lfloor \frac{n}{2} \rfloor} = W_{L_0}(z). \end{aligned}$$

练习 7.5.21 设 L 是一个二元线性码, \hat{L} 是由 L 增加一个奇偶校验位构成的扩充码. 证明:

$$W_{\hat{L}}(z) = \frac{1}{2}[(1+z)W_L(z) + (1-z)W_L(-z)].$$

解: 证明: 设 $W_L(z) = \sum_{i=0}^n A_i z^i$, 则

$$\begin{aligned} W_{\hat{L}}(z) &= \sum_{0 \leq i \leq n+1, i \text{ 是偶数}} (A_i + A_{i-1}) z^i \\ &= \sum_{0 \leq i \leq n+1, i \text{ 是偶数}} A_i z^i + \sum_{0 \leq i \leq n, i \text{ 是奇数}} A_i z^{i+1} \\ &= \frac{1}{2}(W_L(z) + W_L(-z)) + \frac{1}{2}z(W_L(z) - W_L(-z)) \\ &= \frac{1}{2}[(1+z)W_L(z) + (1-z)W_L(-z)]. \end{aligned}$$

第 8 章 Hamming 码

8.1 Hamming 码的定义

8.1.1 Hamming 码的构造

通过构造一个 $r \times n$ 阶校验矩阵来构造 q 元 Hamming 码, 令 $Y_1 = V(r, q)$, 任取非零向量 $y_1 \in Y_1$, 令

$$Y_2 = Y_1 - \{\alpha y_1 \mid \alpha \in F_q, \alpha \neq 0\}$$

任取非零向量 $y_2 \in Y_2$, 令

$$Y_3 = Y_2 - \{\beta y_2, \alpha y_1 \mid \beta, \alpha \in F_q, \beta, \alpha \neq 0\}$$

依次下去, 任取 $y_i \in V(r, q)$, 但 y_i 不是前面已取的 $i-1$ 个向量中的任何一个的非零倍数, 此过程一直进行到 $V(r, q)$ 中没有具有此性质的向量为止.

令 $H = (y_1^T, y_2^T, \dots, y_i^T, \dots)$, 称 H 为 r 阶 Hamming 矩阵, 以 H 为校验矩阵的线性码称为 r 阶 q 元 Hamming 码, 记为 $\text{Ham}(r, q)$.

因为

$$|\{\alpha y_i \mid \alpha \in F_q, \alpha \neq 0\}| = q - 1,$$

所以校验矩阵 H 总共有 $(q^r - 1)/(q - 1)$ 列, 这也就是说, q 元 Hamming 码 $\text{Ham}(r, q)$ 的码长为 $n = (q^r - 1)/(q - 1)$. 事实上, 上述校验矩阵的构造过程相当于把 $V(r, q)$ 中的 $q^r - 1$ 个非零向量分成 $(q^r - 1)/(q - 1)$ 个类, 使得 $V(r, q)$ 中任意两个非零向量线性相关的充分必要条件是它们在同一类中. 在每类中任取一个向量作为校验矩阵的列.

应当指出, 对于给定的参数 r 和 q , 可以选取不同的校验矩阵来定义 q 元 Hamming 码 $\text{Ham}(r, q)$, 但通过列的置换以及某些列与 F_q 中非零元素的相乘, 任意两个不同的校验矩阵可以相互转化. 因此, 在等价的意义上, q 元 Hamming 码 $\text{Ham}(r, q)$ 是唯一的.

8.2 Hamming 码的性质

定理 8.2.1

q 元 Hamming 码 $\text{Ham}(r, q)$ 是 q 元 $[n, n - r, 3]$ 线性码, 其中 $n = \frac{q^r - 1}{q - 1}$.



证明 q 元 Hamming 码 $\text{Ham}(r, q)$ 的检验阵 $H_{r \times n}$ 为 $\text{Ham}(r, q)^\perp$ 的生成阵, 故 $\text{Ham}(r, q)$ 为 $[n, n - r]$ 线性码, 由 H 的构造方法知 $V(r, q)$ 中共有 $q^r - 1$ 个非零向量. $\forall x \in V(r, q) \quad x \neq 0, \quad \alpha \in F_q, \quad \alpha \neq 0, \alpha x$ 与 x 线性相关, αx 共有 $q - 1$ 个, 于是可将 $V(r, q)$ 中 $q^r - 1$ 个非零向量分组, 每组有 $q - 1$ 个, 故共有 $\frac{q^r - 1}{q - 1}$ 组, 每组取一个向量作为 H 的 1 列, 共有 $\frac{q^r - 1}{q - 1}$ 列, 故 $n = \frac{q^r - 1}{q - 1}$, 每

个向量 x 为 r 长, 故 x^T 作为列, H 是 r 行 n 列矩阵. 由 H 的构造可知, H 的任意两列线性无关. 故令 $H = (h_1, h_2, \dots, h_n)$, 有 $h_1 + h_2 \neq 0$ (否则相关) 故 $\exists i$, 有 $h_1 + h_2 = \alpha h_i$ (因非零向量都在 H 的列分组中) 因此 H 中存在 3 列线性相关, 从而 $\text{Ham}(r, q)$ 的最小距离为 3.

例题 8.2.1 $V(2, 3)$ 中, $r = 2, q = 3, |V(2, 3)| = 3^2 = 9$, 故 $V(2, 3)$ 中有 8 个非零向量, 按 Hamming 码的构造方法进行分组,

$$\begin{aligned} & \text{取 } \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{ 则 } \alpha \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \cdot \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}. \\ & \text{取 } \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ 则 } \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\} \\ & \text{取 } \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ 则 } \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\} \\ & \text{取 } \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \text{ 则 } \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\} \end{aligned}$$

于是 $\text{Ham}(2, 3)$ 的校验矩阵为 $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix}$.

$n = \frac{q^r - 1}{q - 1} = \frac{3^2 - 1}{3 - 1} = \frac{8}{2} = 4$, 码长为 4. $\text{Ham}(r, q)$ 是一个 $[4, 2, 3]$ 码.

例题 8.2.2 $V(3, 3)$ 中非零码字个数为 $3^3 - 1 = 26$ 个, H 的列数为 $\frac{q^r - 1}{q - 1} = \frac{26}{2} = 13$.

$$\begin{aligned} & \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right\}, \\ & \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \right\}, \\ & \left\{ \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \right\}, \\ & \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \right\}. \end{aligned}$$

于是有 $H = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \end{pmatrix}$. $\text{Ham}(3, 3)$ 是一个 $[13, 10, 3]$ 码.

定理 8.2.2

q 元 Hamming 码 $\text{Ham}(r, q)$ 是完备的 (达到 Hamming 界).



证明 Ham(r, q) 的最小距离为 3, 故可纠正 1 个错误, 我们有

$$\begin{aligned}
 M & \left(\binom{n}{0} + \binom{n}{1} (q-1) + \binom{n}{2} (q-1)^2 + \cdots + \binom{n}{t} (q-1)^t \right) \\
 & = q^{n-r} \left(1 + \binom{n}{1} (q-1) \right) \\
 & = q^{n-r} \left[1 + \frac{q^r-1}{q-1} (q-1) \right] \\
 & = q^{n-r} \cdot q^r = q^n
 \end{aligned}$$

于是 Hamming 码是完备码.

8.3 Hamming 码的译码方法

Ham(r, q) 是 q 元 $[n, n-r, 3]$ 码, 其中 $n = \frac{q^r-1}{q-1}$, 码字个数 $M = q^{n-r}$, 故 $V(n, q)$ 中不同的 Ham(r, q) 的陪集个数为 $\frac{q^n}{M} = \frac{q^n}{q^{n-r}} = q^r$ ($a+L$ 为陪集, $|a+L| = |b+L| = |L|$, 无交), 故非零陪集头的个数为 $q^r - 1$.

记

$$A = \{0 \cdots 0x_i 0 \cdots 0 \in V(n, q) \mid 0 \neq x_i, 1 \leq i \leq n\}.$$

对 $\forall x = 0 \cdots 0x_i 0 \cdots 0, y = 0 \cdots 0x_j 0 \cdots 0 \in A, i \neq j$, 则:

$$xH^T = x_i h_i^T \neq 0 \quad (\text{域中非零元乘积不等于零})$$

$$yH^T = x_j h_j^T \neq 0$$

$$(x - y)H^T = x_i h_i^T - x_j h_j^T \neq 0$$

(Ham(r, q) 的校验矩阵 H 的任意两列线性无关) 从而 A 中任意两个不同的向量属于不同的陪集, 且 $\forall x \in A, x \notin \text{Ham}(r, q)$, 又

$$|A| = n(q-1) = q^r - 1$$

故 A 中的向量就是 Ham(r, q) 的标准阵的全部陪集头.

q 元 Hamming 码 Ham(r, q) 的译码过程

- (1) 设 x 是接收到的字, 计算 x 的伴随式 $S(x) = xH^T$;
- (2) 如果 $S(x) = 0$, 则 x 没有发生错误, x 即为信道发送端发送的码字;
- (3) 如果 $S(x) \neq 0$, 则 $S(x) = bh_i^T, (0 \cdots 0b0 \cdots 0) \in A$, 其中 h_i 为 Ham(r, q) 的校验矩阵 H 的第 i 列, $0 \neq b \in F_q, 1 \leq i \leq n$, 这时第 i 个位置发生错误, x 破译为码字 $x - a_i$, 其中

$$a_i = (0 \cdots 0b0 \cdots 0)$$

例题 8.3.1 设 Ham(2, 5) 校验矩阵为 $H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 & 3 & 4 \end{pmatrix}$, 设在信道接收端收到的字为 $x = 203031$, 试将 x 译码.

计算

$$S(x) = xH^T = (203031) \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 18 \end{pmatrix} = \begin{pmatrix} 2 & 3 \end{pmatrix} = 2 \begin{pmatrix} 1 & 4 \end{pmatrix}$$

故 $b = 2$, 将 x 译为 $x - a_i = 203031 - 000002 = 203034$

注: 若 $q = 2$, $\text{Ham}(r, 2)$ 是一个二元 $[2^r - 1, 2^r - 1 - r, 3]$ 线性码, 其校验矩阵的列向量是 $V(r, 2)$ 中所有非零向量, $V(r, 2)$ 中所有非零向量是 1 到 $2^r - 1$ 之间的所有整数的二进制表示.

例题 8.3.2 设 $\text{Ham}(2, 2)$ 校验矩阵为 $H = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, 其生成矩阵为 $G = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$, 故 $\text{Ham}(2, 2)$ 是码长为 3 的二元重复码.

例题 8.3.3 $\text{Ham}(3, 2)$ 的校验矩阵为 $H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$, 其中 H 的第 i 列是整数 i 的

二进制表示, $1 \leq i \leq 7$. 如果将 H 的列重新排列, 可得 H 的标准型. $H' = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$

因此, $\text{Ham}(3, 2)$ 的生成矩阵为 $G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$

二元 Hamming 码 $\text{Ham}(r, 2)$ 的非零陪集头的集合为

$$A = \{\underbrace{0 \cdots 0}_{i-1} 1 0 \cdots 0 \in V(n, 2) \mid 1 \leq i \leq n\}.$$

每个陪集头的伴随为

$$\begin{aligned} S(\underbrace{0 \cdots 0}_{i-1} 1 0 \cdots 0) &= (\underbrace{0 \cdots 0}_{i-1} 1 0 \cdots 0) H^T \\ &= h_i^T, \end{aligned}$$

其中 h_i^T 校验矩阵 H 的第 i 列的转置. 如果校验矩阵 H 的第 i 列是整数 i 的二进制表示, $1 \leq i \leq n$, 则二元 Hamming 码的译码过程将非常简明有效. 下面列出二元 Hamming 码 $\text{Ham}(r, 2)$ 的译码过程.

- (1) 设 x 是在信道接收端接收到的向量, 计算其伴随式 $S(x) = xH^T$.
- (2) 如果 $S(x) = 0$, 则没有发生错误, x 就是在信道发送端发送的码字.
- (3) 如果 $S(x) \neq 0$, 则有一个错误发生, $S(x)$ 就是错误位置的二进制表示. 将错误位置上的 0 变为 1, 1 变为 0 即可.

例题 8.3.4 设 $\text{Ham}(3, 2)$ 校验矩阵为 $H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$. 设收到字 $x = 0110110$, 则

$$S(x) = xH^T = 010$$

说明第 2 个位置发生错误, 将 x 译为 0010110.

8.4 二元 Hamming 码的对偶码

定义 8.4.1 (极长码)

二元 Hamming 码 $\text{Ham}(r, 2)$ 的对偶码称为极长码, 记为 \sum_r .



注: $\text{Ham}(r, 2)$ 是一个 $[2^r - 1, 2^r - 1 - r, 3]$ 线性码, 故极长码 \sum_r 是一个二元 $[2^r - 1, r]$ 线性码, 并且其生成矩阵 G_r 正好是二元 Hamming 码 $\text{Ham}(r, 2)$ 的校验矩阵.

例题 8.4.1 极长码 \sum_2 的生成矩阵为 G_2 , $\text{Ham}(2, 2)$ 的对偶码为 \sum_2 , \sum_2 的生成阵即为 $\text{Ham}(2, 2)$ 的校验阵, 故

$$G_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad \sum_2 = \{000, 011, 101, 110\}$$

定理 8.4.1

极长码 \sum_r 具有下列性质:

- (1) \sum_r 中任意一个非零码字的重量都是 2^{r-1} , 因此 \sum_r 是一个二元 $[2^r - 1, r, 2^{r-1}]$ 线性码;
- (2) \sum_r 中任意两个码字的距离都是 2^{r-1} .



证明 (1) 设 \sum_r 的生成矩阵, 即二元 Hamming 码 $\text{Ham}(r, 2)$ 的校验矩阵为

$$H = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ h_{r1} & h_{r2} & \cdots & h_{rn} \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_r \end{pmatrix}$$

其中 $n = 2^r - 1$, h_1, h_2, \dots, h_r 表示 H 的 r 个行向量.

任取一个非零码字 $c \in \sum_r$, 则 c 一定是 h_1, h_2, \dots, h_r 的非零线性组合, 即存在不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_r \in F_2$, 使得

$$c = \sum_{i=1}^r \lambda_i h_i$$

因此,

$$c \text{ 的第 } j \text{ 个坐标为 } 0 \iff \sum_{i=1}^r \lambda_i h_{ij} = 0 \iff \sum_{i=1}^r \lambda_i x_i = 0,$$

其中 $(x_1, x_2, \dots, x_r)^T$ 是 H 的第 j 列. 设 $C_1 \subset V(r, 2)$ 是以 $(\lambda_1, \lambda_2, \dots, \lambda_r)$ 为校验矩阵的线性码, 则 C_1 是一个二元 $[r, r-1]$ 线性码, 并且

$$C_1 = \left\{ x_1 x_2 \cdots x_r \in V(r, 2) \mid \sum_{i=1}^r \lambda_i x_i = 0 \right\}.$$

因为 H 的列向量是 $V(r, 2)$ 中所有的非零向量, 所以码字 c 中为零的分量个数 $n_0(c)$ 就是 C_1 中所有非零向量的个数, 即

$$n_0(c) = |C_1| - 1.$$

由于 $|C_1| = 2^{r-1}$, 所以 $n_0(c) = 2^{r-1} - 1$. 因此,

$$\omega(c) = n - n_0(c) = 2^{r-1}.$$

(2) 设 $x, y \in \sum_r$, 并且 $x \neq y$, 则 $0 \neq x - y \in \sum_r$,

$$d(x, y) = \omega(x - y) = 2^{r-1}.$$

推论 8.4.1

二元 Hamming 码 $\text{Ham}(r, 2)$ 的重量分布多项式为

$$W_L(z) = \frac{1}{2^r} \left((1+z)^n + n(1-z^2)^{\frac{n-1}{2}} (1-z) \right)$$

其中 $n = 2^r - 1$.



证明 事实上, 因为 L 是一个二元 $[2^r - 1, 2^r - 1 - r]$ 线性码, 所以 L^\perp 是一个二元 $[2^r - 1, r]$ 线性码. 由于 L^\perp 中非零码字的重量都是 2^{r-1} , 所以

$$W_{L^\perp}(z) = \sum_{x \in L^\perp} z^{\omega(x)} = (2^r - 1) z^{2^{r-1}} + 1.$$

根据二元线性码的 Mac Williams 恒等式, 我们有

$$\begin{aligned} W_L(z) &= \frac{1}{2^r} (1+z)^n W_{L^\perp} \left(\frac{1-z}{1+z} \right) \\ &= \frac{1}{2^r} (1+z)^n \left(1 + n \left(\frac{1-z}{1+z} \right)^{\frac{n+1}{2}} \right) \\ &= \frac{1}{2^r} \left((1+z)^n + n(1-z^2)^{\frac{n-1}{2}} (1-z) \right). \end{aligned}$$


8.5 习题课

8.5.1 基本概念及方法

1. Hamming 码的构造方法
2. Hamming 码的性质
3. Hamming 码的译码

4. 极长码及性质

8.5.2 课后习题

 **练习 8.5.1** 试求二元 Hamming 码 $\text{Ham}(3, 2)$ 的包含陪集头和对应伴随式列表, 并对在信道接收端接收到的字 0000011, 1111111, 1100110, 1010101 分别进行译码.

解: 二元 Hamming 码 $\text{Ham}(3, 2)$ 的校验矩阵为 $H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$

伴随式列表:


陪集头 x_i	伴随式 $x_i H^T$
1000000	001
0100000	010
0010000	011
0001000	100
0000100	101
0000010	110
0000001	111

$$(00000011)H^T = 001 \quad 0000011 \text{ 译为 } 1000011.$$

$$(11111111)H^T = 000 \quad 1111111 \text{ 译为 } 1111111.$$

$$(1100110)H^T = 011 \quad 1100110 \text{ 译为 } 1110110.$$

$$(1010101)H^T = 000 \quad 1010101 \text{ 译为 } 1010101.$$

 **练习 8.5.2** 试求二元 Hamming 码 $\text{Ham}(4, 2)$ 中重量分别为 1, 2, 3, 4 的码字的个数.


解: 只需求出码 $\text{Ham}(4, 2)$ 的重量分布多项式即可.

$$W_{\text{Ham}(4,2)}(z) = \frac{1}{2^4} \left[(1+z)^{15} + 15(1-z^2)^7(1-z) \right]$$

由二项式展开定理, 计算整理得:

$$W_{\text{Ham}(4,2)}(z) = 1 + 35z^3 + 105z^4$$

于是, 二元 Hamming 码 $\text{Ham}(4, 2)$ 中重量为 1, 2, 3, 4 的码字的个数分别为 0, 0, 35, 105.


 **练习 8.5.3** 写出七元 Hamming 码 $\text{Ham}(2, 7)$ 的校验矩阵 H , 并对在信道接收端接收到的字 35234106 和 10521360 分别进行译码.

解: 七元 Hamming 码 $\text{Ham}(2, 7)$ 的校验矩阵为

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

$$(35234106) \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{pmatrix} = (0 \ 0), \quad (10521360) \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{pmatrix} = (3 \ 6) = 3 \times (1 \ 2)$$

35234106 译为 35234106. 10521360 译为 10561360.

 **练习 8.5.4** 设二元 Hamming 码 $\text{Ham}(4, 2)$ 的校验矩阵为

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$


试对在信道接收端接收到的字 011011001111000 和 001100110011000 分别进行译码.

解:

$$(011011001111000) \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} = (0110), (001100110011000) \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} = (1111)$$

011011001111000 译为 011010001111000.

001100110011000 译为 001100110011001.

 **练习 8.5.5** 写出三元 Hamming 码 $\text{Ham}(3, 3)$ 的校验矩阵 H , 并对在信道接收端接收到的字 0122100110022 和 2211001012020 分别进行译码.

解:

$$H = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 2 & 1 & 2 & 2 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 0 & 2 & 2 & 1 \end{pmatrix}$$

$$(0122100110022) \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \\ 1 & 2 & 1 \end{pmatrix} = (010) = 1 \times (010), (2211001012020) \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \\ 1 & 2 & 1 \end{pmatrix} = (200) = 2 \times (100)$$

2211001012020 译为 2221001012020.

 **练习 8.5.6** 设 q 是一个素数的幂次方, 并且 $3 \leq n \leq q+1$. 证明:

$$A_q(n, 3) = q^{n-2}.$$

解: 证明: 设 C 是一个 $(n, M, 3)$ 码, 在 C 中的码字去掉任意两个坐标位置所得到的向量一定两两不同, 否则 $d(C) \leq 2$, 与 $d(C) = 3$ 矛盾. 因此, $M \leq q^{n-2}$. 因 q 元汉明码 $\text{Ham}(2, q)$ 的码长为

$$n = q^2 - 1/q - 1 = q + 1$$

码字个数为 q^{n-2} 个. 因 $A_q(q+1, 3) = q^{(q+1)-2}$. 对 $3 \leq n < q+1$, 将 $\text{Ham}(2, q)$ 的校验阵 H 去掉 $q+1-n$ 列, 得到一个矩阵 H' , 显然 H' 中任意两列还是线性无关的, 适当选取 H 中去掉的列, 可以保证 H' 中存在 3 列线性相关, 于是 H' 为校验阵的线性码 C' 是一个 q 元 $[n, n-2, 3]$ 线性码, C' 中有 q^{n-2} 个码字. 因此, 对于 $3 \leq n \leq q+1$,

$$A_q(n, 3) = q^{n-2}$$

 **练习 8.5.7** 确定二元 Hamming 码 $\text{Ham}(r, 2)$ 中重量为 3 的码字的个数 A_3 .

解: $W_L(z) = \frac{1}{2^r} \left[(1+z)^n + n(1-z^2)^{\frac{n-1}{2}}(1-z) \right]$, 其中, $n = 2^r - 1$, 由二项式展开定理, 计算整理得:

$$W_L(z) = \frac{1}{2^r} \left[(\cdots + C_n^3 z^3 + \cdots) + n C_{\frac{n-1}{2}}^1 z^3 \right],$$

z^3 的系数为 $\frac{1}{2^r} \left(C_n^3 + n C_{\frac{n-1}{2}}^1 \right) = \frac{n(n-1)}{6}$. 故重量为 3 的码字的个数 $A_3 = \frac{n(n-1)}{6}$.

第 9 章 循环码

9.1 循环码的定义

$$F_q[x] = \{a_0 + a_1x + \cdots + a_mx^m \mid a_i \in F_q\}$$
$$p(x) = a_0 + a_1x + \cdots + a_mx^m \in F_q[x]$$

若 $a_m \neq 0$, 则称 m 为 $p(x)$ 的次数, 记为 $\deg(p(x))$; a_m 称为 $p(x)$ 的首项系数; 如果 $a_m = 1$, $p(x)$ 称为首 1 多项式. $F_q[x]$ 在通常的多项式的加法和乘法运算下构成一个交换环 (有单位元).

设 $p(x) \in F_q[x]$, $\langle p(x) \rangle$ 表示由 $p(x)$ 生成的 $F_q[x]$ 的主理想, 则

$$\langle p(x) \rangle = \{f(x)p(x) \mid f(x) \in F_q[x]\}.$$

设 $f(x), g(x) \in F_q[x]$, 商环 $F_q[x]/\langle p(x) \rangle$ 上的加法和乘法运算:

$$(f(x) + \langle p(x) \rangle) + (g(x) + \langle p(x) \rangle) = (f(x) + g(x)) + \langle p(x) \rangle$$
$$(f(x) + \langle p(x) \rangle)(g(x) + \langle p(x) \rangle) = f(x)g(x) + \langle p(x) \rangle$$

则 $F_q[x]/\langle p(x) \rangle$ 在上述运算下构成一个交换环.

$$F_q[x]/\langle p(x) \rangle = \{f(x) + \langle p(x) \rangle \mid f(x) \in F_q[x]\}$$
$$p(x) = a_0 + a_1x + \cdots + a_mx^m$$
$$f(x) = p(x)q(x) + r(x); \deg(r(x)) < m$$

所以

$$F_q[x]/\langle p(x) \rangle = \{r(x) + \langle p(x) \rangle \mid r(x) \in F_q[x], \deg(r(x)) < m\}$$
$$r(x) = b_0 + b_1x + \cdots + b_{m-1}x^{m-1}, b_i \in F_q, i = 0, 1, \dots, m-1$$

故 $|F_q[x]/\langle p(x) \rangle| = q^m$.

定理 9.1.1

$F_q[x]/\langle p(x) \rangle$ 是域的充分必要条件是 $p(x)$ 为 $F_q[x]$ 中的不可约多项式.



定义 9.1.1 (循环码)

设 $L \subseteq V(n, q)$, L 是一个线性码, 对 $\forall c \in L$, 如果 c 的循环移位仍是一个码字, 即若 $c_0c_1c_2 \cdots c_{n-1} \in L$, 则 $c_{n-1}c_0c_1c_2 \cdots c_{n-2} \in L$, 则称 L 为循环码.



记 $R_n = F_q[x]/\langle x^n - 1 \rangle$, 定义 $\varphi: V(n, q) \rightarrow R_n$

$$c_0c_1 \cdots c_{n-1} \mapsto c_0 + c_1x + \cdots + c_{n-1}x^{n-1}$$

φ 是一一映射. 在此对应下, 线性码 L 可看作 R_n 的子集, $L \subseteq R_n$,

$$\begin{aligned} \text{对 } \forall a(x) &= a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} \in L \\ xa(x) &= (a_0x + a_1x^2 + \cdots + a_{n-1}x^n) \pmod{(x^n - 1)} \\ &= a_{n-1} + a_0x + a_1x^2 + \cdots + a_{n-2}x^{n-1} \end{aligned}$$

$a_0a_1 \cdots a_{n-1} \rightarrow a_{n-1}a_0a_1 \cdots a_{n-2}$ 循环移位.

定理 9.1.2

一个码 $L \subseteq R_n$ 是循环码的充分必要条件为 L 满足下列两个条件.

- (1) 如果 $a(x), b(x) \in L$, 则 $a(x) + b(x) \in L$.
- (2) 如果 $a(x) \in L, r(x) \in R_n$, 则 $r(x)a(x) \in L$.



证明 必要性. 设 L 是循环码, 则 L 是线性码. 因此, (1) 显然成立. 设 $a(x) \in L$,

$$r(x) = r_0 + r_1x + r_2x^2 + \cdots + r_{n-1}x^{n-1} \in R_n.$$

因为 $x^i a(x)$ 等价于将码字循环右移 i 位, $1 \leq i \leq n-1$, 所以

$$r(x)a(x) = r_0a(x) + r_1xa(x) + \cdots + r_{n-1}x^{n-1}a(x) \in L.$$

充分性. 设 (1) 和 (2) 成立. 令 $r(x) = r_0 \in F_q$, 则由 (1) 和 (2) 知, L 是线性码. 取 $r(x) = x$, 则由 (2) 知, 知 L 是循环码.

显然, 定理 9.1.2 等价于说, 一个码 $L \subseteq R_n$ 是循环码的充分必要条件为 L 是 R_n 的理想. 设 $p(x) \in R_n$, 我们知道

$$\langle p(x) \rangle = \{p(x)f(x) \mid f(x) \in R_n\}$$

是 R_n 的理想. 于是, 根据定理 9.1.2, 可得下面的结论.

定理 9.1.3

对于任意 $p(x) \in R_n$, $\langle p(x) \rangle$ 是一个循环码.



注: 在本章, 关于多项式的加法和乘法运算, 若没有特别声明, 均指在 R_n 中的加法和乘法. 即模 $x^n - 1$ 的加法和乘法

9.2 循环码的性质

定理 9.2.1

设 $\{0\} \neq C \subseteq R_n$ 是一个循环码, 则

- (1) C 中存在唯一一个具有最低次数的首 1 多项式 $g(x)$ 生成 C , 即 $C = \langle g(x) \rangle$.
- (2) $g(x) \mid x^n - 1$.
- (3) 如果 $\deg(g(x)) = r$, 则 $\dim C = n - r$. 事实上

$$C = \langle g(x) \rangle = \{r(x)g(x) \mid \deg(r(x)) < n - r, r(x) \in R_n\}.$$

(4) 如果 $g(x) = g_0 + g_1x + g_2x^2 + \cdots + g_rx^r$, 则 $g_0 \neq 0$, 并且 C 的生成矩阵为

$$G = \begin{pmatrix} g_0 & g_1 & g_2 & \cdots & g_r & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & g_2 & \cdots & g_r & 0 & \cdots & 0 \\ 0 & 0 & g_0 & g_1 & g_2 & \cdots & g_r & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & g_0 & g_1 & g_2 & \cdots & g_r \end{pmatrix},$$

其中 G 的每一行都是前一行的循环移位.



证明 (1) 先证明次数最低的首 1 多项式是唯一的.

设 $g_1(x), g_2(x)$ 是 C 中的两个不同的具有最低次数 r 的首 1 多项式, 则 $g_1(x) - g_2(x) \neq 0$, $\deg(g_1(x) - g_2(x)) < r$, 而 $g_1(x) - g_2(x) \in C$, 与 $g_1(x), g_2(x)$ 次数最低矛盾, 因此 C 中只有一个次数最低的首 1 多项式, 设为 $g(x)$, $g(x) \in C$, 故 $\langle g(x) \rangle \subseteq C$, ($\langle g(x) \rangle$ 为包含 $g(x)$ 的最小理想, $g(x) \in C$, C 也是 R_n 的理想, 故 $\langle g(x) \rangle \subseteq C$.)

另一方面, 设 $\forall p(x) \in C, p(x) = g(x)q(x) + r(x)$, 其中 $\deg(r(x)) < \deg(g(x))$, 于是 $r(x) = p(x) - q(x)g(x) \in C$, 与 $g(x)$ 次数最低矛盾, 故 $r(x) = 0$, 即 $p(x) = g(x)q(x) \in \langle g(x) \rangle, C \subseteq \langle g(x) \rangle$, 故 $C = \langle g(x) \rangle$.

(2) 设 $\deg(g(x)) = r$, 用 $g(x)$ 去除 $x^n - 1$, 可得: $x^n - 1 = q(x)g(x) + r(x)$, 其中 $\deg(r(x)) = r$, 在 R_n 中, $x^n - 1 \equiv 0 \pmod{(x^n - 1)} \in C$, 所以 $r(x) = -q(x)g(x) \in C$, 因为 $g(x)$ 为 C 中唯一的具有最低次数的首 1 多项式, 所以有 $r(x) = 0$, 于是有 $g(x) \mid x^n - 1$.

(3) $g(x)$ 生成的理想为

$$\langle g(x) \rangle = \{f(x)g(x) \mid f(x) \in R_n\},$$

现在我们来证明

$$\langle g(x) \rangle = \{r(x)g(x) \mid \deg(r(x)) < n - r, r(x) \in R_n\}.$$

因为 $g(x) \mid x^n - 1$, 所以 $x^n - 1 = g(x)h(x), h(x) \in R_n, \deg(h(x)) = n - r, \forall f(x) \in R_n$, 用 $h(x)$ 去除 $f(x)$ 可得

$$f(x) = q(x)h(x) + r(x), \deg(r(x)) < \deg(h(x)) = n - r$$

于是 $f(x)g(x) = q(x)h(x)g(x) + r(x)g(x) = q(x)(x^n - 1) + r(x)g(x)$ 因此在 R_n 中, $f(x)g(x) = r(x)g(x)$, 即

$$\langle g(x) \rangle = \{r(x)g(x) \mid \deg(r(x)) < n - r, r(x) \in R_n\}$$

于是有 $\{g(x), xg(x), \cdots, x^{n-r-1}g(x)\}$ 生成 C . 由于它们线性无关, 故它们是 C 的一组基, $\dim(C) = n - r$.

(4) 假设 $g_0 = 0$, 则 $g(x) = xg_1(x)$, 其中

$$g_1(x) = g_1 + g_2x + g_3x^2 + \cdots + g_rx^{r-1}, \deg(g_1(x)) < r$$

于是 $g_1(x) = 1 \cdot g_1(x) = x^n g_1(x) \pmod{(x^n - 1)} = x^{n-1}g(x) \in C$ (因 C 是循环码), 与 $g(x)$ 次数

最低矛盾, 因此 $g_0 \neq 0$, 因为 $\{g(x), xg(x), \dots, x^{n-r-1}g(x)\}$ 是 C 的一组基, 故 G 是 C 的生成矩阵.

注: 次数最低的首项系数为 1 的多项式 $g(x)$, 若 $C = \langle g(x) \rangle$, 则称 $g(x)$ 为 C 的生成多项式.

例题 9.2.1 考虑码长为 3 ($n = 3$) 的二元循环码, $C = \langle 1+x \rangle, C \subseteq R_3, R_3 = F_2[x]/\langle x^3-1 \rangle$. $g(x) = 1+x$, $r = \deg(g(x)) = 1, \dim(C) = n-r = 3-1 = 2$, 并且 C 中含有下列码字:

$$\langle 1+x \rangle = \{r(x)g(x) \mid \deg(r(x)) < n-r = 2\}$$


$$0, 1+x, x(1+x) = x+x^2, (1+x)(1+x) = 1+x^2$$

于是 $C = \{0, 1+x, x+x^2, 1+x^2\} = \{000, 110, 011, 101\}$. 事实上, 我们还可以验证

$$\langle 1+x^2 \rangle = \{f(x)(1+x^2) \mid f(x) \in R_3\} = C.$$

这说明循环码 C 也可由 $1+x^2$ 来生成.

定理 9.2.2

R_n 中首 1 多项式 $p(x)$ 是循环码 $C = \langle p(x) \rangle$ 的生成多项式的充要条件为 $p(x) \mid x^n - 1$. 

证明 只证明充分性.

设 $p(x) \mid x^n - 1, g(x)$ 是循环码, $C = \langle p(x) \rangle$ 的生成多项式. 假设 $p(x) \neq g(x)$. 由于 $g(x), p(x)$ 都是首 1 的, 则有 $\deg(p(x)) > \deg(g(x))$

因为 $p(x) \mid x^n - 1$, 故存在 $f(x) \in R_n$, 使得 $x^n - 1 = f(x)p(x)$, 又 $g(x) \in \langle p(x) \rangle$, 所以有

$$g(x) \equiv a(x)p(x) \pmod{(x^n - 1)},$$

其中 $a(x) \in R_n$. 由此可得

$$g(x)f(x) \equiv a(x)p(x)f(x) \equiv a(x)(x^n - 1) \equiv 0 \pmod{(x^n - 1)},$$

但 $\deg(f(x)g(x)) < \deg(f(x)p(x)) = n$, 从而 $f(x)g(x) = 0$, 不可能, 因此 $p(x) = g(x)$.

因此, 只需将 $x^n - 1$ 分解为 F_q 上的首 1 不可约多项式的乘积就可构造出码长为 n 的所有 q 元循环码.

例题 9.2.2 试找出 R_3 中的所有二元循环码.

解: 二元域 F_2 中 $x^3 - 1 = (x+1)(x^2+x+1)$, 其中 $x+1$ 和 x^2+x+1 都是 F_2 上的不可约多项式.

生成多项式	R_3 中的码	$V(3, 2)$ 中的码
1	R_3	$V(3, 2)$
$1+x$	$\{0, 1+x, x+x^2, 1+x^2\}$	$\{000, 110, 011, 101\}$
$1+x+x^2$	$\{0, 1+x+x^2\}$	$\{000, 111\}$
x^3-1	$\{0\}$	$\{000\}$

例题 9.2.3 写出 R_4 中的码长为 4 的所有三元循环码.

解: 在二元域 F_3 上

$$x^4 - 1 = (x-1)(x+1)(1+x^2) = (x+2)(x+1)(1+x^2),$$

其中 $x+2, x+1, x^2+1$ 在 F_3 上不可约.

列表:

生成多项式	R_4 中的码	$V(4, 3)$ 中的码
1	R_4	$V(4, 3)$
$2 + x$	$3^3 = 27$ 个元	{0000, 2100, 0210, 0021, 1002, 1200, 0120, 0012, 2001, 1020, 0102, 2010, 0201, 1110, 0111, 1011, 1101, 2220, 0222, 2022, 2202, 2211, 1221, 1122, 2112, 2121, 1212}
$1 + x$	$3^3 = 27$ 个元	{0000, 1100, 0110, 0011, 1001, 2200, 0220, 0022, 2002, 1020, 0102, 2010, 0201, 1210, 0121, 1012, 2101, 2120, 0212, 2021, 1202, 2112, 2211, 1221, 1122, 1111, 2222}
$1 + x^2$	$3^2 = 9$ 个元	{0000, 1010, 0101, 2020, 0202, 1212, 2121, 1111, 2222}
$2 + x^2$	$3^2 = 9$ 个元	{0000, 2010, 0201, 1020, 0102, 1122, 2112, 2211, 1221}
$2 + x + 2x^2 + x^3$	$3^1 = 3$ 个元	{0000, 2121, 1212}
$1 + x + x^2 + x^3$	$3^1 = 3$ 个元	{0000, 1111, 2222}
$x^4 - 1$	{0}	{000}

$(a + bx + cx^2)(2 + x), a, b, c \in F_3, 3^3 = 27$ 个元;

$(a + bx)(1 + x^2), a, b \in F_3, 3^2 = 9$ 个元;

$a(2 + x + 2x^2 + x^3), a \in F_3, 3^1 = 3$ 个元.

9.3 循环码的校验矩阵及其对偶码

设 $C = \langle g(x) \rangle$, $\deg(g(x)) = r$, C 是一个 $[n, n - r]$ 循环码, 并且存在 $h(x) \in R_n$, 使得

$$x^n - 1 = h(x)g(x).$$

$h(x)$ 为首 1 多项式, $\deg(h(x)) = n - r$, $h(x)$ 称为循环码 C 的校验多项式.

定理 9.3.1

设 $h(x)$ 是循环码 $C = \langle g(x) \rangle$ 的校验多项式, $\deg(g(x)) = r$, 则

(1) $C = \{p(x) \in R_n \mid p(x)h(x) \equiv 0 \pmod{(x^n - 1)}\}$;

(2) 如果 $h(x) = h_0 + h_1x + \cdots + h_{n-r}x^{n-r}$, 则 C 的校验矩阵为

$$H = \begin{pmatrix} h_{n-r} & \cdots & h_0 & 0 & 0 & \cdots & 0 \\ 0 & h_{n-r} & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & 0 & h_{n-r} & \cdots & h_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & h_{n-r} & \cdots & h_0 \end{pmatrix};$$

(3) 循环码 C 的对偶码 C^\perp 是一个 $[n, r]$ 循环码, 其生成多项式为

$$\begin{aligned} h^\perp(x) &= h_0^{-1}x^{n-r}h(x^{-1}) \\ &= h_0^{-1}(h_0x^{n-r} + h_1x^{n-r-1} + \cdots + h_{n-r}). \end{aligned}$$



证明 (1) 如果 $p(x) \in C$, 则存在 $f(x) \in R_n$, 使得 $p(x) = f(x)g(x)$. 因此,

$$p(x)h(x) = f(x)g(x)h(x) = f(x)(x^n - 1) \equiv 0 \pmod{(x^n - 1)}.$$

另一方面, 如果 $p(x) \in R_n$, 并且 $p(x)h(x) \equiv 0 \pmod{(x^n - 1)}$, 则用 $g(x)$ 去除 $p(x)$ 可得

$$p(x) = q(x)g(x) + r(x),$$

其中 $\deg(r(x)) < \deg(g(x)) = r$. 于是, 我们有

$$\begin{aligned} p(x)h(x) &= q(x)g(x)h(x) + r(x)h(x) = q(x)(x^n - 1) + r(x)h(x), \\ p(x)h(x) &\equiv r(x)h(x) \equiv 0 \pmod{(x^n - 1)}. \end{aligned}$$

由于 $\deg(r(x)h(x)) < n$, 所以一定有 $r(x)h(x) = 0$. 因此, $r(x) = 0$. 于是, $p(x) = q(x)g(x) \in C$.

(2) 设

$$c(x) = c_0 + c_1x + c_2x^2 + \cdots + c_{n-1}x^{n-1} \in C,$$

则 $c(x)h(x) \equiv 0 \pmod{(x^n - 1)}$. 由于 $\deg(c(x)h(x)) < 2n - r$, 所以存在 $q(x) \in R_n$,

$$\deg(q(x)) < n - r$$

使得

$$c(x)h(x) = q(x)(x^n - 1) = q(x)x^n - q(x).$$

由此可知, $C(x)h(x)$ 中 $x^{n-r}, x^{n-r+1}, \dots, x^{n-1}$ 的系数一定为零, 即

$$\begin{aligned} c_0h_{n-r} + c_1h_{n-r-1} + \cdots + c_{n-r}h_0 &= 0, \\ c_1h_{n-r} + c_2h_{n-r-1} + \cdots + c_{n-r+1}h_0 &= 0, \\ &\vdots \\ c_{r-1}h_{n-r} + c_rh_{n-r-1} + \cdots + c_{n-1}h_0 &= 0. \end{aligned}$$

这等价于

$$(c_0c_1c_2 \cdots c_{n-1})H^T = 0$$

设 C' 为以 H 为生成矩阵的线性码, 则 $C' \subseteq C^\perp$. 因为 $h_{n-r} \neq 0$, $\dim(C') = \dim(C^\perp) = r$, 所以 $C' = C^\perp$, 即 H 是 C 的校验矩阵.

(3) 因为 $h(x)g(x) = x^n - 1$, 所以有

$$\begin{aligned} h(x^{-1})g(x^{-1}) &= x^{-n} - 1, \\ x^{n-r}h(x^{-1})x^r g(x^{-1}) &= 1 - x^n. \end{aligned}$$

于是, $h^\perp(x) \mid (x^n - 1)$. 因此 $h^\perp(x)$ 是循环码 $\langle h^\perp(x) \rangle$ 的生成多项式, 其生成矩阵为 H , 因此 $\langle \langle h^\perp(x) \rangle \rangle = C'$.

由上述定理可知, 循环码的对偶码也是循环码. 需要注意的是, 如果 $g(x)$ 和 $h(x)$ 分别是循环码 C 的生成多项式和校验多项式, 则对偶码 C^\perp 的生成多项式不是 $h(x)$, 而是 $h^\perp(x)$. 这是因为

$$g(x)h(x) \equiv 0 \pmod{(x^n - 1)}$$

并不等价于对应的 $V(n, q)$ 中的向量正交.

设 $h(x)$ 是 R_n 中的一个多项式,

$$h(x) = h_0 + h_1x + h_2x^2 + \cdots + h_kx^k, \quad h_k \neq 0.$$

多项式

$$\bar{h}(x) = x^k h(x^{-1}) = h_0x^k + h_1x^{k-1} + \cdots + h_k$$

称为 $h(x)$ 的互反多项式. 定理 9.3.1 中的 $h^\perp(x)$ 就是校验多项式 $h(x)$ 的首 1 互反多项式.

例题 9.3.1 在二元域 F_2 中, 多项式 $x^4 - 1$ 可分解为

$$x^4 - 1 = (x - 1)(x^3 + x^2 + x + 1) = (x - 1)(x + 1)(1 + x^2)$$

循环码 $C = \langle x - 1 \rangle$ 的生成矩阵为

$$G = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

校验多项式

$$h(x) = (x + 1)(x^2 + 1) = x^3 + x^2 + x + 1$$

C 的对偶码 C^\perp 的生成多项式为

$$h^\perp(x) = x^3 h(x^{-1}) = 1 + x + x^2 + x^3.$$

因此, 码 C 的校验矩阵, 即 C^\perp 的生成矩阵为

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$$

于是, $C^\perp = \{0000, 1111, 2222\}$.

下面我们来证明, 二元 Hamming 码和它的对偶码与循环码等价.

定理 9.3.2

二元 Hamming 码 $\text{Ham}(r, 2)$ 等价于循环码.



证明 设 $p(x)$ 是 $F_2[x]$ 中的一个 r 次不可约多项式. 则 $F_2[x]/\langle p(x) \rangle$ 是一个 2^r 阶域. 因此, 存在

本原元 $\alpha \in F_2[x]/\langle p(x) \rangle$, 使得

$$F_2[x]/\langle p(x) \rangle = \{0, 1, \alpha, \alpha^2, \dots, \alpha^{2^r-2}\}.$$

将 $F_2[x]/\langle p(x) \rangle$ 中的每一个多项式 $a_0 + a_1x + \dots + a_{r-1}x^{r-1}$ 看成一个列向量 $\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{r-1} \end{pmatrix}$.

令

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{2^r-2} \end{pmatrix}$$

H 是一个 $r \times (2^r - 1)$ 阶矩阵. 设 C 是以 H 为校验矩阵的线性码. 因为 H 的全部列向量正好是 $V(n, 2)$ 中的所有非零向量, 所以 C 就是 $\text{Ham}(r, 2)$. 令 $n = 2^r - 1$, 我们有

$$\begin{aligned} C &= \{c_0c_1 \cdots c_{n-1} \in V(n, 2) \mid c_0 + c_1\alpha + \cdots c_{n-1}\alpha^{n-1} = 0\} \\ &= \{c(x) \in R_n \mid c(\alpha) \equiv 0 \pmod{p(x)}\}. \end{aligned}$$

如果 $c(x) \in C, r(x) \in R_n$, 则

$$r(\alpha)c(\alpha) \pmod{p(x)} = r(\alpha) \cdot 0 \pmod{p(x)} = 0.$$

从而 $r(x)c(x) \in C$, 因此, C 是循环码. 因为循环码的对偶码也是循环码, 所以二元 Hamming 码的对偶码也与循环码等价.

9.4 循环码的编码方法

设 $C = \langle g(x) \rangle$ 是一个 q 元 $[n, n-r]$ 循环码, $\deg(g(x)) = r, C$ 有 $n-r$ 个信息位, 则它可以对 $V(n-r, q)$ 中的数字信息进行编码.

1. 非系统的编码方法

对任意信息串 $a_0a_1 \cdots a_{n-r-1} \in V(n-r, q)$, 可得到一个信息多项式

$$a(x) = a_0 + a_1x + \cdots + a_{n-r-1}x^{n-r-1},$$

将 $a(x)$ 编成 C 中的码字 $a(x)g(x)$.

2. 系统的编码方法

对 $\forall a_0a_1 \cdots a_{n-r-1} \in V(n-r, q)$, 构造多项式

$$\bar{a}(x) = a_0x^{n-1} + a_1x^{n-2} + \cdots + a_{n-r-1}x^r,$$

显然, $r \leq \deg(\bar{a}(x)) \leq n-1$, 用 $g(x)$ 去除 $\bar{a}(x)$, 得

$$\bar{a}(x) = q(x)g(x) + r(x),$$

其中 $\deg(r(x)) < \deg(g(x)) = r$, 信息串 $a_0a_1 \cdots a_{n-r-1}$ 可编为 C 中的码字

$$c(x) = \bar{a}(x) - r(x) = q(x)g(x).$$

因为 $\bar{a}(x)$ 与 $r(x)$ 中没有相同的项, 所以这种编码是系统编码 (去掉后面 r 个位置后, 恰好为 $V(n-r, q)$ 中的全部向量). 事实上, 如果 $c(x)$ 中 x 的项以降幂排列, 则前 $n-r$ 个是信息位,

后 r 位是校验位.

例题 9.4.1 设 C 是一个二元 $[7, 4]$ 循环码, 它的生成多项式为 $g(x) = x^3 + x + 1$, 因为

$$h(x) = \frac{x^7 - 1}{x^3 + x + 1} = x^4 + x^2 + x + 1,$$

$$h^\perp(x) = x^4 h(x^{-1}) = 1 + x^2 + x^3 + x^4,$$

所以码 C 的校验矩阵为 $H = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$. 这与二元 Hamming 码 $\text{Ham}(3, 2)$ 的

校验矩阵相同, 因此, C 就是 $\text{Ham}(3, 2)$. C 可以纠正一个错误. 设信息串为 0101, 用非系统编码方法可得:

$$a(x) = x + x^3$$

对 $a(x)$ 编码所对应的码字为

$$c(x) = a(x)g(x) = (x + x^3)(x^3 + x + 1) = x^6 + x^3 + x^2 + x$$

用系统编码方法对信息串 0101 进行编码, 有

$$\bar{a}(x) = x^5 + x^3$$

$$x^5 + x^3 = x^2(x^3 + x + 1) + x^2$$

所以 0101 被编为码字

$$c(x) = \bar{a} - r(x) = x^5 + x^3 + x^2.$$

9.5 循环码的检错性能

循环码作为检错码时, 可以检查出成区间的错误.

定义 9.5.1

设 $e(x) = e_0 + e_1x + \cdots + e_{n-1}x^{n-1} \in R_n$, 如果 $e(x)$ 的系数有连续 b 个, 如第 $m+1, m+2, \cdots, m+b$ 个, 使得当 $j < m+1$ 或 $j > m+b$ 时, $e_j = 0$ 但 $e_{m+1} \neq 0, e_{m+b} \neq 0$, 则称 $e(x)$ 为一个长为 b 的成区间差错模式.



设在信道发送端发送的码字为 $c(x)$, 在信道接收端接收到的字为 $u(x)$. 如果

$$e(x) = u(x) - c(x)$$

是一个长为 b 的成区间的差错模式, 则称 $c(x)$ 在传输的过程中出现了一个长为 b 的成区间的错误.

定理 9.5.1

设 C 是一个 q 元 $[n, n-r]$ 循环码, 则 C 可以检查出任意一个长为 $b \leq r$ 的成区间错误.



证明 设

$$c(x) = c_0 + c_1x + \cdots + c_{n-1}x^{n-1} \in C \subseteq R_n$$

是发送的码字, 在传送的过程中出现了一个长为 $b \leq r$ 的成区间错误, 在信道接收端接收到的字为 $u(x)$, 则

$$e(x) = u(x) - c(x) = e_0 + e_1x + \cdots + e_{n-1}x^{n-1}$$

中存在连续 b 个系数, 设为第 $m+1, m+2, \cdots, m+b$ 个, 使得当 $j < m+1$ 或 $j > m+b$ 时, $e_j = 0$, 而 $e_{m+b} \neq 0$. 于是

$$\begin{aligned} e(x) &= e_0 + e_1x + \cdots + e_{n-1}x^{n-1} \\ &= x^{m+1} (e_{m+1} + e_{m+2}x + \cdots + e_{m+b}x^{b-1}). \end{aligned}$$

令

$$a(x) = e_{m+1} + e_{m+2}x + \cdots + e_{m+b}x^{b-1},$$

则

$$\deg(a(x)) = b-1 \leq r-1.$$

设 $g(x)$ 是循环码 C 的生成多项式, $\deg(g(x)) = r$, 则显然 $g(x)$ 不能整除 $a(x)$. 因为 $g(x)$ 的零次项的系数不为零. 所以 $g(x)$ 不能整除 $e(x)$. 否则, 假设 $g(x) \mid e(x)$, 即 $g(x) \mid x^{m+1}a(x)$, 因为 $g(x)$ 与 x^{m+1} 互素, 所以 $g(x) \mid a(x)$, 导致矛盾. 又因为 $c(x)$ 是码字, 所以 $g(x) \mid c(x)$. 因此, 我们有 $g(x)$ 不能整除 $c(x) + e(x)$. 这说明 $c(x) + e(x)$ 不是码字, 即在信道接收端接收到的字 $u(x) = c(x) + e(x)$ 不是码字. 因此, 循环码 C 可以检查出任意长为 $b \leq r$ 的成区间错误.

9.6 习题课

9.6.1 基本概念

循环码的定义, 循环码的生成多项式, 循环码的校验多项式, 循环码的非系统编码方法与系统编码方法, 成区间的差错模式

9.6.2 基本性质与结论

1. 一个码 $L \subseteq R_n$ 是循环码的充分必要条件为 L 满足下列两个条件.

(1) 如果 $a(x), b(x) \in L$, 则 $a(x) + b(x) \in L$.

(2) 如果 $a(x) \in L, r(x) \in R_n$, 则 $r(x)a(x) \in L$.

2. 一个码 $L \subseteq R_n$ 是循环码的充分必要条件为 L 是 R_n 的理想.

3. 定理9.2.1


4. R_n 中首 1 多项式 $p(x)$ 是循环码 $C = \langle p(x) \rangle$ 的生成多项式的充要条件为 $p(x) \mid x^n - 1$.

5. 定理9.3.1

6. 二元 Hamming 码 $\text{Ham}(r, 2)$ 等价于循环码.

7. 设 C 是一个 q 元 $[n, n-r]$ 循环码, 则 C 可以检查出任意一个长为 $b \leq r$ 的成区间错误.

9.6.3 课后习题

 **练习 9.6.1** 设 p 是一个素数.

(1) 在 F_p 上将 $x^p - 1$ 分解成不可约多项式的乘积.

(2) 在 F_p 上将 $x^{p-1} - 1$ 分解成不可约多项式的乘积.

解: (1) $x^p - 1$ 在 \mathbb{F}_p 上的分解: 在有限域 \mathbb{F}_p 中, 由于域的特征是 p , 可以应用二项式定理, 得到:


$$x^p - 1 = (x - 1)^p$$

这是因为在 \mathbb{F}_p 中, 除了 -1 和 1 的项外, 二项式展开中所有的组合数 $\binom{p}{k}$ 乘以 $x^{p-k}(-1)^k$ (其中 $1 < k < p-1$) 都会被 p 整除, 从而在 \mathbb{F}_p 中为 0 . 这意味着 $x^p - 1$ 与 $(x - 1)^p$ 等价.

(2) $x^{p-1} - 1$ 在 \mathbb{F}_p 上的分解: 对于有限域 \mathbb{F}_p 中的任何非零元素 a , 根据费马小定理, 我们知道 $a^{p-1} = 1$. 这表明 $x^{p-1} - 1$ 的每个非零元素 a 都是一个根, 因此:

$$x^{p-1} - 1 = (x - 1)(x - 2) \cdots (x - (p - 1))$$


这里, 我们有 $p-1$ 个根, 每个根对应于 \mathbb{F}_p 中的非零元素, 从而给出了 $x^{p-1} - 1$ 的完全分解.

 **练习 9.6.2** 在 F_3 上将 $x^4 - 1$ 分解成不可约多项式的乘积, 确定所有码长为 4 的三元循环码, 并写出每一个码的生成矩阵和校验矩阵.

解： 长为 4 的三元循环码的生成多项式, 生成矩阵和校验矩阵如下:

特征多项式	特征矩阵	校验多项式	校验矩阵
1	I_4	$x^4 - 1 = 0$	$\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$
$x - 1$	$\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}$	$x^3 + x^2 + x + 1$	$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$
$x + 1$	$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$	$x^3 - x^2 + x - 1$	$\begin{pmatrix} 1 & -1 & 1 & -1 \end{pmatrix}$
$x^2 + 1$	$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$	$x^2 - 1$	$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$
$x^2 - 1$	$\begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$	$x^2 + 1$	$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$
$x^3 - x^2 + x - 1$	$\begin{pmatrix} -1 & 1 & -1 & 1 \end{pmatrix}$	$x + 1$	$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$
$x^3 + x^2 + x + 1$	$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$	$x - 1$	$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$
$x^4 - 1 = 0$	$\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$	1	I_4


每个生成多项式所对应的具体码字的形式见例题 9.2.3.

 **练习 9.6.3** 在 F_2 上将 $x^5 - 1$ 分解成不可约多项式的乘积, 确定所有码长为 5 的二元循环码, 并写出每个码的生成矩阵和校验矩阵.

解： 在 F_2 上, $x^5 - 1 = (x + 1)(x^4 + x^3 + x^2 + x + 1)$ 所有码长为 5 的二元循环码的生成多项

式, 生成矩阵和校验矩阵如下:

生成多项式	生成矩阵	校验矩阵	$V(5, 2)$ 中的码
1	I_5	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	$V(5, 2)$
$x + 1$	$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$	$\left\{ \begin{array}{l} 00000, 11000 \\ 01100, 00110 \\ 00011, 10001 \\ 10010, 01001 \\ 10100, 01010 \\ 00101, 11110 \\ 01111, 10111 \\ 11011, 11101 \end{array} \right\}$
$x^4 + x^3 + x^2 + x + 1$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$	$\{00000, 11111\}$
$x^5 - 1$	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	I_5	$\{00000\}$

 **练习 9.6.4** 证明: 在一个 q 元 $[n, k]$ 循环码中, 任意 k 个连续坐标位置都可构成信息位.

解:

设 C 是一个 q 元 $[n, k]$ 循环码, $g(x)$ 是 C 的生成多项式, 对于任意信息串 $a_0a_1 \cdots a_{k-1} \in V(k, q)$, 构造信息多项式

$$\bar{a}(x) = a_0x^{n-1} + a_1x^{n-2} + \cdots + a_{k-1}x^{n-k}.$$

显然, $n - k \leq \deg(\bar{a}(x)) \leq n - 1$, 用 $g(x)$ 去除 $\bar{a}(x)$, 得


$$\bar{a}(x) = q(x)g(x) + r(x)$$

其中 $\deg(r(x)) < \deg(g(x)) = n - k$, 信息串 $a_0a_1 \cdots a_{k-1}$ 可编为 C 中的码字

$$c(x) = \bar{a}(x) - r(x) = q(x)g(x)$$

因为 $\bar{a}(x)$ 与 $r(x)$ 中没有相同的项, 所以这种编码是系统编码 (去掉后面 $n - k$ 个位置后, 恰好为 $V(k, q)$ 中的全部向量). 事实上, 如果 $c(x)$ 中 x 的项以降幂排列, 则前 k 个是信息位, 后 $n - k$ 位是校验位.

将码 C 中的码字均向右循环移位 t 位, 由于 C 是循环码, 那么移位后的 q^k 个字恰好还是 C 的全部码字. 此时 q^k 个码字中的每个码字从第 t 位起至第 $t + k - 1$ 的长为 k 的字符串恰为 $V(k, q)$ 中的全部向量, 因而从第 t 位起至第 $t + k - 1$ 位连续 k 个坐标位构成信息位.

 **练习 9.6.5** 设 $C_1 = \langle\langle g_1(x) \rangle\rangle$ 和 $C_2 = \langle\langle g_2(x) \rangle\rangle$ 是 R_n 中的循环码. 令

$$C_1 + C_2 = \{c_1 + c_2 \mid c_1 \in C_1, c_2 \in C_2\}$$

则

$$(1) C_1 \subset C_2 \iff g_2(x) \mid g_1(x).$$

$$(2) C_1 \cap C_2 = \langle \langle \text{lcm}(g_1(x), g_2(x)) \rangle \rangle.$$

$$(3) C_1 + C_2 = \langle \langle \text{gcd}(g_1(x), g_2(x)) \rangle \rangle.$$

解:

证明: (1)

\Rightarrow 由于 $C_1 \subset C_2$, 则 $\forall c(x) \in C_1$, 有 $c(x) \in C_2$, 特别地取 $c(x) = g_1(x)$, 则 $g_1(x) \in C_2$, $\exists a(x) \in R_n$, 使得 $g_1(x) = a(x)g_2(x)$, 即 $g_2(x) \mid g_1(x)$.

\Leftarrow 由于 $g_2(x) \mid g_1(x)$, 则 $\exists h(x) \in R_n$, 使得 $g_1(x) = h(x)g_2(x)$. $\forall c(x) \in C_1$, 即 $\exists b(x) \in R_n$, 使得 $c(x) = b(x)g_1(x)$, 从而 $c(x) = b(x)h(x)g_2(x)$. 因此 $c(x) \in C_2$, 即 $C_1 \subset C_2$.

(2) 由于 $C_1 = \langle \langle g_1(x) \rangle \rangle$, $C_2 = \langle \langle g_2(x) \rangle \rangle$, 则 $\forall c(x) \in C_1 \cap C_2$, $\exists a(x), b(x) \in R_n$, 使得 $c(x) = a(x)g_1(x)$ 且 $c(x) = b(x)g_2(x)$, 于是 $\exists d(x) \in R_n$ 使得

$$c(x) = d(x) \text{lcm}(g_1(x), g_2(x)),$$

即 $c(x) \in \langle \langle \text{lcm}(g_1(x), g_2(x)) \rangle \rangle$, 因此

$$C_1 \cap C_2 \subseteq \langle \langle \text{lcm}(g_1(x), g_2(x)) \rangle \rangle$$

反之, 对 $\forall c(x) \in \langle \langle \text{lcm}(g_1(x), g_2(x)) \rangle \rangle$, 则 $\exists t(x) \in R_n$, 使得

$$c(x) = t(x) \text{lcm}(g_1(x), g_2(x))$$

则 $g_1(x) \mid c(x)$, $g_2(x) \mid c(x)$, 于是 $c(x) \in C_1$, $c(x) \in C_2$. 即 $C(x) \in C_1 \cap C_2$. 因此


$$C_1 \cap C_2 \supseteq \langle \langle \text{lcm}(g_1(x), g_2(x)) \rangle \rangle.$$

综上 $C_1 \cap C_2 = \langle \langle \text{lcm}(g_1(x), g_2(x)) \rangle \rangle$.

(3) 设 $\text{gcd}(g_1(x), g_2(x)) = d(x)$, 则 $\exists f_1(x), f_2(x) \in R_n$,

使得 $d(x) = f_1(x)g_1(x) + f_2(x)g_2(x)$ 对 $\forall c(x) \in \langle \langle \text{gcd}(g_1(x), g_2(x)) \rangle \rangle$, 则 $\exists a(x) \in R_n$, 满足 $c(x) = a(x)d(x)$. 即 $c(x) = a(x)f_1(x)g_1(x) + a(x)f_2(x)g_2(x) \triangleq c_1(x) + c_2(x)$. 显然 $c_1(x) \in C_1$, $c_2(x) \in C_2$ 即 $c(x) \in C_1 + C_2$. 故 $C_1 + C_2 \supseteq \langle \langle \text{gcd}(g_1(x), g_2(x)) \rangle \rangle$.

反之, $\forall c(x) \in C_1 + C_2$, 则 $\exists c_1(x), c_2(x) \in R_n$, 使得 $c(x) = c_1(x)g_1(x) + c_2(x)g_2(x)$, 又 $\text{gcd}(g_1(x), g_2(x)) = d(x)$, 则 $\exists a_1(x), a_2(x) \in R_n$, 使得 $g_1(x) = a_1(x)d(x)$, $g_2(x) = a_2(x)d(x)$, 从而 $c(x) = (c_1(x)a_1(x) + c_2(x)a_2(x))d(x)$, 即 $c(x) \in \langle \langle \text{gcd}(g_1(x), g_2(x)) \rangle \rangle$. 故 $C_1 + C_2 \subseteq \langle \langle \text{gcd}(g_1(x), g_2(x)) \rangle \rangle$. 综上有 $C_1 + C_2 = \langle \langle \text{gcd}(g_1(x), g_2(x)) \rangle \rangle$.

 **练习 9.6.6** 设 E_n 是 $V(n, 2)$ 中所有具有偶数重量的向量的集合, C 是一个码长为 n 的二元循环码, 其生成多项式为 $g(x)$, 则

$$(1) E_n = \langle \langle x - 1 \rangle \rangle.$$

$$(2) C = \langle \langle g(x) \rangle \rangle \subset E_n \iff (x - 1) \mid g(x).$$

解: 证明 (1) 首先证明 E_n 是循环码. $\forall x, y \in E_n$, $\omega(x + y) = \omega(x) + \omega(y) - 2\omega(x \cap y)$, 因此, $\omega(x + y)$ 为偶数, $x + y \in E_n$, 即 E_n 是线性码, 若 $\omega(x)$ 为偶数, 则 x 的循环移位的重量不变, 仍然为偶重量, 因此 x 的循环移位仍属于 E_n , 即 E_n 是循环码. $x - 1$ 的重量为 2, 则 $x - 1 \in E_n$, 且 $x - 1$ 是 E_n 中次数最低的首 1 多项式, 因此 $E_n = \langle x - 1 \rangle$.

方法二: 设循环码 $C = \langle x-1 \rangle$, 则 C 的校验多项式为

$$h(x) = \frac{x^n - 1}{x - 1} = x^{n-1} + x^{n-2} + \cdots + x + 1.$$


于是,

$$C^\perp = \langle x^{n-1} + x^{n-2} + \cdots + x + 1 \rangle = \{0, 1\}.$$

显然, $C = (C^\perp)^\perp = \{0, 1\}^\perp = E_n$.

(2) 若 $C = \langle \langle g(x) \rangle \rangle \subset E_n$, 则 $g(x) \in E_n = \langle \langle x-1 \rangle \rangle$, 即存在 $q(x) \in R_n$ 使得 $g(x) = q(x)(x-1)$, 因此 $(x-1) \mid g(x)$. 反之, 若 $(x-1) \mid g(x)$, 则 $g(x) \in \langle \langle x-1 \rangle \rangle$, 从而有

$$C = \langle \langle g(x) \rangle \rangle \subset \langle \langle x-1 \rangle \rangle = E_n.$$

 **练习 9.6.7** 设 $C_1 = \langle \langle x^3 + x + 1 \rangle \rangle$ 是一个二元 $[7, 4]$ 循环码, $C_2 = \langle \langle x^4 + x^3 + x^2 + 1 \rangle \rangle$ 是一个二元 $[7, 3]$ 循环码. 证明: C_1 和 C_2 互为对偶码.

解:

证明: 由已知得, 码 C_1 的生成矩阵为
$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix},$$
 设码 C_2 的校验多项式为

$h(x)$, 则有


$$h(x)(x^4 + x^3 + x^2 + 1) = x^7 - 1,$$

即 $h(x) = x^3 + x^2 + 1$. 显然, 码 C_2 的校验矩阵与码 C_1 生成矩阵相同, 则 C_1 和 C_2 互为对偶码.

方法二: 由于 0 和 1 都不是多项式 $x^3 + x + 1$ 的根, 所以 $x^3 + x + 1$ 一定是二元域 $\text{GF}(2)$ 上的不可约多项式. 因此, $x^3 + x + 1$ 一定是循环码 $C_1 = \langle x^3 + x + 1 \rangle$ 的生成多项式. C_1 的校验多项式为

$$\frac{x^7 - 1}{x^3 + x + 1} = x^4 + x^2 + x + 1.$$

于是, 由定理知, $C_1^\perp = \langle 1 + x^2 + x^3 + x^4 \rangle = C_2$.

 **练习 9.6.8** 在 F_2 上将 $x^7 - 1$ 分解成不可约因式的乘积,

$$x^7 - 1 = (x - 1)(x^3 + x + 1)(x^3 + x^2 + 1)$$

确定所有码长为 7 的循环码, 并且准确描述这些码的特性.

解: 在 F_2 上 $x - 1 = x + 1$, 则所有码长为 7 的二元循环码的生成多项式, 生成矩阵和校验矩阵如下:

生成多项式	生成矩阵	校验矩阵
1	I_7	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
$x+1$	$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$
x^3+x+1	$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$
x^3+x^2+1	$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$
x^4+x^2+x+1	$\begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$
$x^4+x^3+x^2+1$	$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$
$x^6+x^5+x^4+x^3+x^2+x+1$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$
x^7-1	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	I_7

将上述循环码依次记作 C_1, C_2, \dots, C_8 , 从各个循环码的生成矩阵和校验矩阵可以看出, C_1 与 C_8, C_2 与 C_7, C_3 与 C_6, C_4 与 C_5 为四组对偶码.