中国科学: 数学 2021 年 第 51 卷 第 10 期: 1495～1512
SCIENTIA SINICA Mathematica
《中国科学》杂志社
SCIENCE CHINA PRESS
CrossMark
click for updates

综 述

最优局部修复码研究综述

献给冯克勤教授 80 华诞

马立明[1], 邢朝平[2*]

1. 中国科学技术大学数学学院, 合肥 230026;
2. 上海交通大学电子信息与电气工程学院, 上海 200240
E-mail: lmma20@ustc.edu.cn, xingcp@sjtu.edu.cn

收稿日期: 2020-09-28; 接受日期: 2021-01-05; 网络出版日期: 2021-02-09; * 通信作者
国家自然科学基金 (批准号: 12031011) 和华为 - 上海交大联合项目资助项目

摘要 局部修复码或局部恢复码 (locally repairable/recoverable codes, LRC) 是为分布式存储系统和云存储系统的应用而设计的. 与经典分组码相似, 局部修复码也有一个重要的界, 称为 Singleton- 型上界. 在本文中, 达到 Singleton- 型上界的分组码称为最优局部修复码. 类似于经典的极大距离可分码 (maximum distance separable codes, MDS 码), 最优局部修复码具有一些非常好的组合结构. 自从引入局部修复码的 Singleton- 型上界后, 人们在构造最优局部修复码方面做了很多有意义的工作. 本文将综述性地介绍目前国际上关于最优局部修复码的若干界和构造方面的工作.

关键词 局部修复码 函数域 线性码 校验矩阵

MSC (2020) 主题分类 94B05, 68P20, 11T71, 14G50, 94B65

1 引言

大规模云存储和分布式文件系统, 如 Amazon EBS (elastic block store) 和 GoogleFS (Google file system), 在达到如此大的规模后, 磁盘故障是常态, 而不是例外. 在这些系统中, 为了保护数据免受磁盘故障的影响, 最简单的解决方案是跨不同磁盘直接复制数据包. 然而, 这种解决方案的存储开销很大. 作为替代解决方案, $[n, k]$ MDS 码 (maximum distance separable codes) 被用作存储码, 其将 k 个信息字符编码成 n 个字符, 并将它们存储在 n 个磁盘上. 与复制相比, 使用 MDS 码可以显著优化冗余度. 然而, 对于 MDS 码, 当一个节点发生故障时, 系统需要访问 k 个幸存的节点字符作为代价来恢复它, 从而使修复过程复杂化. 为了有效地提高修复效率, 文献 [1] 引入了局部修复码来减少在修复故障节点过程中访问的节点数. 更准确地说, 局部修复码确保了一个故障节点可以通过仅访问 $r \ll k$ 个其他节点来恢复 [1].

英文引用格式: Ma L M, Xing C P. A survey on optimal locally repairable codes (in Chinese). Sci Sin Math, 2021, 51: 1495–1512, doi: 10.1360/SSM-2020-0293

© 2021《中国科学》杂志社

自从 2007 年引入局部修复码以来, 这类码在过去十几年中被许多研究者广泛研究 (参见文献 [1–8]). 通俗地讲, 一个局部修复码只是附加了局部修复性的分组码. 由于大多数经典分组码不具有良好的局部修复性, 人们不得不研究具有良好局部修复性的分组码.

具有良好局部修复性的分组码最初在文献 [1,9] 中被研究. 文献 [1] 考虑了局部修复性的稍微不同的定义, 称为信息局部修复性. 文献 [2,10] 也研究了码的信息局部修复的性质. 对于码长为 n 的具有 k 个信息位和具有局部修复性 r 的局部修复码, 文献 [2] 证明了该码的最小距离 $d(C)$ 满足上界

$$d(C) \leqslant n - k - \left\lceil \frac{k}{r} \right\rceil + 2. \tag{1.1}$$

局部修复码的最小距离的上界 (1.1) 被称为 Singleton- 型上界, 它的证明可由经典 Singleton 界的证明推广得到. 像经典 Singleton 界一样, 局部修复码的 Singleton- 型上界 (1.1) 也没有考虑字母表 q 的大小. 为了补充这个结果, 最近的文献 [11] 建立了依赖于 q 的局部修复码最小距离的界, 对部分情形改进了已有的界. 达到上述任何一个界的局部修复码都称为最优的. 本文特指达到界 (1.1) 的分组码为最优局部修复码.

1.1 早先的结果

显然, 构造最优局部修复码, 即达到界 (1.1) 的分组码, 具有非常重要的理论意义和实际价值. 然而, 由于这个问题的困难性, 目前已知的最优局部可修复码只有少数几类. 早先构造的一类码—金字塔码被证明是最优的 [1]. Silberstein 等 [12] 提出了一种基于 Gabidulin 码与单个奇偶校验码 $(r+1, r)$ 相结合的两层结构. 文献 [13] 利用两层 MDS 码、Reed-Solomon 码和特殊的 $(r+1, r)$-MDS 码, 给出了另一种构造. 这些构造的一个共同缺点是, 字母表的大小均为码长的指数函数, 从而难以实现. 文献 [14] 给出了一类码长与字母表大小相当的最优局部修复码的构造. 然而, 文献 [14] 中构造的局部修复码的码长 n 只能为特殊取值 $n = \lceil \frac{k}{r} \rceil (r+1)$, 即该码的信息率接近于 1. 文献 [14,15] 也给出了一些对局部修复性 r 限制很少的最优局部修复码的存在性结果, 但这两个结果的字母表大小均为码长的指数函数.

第一个突破性的构造在文献 [15] 中给出. 这个构造自然地推广了 Reed-Solomon 码的构造, 因而, 码长依赖于字母表的大小. 这个构造背后的想法非常好. 唯一的缺点是, 对局部修复性 r 的限制, 也就是说 $r+1$ 必须是 $q-1$ 或 q 的因子, 或 $r+1$ 等于 $q-1$ 的因子与 q 的因子的乘积. 文献 [15] 的缺点是, 对局部修复性 r 的选择有一定的限制, 码长以字母表大小 q 为上界. Jin 等 [16] 利用有理函数域的自同构群推广了文献 [15] 中的构造, 得到了具有更灵活的局部修复性以及码长可以达到 $q+1$ 的最优局部修复码.

基于经典的 MDS 猜想, 人们很自然会好奇 q 元最优局部修复码的长度是否能超过 $q+1$. 令人惊讶的是, 文献 [17] 给出了码长超过 $q+1$ 的 q 元最优局部修复码. 虽然文献 [17] 中只产生了少数几个长度大于 $q+1$ 的 q 元最优局部修复码, 但它为人们继续寻找这类码铺平了道路. 文献 [18] 利用椭圆曲线构造了长度达到或接近 $q+2\sqrt{q}$ 的最优局部修复码, 并且构造出的最优局部修复码对最小距离没有限制. 文献 [19] 给出了具有小的最小距离的长度为 $n \approx \frac{r+1}{r} q$ 的最优局部修复码的构造.

鉴于文献 [17,18] 中的结果, 一个自然的问题是, 一个 q 元最优局部修复码的码长最大能有多长. 文献 [20,21] 回答了这个问题. 确切地说, 当最小距离 $d = 3, 4$ 时, 可以利用循环码构造任意长度的最优局部修复码 [21]; 文献 [20] 对于 $d \geqslant 5$ 的情形, 证明了具有局部修复性 r、距离为 d 的 q 元最优局

部修复码的长度由 $O_r(dq^3)$ 界定. 此外, 文献 [20] 还构造了码长为 q 的超线性函数的 q 元最优局部修复码.

局部修复码最初的定义是针对发生一次删除 (即一个节点失败) 的情形. 在过去的几年里, 人们对局部修复提出了各种各样新的定义. 其中一种推广是纠正多个删除错误的码 [22–25], 特别地, 文献 [22] 将文献 [20] 的结果从一个删除推广到多个删除的情形.

1.2 论文的结构

第 2 节介绍一些相关的基本概念和必要的基础知识, 如局部修复码的定义、有限域上的函数域、自同构群、函数域的分歧理论、循环码、代数几何码和二元常重码等. 第 3 节介绍文献 [15] 中利用 Reed-Solomon 码和好的多项式给出的最优局部修复码的构造. 第 4 节介绍文献 [16,18] 中通过有理函数域和椭圆曲线的自同构群构造最优局部修复码的方法. 第 5 节介绍文献 [20,21] 中给出的最优局部修复码的码长上界. 最后一节介绍文献 [20,26,27] 中关于码长为 q 的超线性函数的 q 元最优局部修复码的构造性工作.

2 预备知识

本节介绍局部修复码、有限域上的函数域、有理函数域和椭圆曲线的自同构群、函数域的分歧理论、循环码、代数几何码和二元常重码中的相关概念和理论.

2.1 局部修复码

令 $\mathbb{F}_q$ 是由 q 个元素组成的有限域. 用 $[n]$ 表示集合 $\{1,2,\ldots,n\}$. 通俗来讲, 如果一个给定码字的每个坐标可以通过访问这个码字的其他最多 r 个坐标来恢复, 则它是一个具有局部修复性 r 的分组码. 在各文献中, 局部修复码有几种等价的定义. 局部修复码的正式定义可以按以下方式给出.

定义 2.1 设 $C\subseteq\mathbb{F}_q^n$ 是长度为 n 的 q 元分组码, 对于每个 $\alpha\in\mathbb{F}_q$ 和 $i\in[n]$, 定义

$$C(i,\alpha):=\{\mathbf{c}=(c_1,\ldots,c_n)\in C: c_i=\alpha\}.$$

对于子集 $I\subseteq[n]\setminus\{i\}$, 我们用 $C_I(i,\alpha)$ 表示 $C(i,\alpha)$ 在 I 上的投影. 如果对于每个 $i\in[n]$, 存在子集 $I_i\subseteq[n]\setminus\{i\}$ 满足 $|I_i|\leqslant r$, 使得对任意 $\alpha\neq\beta$, 码 $C_{I_i}(i,\alpha)$ 和 $C_{I_i}(i,\beta)$ 都不相交, 则称 C 为具有局部修复性 r 的局部修复码.

本文只考虑线性局部修复码. 因此, 码长为 n、维数为 k、最小距离为 d 和局部修复性为 r 的 q 元局部修复码可被记为具有局部修复性 r 的 q 元 $[n,k,d]$ 线性码.

2.2 有限域上的函数域

域 $F/\mathbb{F}_q$ 称为 $\mathbb{F}_q$ 上的一元代数函数域, 是指存在 $\mathbb{F}_q$ 上的超越元 x 使得 $F/\mathbb{F}_q(x)$ 为有限代数扩张. 函数域 F 上所有的位组成的集合记为 $\mathbb{P}_F$. 函数域 $F/\mathbb{F}_q$ 上的一个除子 D 可表示为形式和

$$D=\sum_{P\in\mathbb{P}_F}m_PP,$$

其中 m_P 为整数且至多有限多个系数不为 0. 除子 D 的次数为

$$\deg(D) = \sum_{P \in \mathbb{P}_F} m_P \deg(P).$$

令 ν_P 是 F 关于 P 的标准离散赋值, 则 x 的主除子可由下式给出:

$$\mathrm{div}(x) = \sum_{P \in \mathbb{P}_F} \nu_P(x) P = (x)_0 - (x)_\infty,$$

其中 $(x)_0$ 和 $(x)_\infty$ 分别是 x 的零点除子和极点除子. 主除子 $\mathrm{div}(x)$ 的次数等于 0.

对于 F 的任意除子 D, 与之相关联的 Riemann-Roch 空间可由下式定义:

$$\mathcal{L}(D) = \{x \in F^* \mid \mathrm{div}(x) + D \geqslant 0\} \cup \{0\}.$$

事实上, $\mathcal{L}(D)$ 是 $\mathbb{F}_q$ 上的有限维向量空间, 我们用 $l(D)$ 表示它的维数. 函数域 $F/\mathbb{F}_q$ 的亏格 g 可由下式定义:

$$g = g(F) := \max\{\deg(D) - l(D) + 1 \mid D \in \mathrm{Div}(F)\}.$$

事实上, 亏格是函数域中最重要的不变量之一.

命题 2.1 (Riemann 定理) 若 $F/\mathbb{F}_q$ 是一个亏格为 g 的代数函数域, 则对于 F 的任意除子 D, 有

$$l(D) \geqslant \deg(D) + 1 - g.$$

更进一步, 若 $\deg(D) \geqslant 2g - 1$, 则有

$$l(D) = \deg(D) + 1 - g.$$

2.3 自同构群

本小节给出有理函数域和椭圆曲线的自同构群的相关结论. 令 F 是有理函数域 $\mathbb{F}_q(x)$, 其中 x 在 $\mathbb{F}_q$ 上超越. 有理函数域 $\mathbb{F}_q(x)$ 恰好有 $q+1$ 个有理位, 分别为对应 $x-\alpha$ (其中 $\alpha \in \mathbb{F}_q$) 的有限点 $P_{x-\alpha}$ 和对应 $1/x$ 的无穷远点 P_∞. 我们用 $\mathrm{Aut}(F/\mathbb{F}_q)$ 表示由 F 的 $\mathbb{F}_q$- 自同构组成的群, 即

$$\mathrm{Aut}(F/\mathbb{F}_q) = \{\sigma : F \to F \mid \sigma \text{ 是 } F \text{ 的 } \mathbb{F}_q\text{- 自同构}\}.$$

显然, F 的任意自同构 $\sigma \in \mathrm{Aut}(F/\mathbb{F}_q)$ 都可由 $\sigma(x)$ 唯一确定, 并可由下式给出:

$$\sigma(x) = \frac{ax+b}{cx+d},$$

其中 $a, b, c, d \in \mathbb{F}_q$ 满足 $ad - bc \neq 0$. 事实上, 自同构群 $\mathrm{Aut}(F/\mathbb{F}_q)$ 与射影一般线性群 $\mathrm{PGL}_2(q)$ 同构 (参见文献 [28]). 因此, 我们可以用 $\mathrm{PGL}_2(q)$ 来表示 $\mathrm{Aut}(F/\mathbb{F}_q)$.

射影一般线性群 $\mathrm{PGL}_2(q)$ 的所有子群结构都在文献 [16] 中具体列出. 仿射线性群是 $\mathrm{PGL}_2(q)$ 的最重要的一类极大子群, 其定义如下:

$$\mathrm{AGL}_2(q) := \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : \ a \in \mathbb{F}_q^*, b \in \mathbb{F}_q \right\}.$$

每个

$$A = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \in \mathrm{AGL}_2(q)$$

可诱导出函数域 F 的一个自同构 $\sigma(x) = ax + b$.

一条椭圆曲线可视为一个对 $(\mathfrak{E}, O)$, 其中 $\mathfrak{E}$ 由如下的非退化的 Weierstrass 方程定义:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

其中 $a_i \in \mathbb{F}_q$; O 是 x 和 y 的公共极点, 即 $\mathfrak{E}$ 的无穷远点. 令 $\mathrm{Aut}(\mathfrak{E})$ 为固定 $\mathfrak{E}$ 中无穷远点 O 的 $\bar{\mathbb{F}}_q$-自同构群. 下面的结果可以在文献 [29, 定理 III.10.1] 中找到.

引理 2.1 设 $\mathfrak{E}/\mathbb{F}_q$ 是一条椭圆曲线, 则 $\mathrm{Aut}(\mathfrak{E})$ 的阶整除 24. 更准确地, $\mathrm{Aut}(\mathfrak{E})$ 的阶只有以下 5 种情形:

(i) $|\mathrm{Aut}(\mathfrak{E})| = 2$, 如果 $j(\mathfrak{E}) \neq 0, 1728$;

(ii) $|\mathrm{Aut}(\mathfrak{E})| = 4$, 如果 $j(\mathfrak{E}) = 1728$ 且 $\mathrm{char}(\mathbb{F}_q) \neq 2, 3$;

(iii) $|\mathrm{Aut}(\mathfrak{E})| = 6$, 如果 $j(\mathfrak{E}) = 0$ 且 $\mathrm{char}(\mathbb{F}_q) \neq 2, 3$;

(iv) $|\mathrm{Aut}(\mathfrak{E})| = 12$, 如果 $j(\mathfrak{E})$ 等于 0 或 1728 且 $\mathrm{char}\ (\mathbb{F}_q) = 3$;

(v) $|\mathrm{Aut}(\mathfrak{E})| = 24$, 如果 $j(\mathfrak{E})$ 等于 0 或 1728 且 $\mathrm{char}(\mathbb{F}_q) = 2$.

为了构造具有良好参数的局部修复码, 我们需要定义在 $\mathbb{F}_q$ 上的极大椭圆曲线及其 $\mathbb{F}_q$- 自同构群. 令 $\mathrm{Aut}(\mathfrak{E}/\mathbb{F}_q)$ 是 $\mathrm{Aut}(\mathfrak{E})$ 中的所有 $\mathbb{F}_q$- 自同构组成的子群. 文献 [18, 第 3 节] 枚举了很多具有确定的固定无穷远点的自同构群的极大椭圆曲线.

2.4 函数域的分歧理论

设 $E/\mathbb{F}_q$ 是一个具有全常数域 $\mathbb{F}_q$ 的函数域. 设 F 是 E 的一个子域, 具有相同的全常数域 $\mathbb{F}_q$, 并且 E/F 是可分的. 由 Hurwitz 亏格公式 (参见文献 [30, 定理 3.4.13]) 可得出

$$2g(E) - 2 = [E : F](2g(F) - 2) + \deg \mathrm{Diff}(E/F),$$

其中 $\mathrm{Diff}\ (E/F)$ 为 E/F 的差分.

引理 2.2 设 $\mathfrak{E}/\mathbb{F}_q$ 为椭圆曲线, 其函数域记为 $E = \mathbb{F}_q(\mathfrak{E})$. 设 F 是 E 的一个子域, E/F 是有限可分扩张, 并且 E 中存在位 Q 在域扩张 E/F 中具有分歧指数 $e_Q > 1$, 则 F 是有理函数域.

令 $\mathrm{Aut}(E/\mathbb{F}_q)$ 是函数域 E 的 $\mathbb{F}_q$- 自同构群, 即

$$\mathrm{Aut}(E/\mathbb{F}_q) = \{\sigma : E \to E \mid \sigma \text{ 是 } E \text{ 的 } \mathbb{F}_q\text{- 自同构}\}.$$

令 $\mathcal{G}$ 是 $\mathrm{Aut}(E/\mathbb{F}_q)$ 的一个子群. 记

$$E^{\mathcal{G}} = \{z \in E : \sigma(z) = z \text{ 对任意的 } \sigma \in \mathcal{G} \text{ 成立}\}$$

为 E 中对应 $\mathcal{G}$ 的固定子域. 由 Galois 理论可知, $E/E^{\mathcal{G}}$ 为 Galois 扩张且有 $\mathrm{Gal}(E/E^{\mathcal{G}}) = \mathcal{G}$. 函数域扩张 $E/E^{\mathcal{G}}$ 的差分可以由 Dedekind 差分定理 (参见文献 [30, 定理 3.5.1]) 和 Hilbert 差分公式 (参见文献 [30, 定理 3.8.7]) 来确定.

2.5 循环码

线性空间 $\mathbb{F}_q^n$ 的每个线性子空间称为一个码长为 n 的线性码. 码长为 n 和维数为 k 的线性码记为 $[n,k]$ 码. 最小距离为 d 的 $[n,k]$ 码记为 $[n,k,d]$ 码. 下面的结果对确定线性码的最小距离非常有用 (参见文献 [31, 定理 4.5.6]).

命题 2.2 设 C 是 $\mathbb{F}_q$ 上的一个 $[n,k]$ 线性码, H 是 C 的一个校验矩阵, 则 C 的最小距离至少为 d 当且仅当 H 的任意 $d-1$ 列线性无关.

对线性码 C, 若码字 $(c_1,c_2,\ldots,c_{n-1},c_n)\in C$ 成立, 即有 $(c_n,c_1,c_2,\ldots,c_{n-1})\in C$ 成立, 则称 C 为循环码. 长度为 n 的 q 元循环码 C 和主理想环 $\mathbb{F}_q[x]/(x^n-1)$ 的理想一一对应. 设 $C=\langle g(x)\rangle$ 是长度为 n 的 q 元循环码, 其中 $g(x)$ 整除 x^n-1, 则 C 的维数是 $n-\deg g(x)$. 设 $\beta\in\bar{\mathbb{F}}_q$ 是 n 次本原单位根. 如果存在整数 t 和正整数 δ, 使得 $\beta^t,\beta^{t+1},\ldots,\beta^{t+\delta-2}$ 均是 $g(x)$ 的根, 则由文献 [31, 定理 8.1.18] 可知 C 的最小距离至少为 δ.

有限域 $\mathbb{F}_q$ 上的 Reed-Solomon 码可看成是长度为 $q-1$ 的 $\mathbb{F}_q$ 上的 BCH (Bose-Chaudhuri-Hocquenghen) 码, 其生成多项式为

$$g(x)=(x-\alpha^a)(x-\alpha^{a+1})\cdots(x-a^{a+\delta-2}),$$

其中 α 是 $\mathbb{F}_q$ 的本原元, $a\geqslant 1$ 且 $2\leqslant\delta\leqslant q-1$. 当 $a=1$ 时, 生成多项式为

$$g(x)=(x-\alpha)(x-\alpha^2)\cdots(x-a^{\delta-1})$$

的狭义 BCH 码可表示成如下形式:

$$\{(f(1),f(\alpha),f(\alpha^2),\ldots,f(\alpha^{q-2})):f(x)\in\mathbb{F}_q[x]\ \text{满足}\ \deg(f(x))<q-\delta\}.$$

事实上, 这是一个参数为 $[q-1,q-\delta,\delta]$ 的极大距离可分码 (MDS 码).

2.6 代数几何码

代数几何码是 Reed-Solomon 码的自然推广. 关于代数几何码的构造, 可参见文献 [32–35] 了解更多细节. 设 $F/\mathbb{F}_q$ 是一个具有全常数域 $\mathbb{F}_q$ 的函数域. 设 G 是 F 上的一个除子, Riemann-Roch 空间

$$\mathcal{L}(G)=\{z\in F^*:(z)\geqslant -G\}\cup\{0\}$$

是 $\mathbb{F}_q$ 上的有限维向量空间, 由 Riemann 定理可知它的维数至少是 $\deg(G)-g(F)+1$. 令 $\mathcal{P}=\{P_1,\ldots,P_n\}$ 是 F 的 n 个不同有理位的集合, G 是 F 上的除子, 满足

$$0<\deg(G)<n\quad\text{且}\quad \mathrm{supp}(G)\cap\mathcal{P}=\emptyset,$$

则与 $\mathcal{P}$ 和 G 关联的代数几何码可定义如下:

$$C(\mathcal{P},G):=\{(f(P_1),\ldots,f(P_n)):f\in\mathcal{L}(G)\}.$$

显然, $C(\mathcal{P},G)$ 是一个 $[n,k,d]$ 线性码, 其中维数 $k=\ell(G)$, 最小距离 $d\geqslant n-\deg(G)$. 设 V 是 $\mathcal{L}(G)$ 的子空间, 我们可定义 $C(\mathcal{P},G)$ 的子码

$$C(\mathcal{P},V):=\{(f(P_1),\ldots,f(P_n)):f\in V\},$$

则 $C(\mathcal{P},V)$ 的维数是 $\mathbb{F}_q$ 上向量空间 V 的维数, 且 $C(\mathcal{P},V)$ 的最小距离仍然以 $n-\deg(G)$ 为下界.

2.7 二元常重码

码长为 n 的二元常重码是 $\mathbb{F}_2^n$ 的子集且每个码字具有相同的 Hamming 重量. 令 $(n, M, d; w)$ 表示码长为 n、码字个数为 M、最小距离为 d 和重量为 w 的二元常重码. 记 $A(n, d, w)$ 是二元 $(n, M, d; w)$ 常重码的最大码字个数.

一般来说, 确定 $A(n, d, w)$ 的具体值是非常有挑战性的. 事实上, 在构造最优局部修复码时, 我们只需要知道 $A(n, 2r, r+1)$ 的值.

引理 2.3 设 ℓ 是素数方幂, t 是正整数, 则有

$$A(\ell^t, 2\ell-2, \ell) = \frac{\ell^{t-1}(\ell^t-1)}{\ell-1}.$$

满足上式的二元常重码可以通过基于射影几何的 Steiner 系统明确地构造出来 (参见文献 [36,37]).

引理 2.4 [38] 对于素数方幂 q 和整数 $r \geqslant 3$, 有

$$A(q, 2r, r+1) \geqslant \frac{\binom{q}{r+1}}{q^{r-1}-1}.$$

3 通过好的多项式构造

通过多项式和 Reed-Solomon 码构造局部修复码的想法来自文献 [15]. 在 $\mathbb{F}_q$ 中选择 m 个大小为 $r+1$ 且两两不相交的子集 $A_1, A_2, \ldots, A_m$. 选择一个次数为 $r+1$ 的多项式 $g(x) \in \mathbb{F}_q[x]$, 使得对任意 $1 \leqslant i \leqslant m$, $g(x)$ 在每个 A_i 上取值为常数. 这样的多项式 $g(x)$ 称为 "好的". 对每个 $1 \leqslant i \leqslant m$, 令 $A_i = \{\alpha_{i1}, \ldots, \alpha_{i,r+1}\}$ 并考虑下面式子给出的码:

$$C = \{(f(\alpha_{ij}))_{1\leqslant i\leqslant m, 1\leqslant j\leqslant r+1} : f \in V\}, \tag{3.1}$$

其中

$$V = \left\{ \sum_{i=0}^{r-1} \left(\sum_{j=0}^{t-1} a_{ij} g(x)^j \right) x^i : \ a_{ij} \in \mathbb{F}_q \right\}.$$

很容易看出 $\mathbb{F}_q$ 上向量空间 V 的维数是 rt, 且 C 是 Reed-Solomon 码的子码. 文献 [15, 定理 3.1] 给出了以下结论.

引理 3.1 上述 (3.1) 中定义的码 C 是一个最优的具有局部修复性 r 的 $[n, k, d]$ 线性码, 其中

$$n = m(r+1), \quad k = rt \quad \text{且} \quad d = n-(r+1)t+2.$$

证明 对任意多项式 $f \in V$, 由定义可知 f 的次数不超过 $(t-1)(r+1)+(r-1)$. 故码字

$$(f(\alpha_{ij}))_{1\leqslant i\leqslant m, 1\leqslant j\leqslant r+1}$$

中的非零分量至少有 $n-(t-1)(r+1)-(r-1)$ 个. 因此, 码 C 的最小距离 $d = n-(r+1)t+2$ 恰好达到 Singleton- 型上界.

下面只需证明 C 具有局部修复性 r. 假设在 $\alpha_{i,j}$ 处的值 $f(\alpha_{i,j})$ 被删除. 由于 $g(x)$ 在 A_i 上的取值是常数, 因此多项式 $\sum_{i=0}^{r-1}(\sum_{j=0}^{t-1} a_{ij} g(x)^j) x^i$ 限制在 A_i 上可以看成是次数不超过 $r-1$ 的多项式. 利用 Lagrange 插值公式, 多项式的系数可由 A_i 的其他 r 个位置上的取值确定. 因此可以恢复在 $\alpha_{i,j}$ 处的值 $f(\alpha_{i,j})$. □

设 H 是乘法群 $\mathbb{F}_q^*$ 或加法群 $\mathbb{F}_q$ 的子群, 文献 [15, 性质 3.2] 证明了多项式 $g(x)=\prod_{h\in H}(x-h)$ 在 H 的每个陪集上取值均为常数. 因此, 由引理 3.1 可知, 存在一个具有局部修复性 $r=|H|-1$、码长 $n\leqslant q-1$ (考虑乘法群) 或码长 $n\leqslant q$ (考虑加法群) 的最优局部可修复码. 若将 $\mathbb{F}_q$ 的加法结构和乘法结构相结合对 $\mathbb{F}_q$ 进行子集划分得到 $\{A_i\}_{i=1}^m$, 并构造得到好的多项式 $g(x)$, 则有如下结论 (参见文献 [15, 定理 3.3]).

定理 3.1 令 p 是由 $q=p^s$ 个元素组成的有限域 $\mathbb{F}_q$ 的特征. 令 l 和 u 为整数, 满足 $l\mid s$ 和 $p^l\equiv 1 \pmod u$. 设 H 是 $\mathbb{F}_q$ 的一个加法子群, 且对 $\mathbb{F}_{p^l}$ 的乘法运算封闭, 设 $\alpha_1,\alpha_2,\ldots,\alpha_u$ 是 $\mathbb{F}_{p^s}$ 中的 u 次单位根, 则对任意 $b\in\mathbb{F}_{p^s}$, 多项式

$$g(x)=\prod_{i=1}^{u}\prod_{h\in H}(x+h+\alpha_i)$$

在 H 的陪集之并 $\bigcup_{1\leqslant i\leqslant u}H+b\alpha_i$ 上取值是常数. 特别地, 假设 v $(\leqslant s)$ 是 l 的倍数, 取 $H=\mathbb{F}_{p^v}$, 即 $u\mid \gcd(p^v-1,q-1)$, 则存在一个最优的具有局部修复性 $r=up^v-1$ 的 $[n,k,d]$ 线性码, 其中码长 $n\leqslant q-up^v$ 且 $up^v\mid n$.

4 通过函数域的自同构群构造

在第 3 节中给出的构造中, 首先, 局部修复性 r 是有限制的, 这是因为集合 A_i 的大小不能任意选取; 其次, 构造出的最优局部修复码的码长不能超过 q. 上述局部修复码的构造可以很自然地从多项式推广到有限域上的代数曲线 (参见文献 [8,17]). 接下来介绍文献 [16,18] 中利用有理函数域和椭圆曲线来构造最优局部可修复码的工作.

4.1 有理函数域

本小节通过有理函数域的自同构群构造最优局部修复码. 事实证明, 第 3 节给出的最优局部修复码的所有构造都可看成本小节中构造的特例. 此外, 我们还得到了一些不能从文献 [15] 中得到的最优局部修复码.

我们的构造方法如下. 令 $F/\mathbb{F}_q$ 为有理函数域. 设 $\mathcal{G}$ 是 $\mathrm{Aut}(F/\mathbb{F}_q)$ 的 $r+1$ 阶的子群. 从而存在 F 的子域 E, 使得 F/E 是 Galois 扩张, 对应的 Galois 群为 $\mathrm{Gal}(F/E)=\mathcal{G}$. 设 $Q_1,Q_2,\ldots,Q_m$ 是 E 的有理位, 并且它们都在 F/E 中完全分裂. 对任意的 $1\leqslant i\leqslant m$, 令 $P_{i,1},P_{i,2},\ldots,P_{i,r+1}$ 表示 F 中位于 Q_i 之上的 $r+1$ 个有理位. 令 $n=(r+1)m$, 选取 $\mathcal{P}=\{P_{ij}\}_{1\leqslant i\leqslant m,1\leqslant j\leqslant r+1}$. 设整数 t 满足 $1\leqslant t\leqslant m$, 令 $k=rt$, 选取 G 为 F 中的某个 $k+t-2$ 次的除子, z 为 E 中满足 $\deg(z)_\infty^E=1$ 的函数, 以及 x 为 F 中满足 $\deg(x)_\infty=1$ 的函数, 则码

$$\left\{(f(P))_{P\in\mathcal{P}}: f\in\sum_{i=0}^{r-1}\left(\sum_{j=0}^{t-1}a_{ij}z^j\right)x^i; a_{ij}\in\mathbb{F}_q\right\}$$

是一个最优的具有局部修复性 r 的 q 元 $[n,k,d]$ 线性码, 其中 $k=rt$, $d=n-k-\frac{k}{r}+2$. 事实上, 容易看出 $E=\mathbb{F}_q(z)$ 且生成元 z 起着与第 3 节中好的多项式 $g(x)$ 相同的作用.

通过估计扩张 $F/F^{\mathcal{G}}$ 中分歧有理位的个数, 我们可以利用有理函数域的自同构群得到最优局部可修复码的一般构造.

定理 4.1 令 $F/\mathbb{F}_q$ 为有理函数域. 设 $\mathcal{G}$ 是 $\mathrm{Aut}(F/\mathbb{F}_q)$ 的 $r+1$ 阶子群. 若正整数 m 满足 $m \leqslant \lfloor \frac{q+1-2r}{r+1} \rfloor$, 令 $n = m(r+1)$, 则对满足 $1 \leqslant t \leqslant m$ 的整数 t, 存在 q 元最优的具有局部修复性 r 的 $[n,k,d]$ 线性码, 其中

$$k = rt, \quad d = n - k - \frac{k}{r} + 2.$$

文献 [16, 命题 2.1 和 2.2] 给出了 $\mathrm{PGL}_2(q)$ 的完整子群结构. 因此, 通过有理函数域构造的最优局部修复码的局部修复性 r 可以是以下情形之一:

(i) $r+1$ 是 $q-1$ 的因子;

(ii) $r+1$ 是 q 的因子;

(iii) $r+1 = up^v$, 其中 $u \geqslant 2$, $v \geqslant 1$, u 是 $q-1$ 和 p^v-1 的公因子;

(iv) $r+1$ 是 $q+1$ 的因子;

(v) $r+1 = 2h$, 其中 $h \geqslant 2$, 且当 q 是偶数时 $h \mid (q \pm 1)$ 或当 q 是奇数时 $h \mid \frac{q\pm1}{2}$;

(vi) $r+1 = \ell(\ell^2-1)$, 其中 q 是 ℓ 的方幂;

(vii) $r+1 = \ell(\ell^2-1)/2$, 其中 q 为奇数且 q 是 ℓ 的方幂;

(viii) $r+1 = 12$, 当 q 是 2 的偶次方幂或 q 为奇数时;

(ix) $r+1 = 24$, 当 q 为奇数时;

(x) $r+1 = 60$, 当 q 是 2 的偶次方幂或 $q \equiv \pm 1 \pmod{10}$ 时.

由于 $\mathrm{PGL}_2(q)$ 的子群结构更丰富, 从而与文献 [15] 相比, 在给定的有限域 $\mathbb{F}_q$ 上, 局部修复性 r 的选择更加灵活多变. 更进一步, 设 F 中在函数域扩张 $F/F^{\mathcal{G}}$ 分歧的有理位个数为 s, 则最优局部修复码的长度 n 可以达到 $q+1-s$. 因此, 我们可以通过研究 $\mathrm{Aut}(F/\mathbb{F}_q)$ 的子群结构和相应域扩张的分歧性质, 具体地构造码长更大的最优局部修复码.

特别地, 当我们考虑 $\mathrm{PGL}_2(q)$ 的仿射线性子群时, 即

$$\mathrm{AGL}_2(q) \cong \{\sigma \in \mathrm{Aut}(\mathbb{F}_q(x)/\mathbb{F}_q) \mid \sigma(x) = ax+b, a \in \mathbb{F}_q^*, b \in \mathbb{F}_q\},$$

可以证明, 在我们的框架下, 第 3 节给出的构造等价于通过仿射线性群的子群和代数几何码的子码构造 (参见文献 [16, 第 4 节]).

定理 4.2 令 $\mathcal{G}$ 是 $\mathrm{AGL}_2(q)$ 的一个阶为 $up^v = r+1$ 的子群, 其中 u 是 $q-1$ 的因子. 设 $n = m(r+1)$ 为正整数, 其中 m 满足

$$m \leqslant \begin{cases} \dfrac{q-1}{r+1}, & \text{当 } v = 0 \text{ 时}, \\ \dfrac{q}{r+1}, & \text{当 } u = 1 \text{ 时}, \\ \dfrac{q-p^v}{r+1}, & \text{其他}, \end{cases}$$

则对满足 $1 \leqslant t \leqslant m$ 的整数 t, 存在 q 元最优的具有局部修复性 r 的 $[n,k,d]$ 线性码, 其中

$$k = rt, \quad d = n - k - \frac{k}{r} + 2.$$

文献 [16, 引理 5.1] 明确给出了 $\mathrm{PGL}_2(q)$ 的一个 $q+1$ 阶循环子群, 并且文献 [16, 性质 5.2] 证明了该循环扩张中的分歧的有理位个数为 0. 因此, 在文献 [16, 定理 5.3] 中利用改进的代数几何码可以给出长度为 $n = q+1$ 的最优局部修复码的构造.

定理 4.3 令 r 是正整数, 使得 $(r+1) \mid (q+1)$. 对任意正整数 $m \leqslant \frac{q+1}{r+1}$, 令 $n = m(r+1)$, 则对满足 $1 \leqslant t \leqslant m$ 的整数 t, 存在 q 元最优的具有局部修复性 r 的 $[n,k,d]$ 线性码, 其中 $k = rt$, $d = n-k-\frac{k}{r}+2$.

4.2 椭圆曲线

本小节利用椭圆曲线的自同构群构造一族 q 元最优局部修复码[18], 其码长可达到 $q+2\sqrt{q}$. 由引理 2.1 知, 固定椭圆曲线无穷远点的自同构群的阶只能是 2、4、6、12 和 24 中的一个, 因此, 局部修复性 r 必然来自集合 $\{1,2,3,5,7,11,23\}$. 首先, 构造局部修复性 $r=2$ 的最优局部修复码.

定理 4.4 令 $\mathfrak{E}/\mathbb{F}_q$ 为椭圆曲线, 对应函数域为 $E = \mathbb{F}_q(\mathfrak{E})$. 假设 $\mathrm{Aut}(\mathfrak{E}/\mathbb{F}_q)$ 包含一个 3 阶子群 $\mathcal{G}$. 令 F 是 E 中对应 $\mathcal{G}$ 的固定子域, 假设在 F 中有 ℓ 个有理位在 E/F 中完全分裂, 则对于满足 $0 \leqslant t < \ell$ 的整数 t, 存在一个 q 元最优的具有局部修复性 $r=2$ 的

$$[n = 3\ell, k = 2t+1, d = n-3t]$$

线性码.

证明 令 O 是 $\mathfrak{E}$ 的无穷远点, 则 O 在 E/F 中是完全分歧的, 由 Hurwitz 亏格公式或引理 2.2 可得 F 为有理函数域. 设 O' 为 O 在 F 上的局限. 在 F 中选择 z, 使得 $(z)_\infty^F = O'$, 则有

$$(z)_\infty = 3O.$$

选择 $x \in E$, 使得 $(x)_\infty = 2O$. 考虑由下式定义的 $\mathbb{F}_q$- 线性空间 V_t:

$$V_t := \{f_0(z) + f_1(z)x : \text{其中 } f_i(z) \in \mathbb{F}_q[z], i = 0,1; \deg(f_0(z)) \leqslant t;\ \deg(f_1(z)) \leqslant t-1\}.$$

设 $\{P_{i1}, P_{i2}, P_{i3}\}_{i=1}^{\ell}$ 为 E 上 n 个有理位, 使得对于每个 i, P_{i1}、P_{i2} 和 P_{i3} 位于 F 的相同有理位上. 我们可知代数几何码的子码

$$C_t := \{(f(P_{i1}), f(P_{i2}), f(P_{i3}))_{i=1}^{\ell} : f \in V_t\}$$

是最优的具有局部修复性 $r=2$ 的 q 元 $[3\ell, 2t+1, 3\ell-3t]$ 线性码. □

通过使用在文献 [18, 第 3 节] 中构造的极大椭圆曲线, 我们可以得到以下结果 (参见文献 [18, 定理 1]).

定理 4.5 设 $q = p^a$, 其中 p 为素数, $a > 0$ 为偶数. 如果 $p = 3$ 或 $p \equiv 2 \pmod 3$, 则对于满足 $0 \leqslant t < l \leqslant \lfloor \frac{q+2\sqrt{q}}{3} \rfloor$ 的正整数 t, 存在 q 元最优的具有局部修复性 $r=2$ 的

$$[n = 3\ell, k = 2t+1, d = n-3t]$$

线性码.

现在考虑 r 是奇数的情形, 即 $r = 3,5,7,11,23$. 设 $\mathfrak{E}/\mathbb{F}_q$ 是定义在 $\mathbb{F}_q$ 上的椭圆曲线, $E = \mathbb{F}_q(x,y)$ 为对应的函数域 $\mathbb{F}_q(\mathfrak{E})$, 其中 x 和 y 满足 Weierstrass 方程. 设 $\mathcal{G}$ 是 $\mathrm{Aut}(\mathfrak{E}/\mathbb{F}_q)$ 的阶为 $r+1 = 2s$ 的子群, 其中正整数 $s \geqslant 2$, $r < q$. 此外, 假设集合 $\{\sigma(x) : \sigma \in \mathcal{G}\}$ 大小为 s. 设 $F = E^{\mathcal{G}}$ 是 E 关于 $\mathcal{G}$ 的固定子域. 由文献 [18, 性质 20] 可知, 我们有以下两个事实:

(i) 存在函数 $z \in E$ 和 $r+1$ 个不同的有理位 P_i $(1 \leqslant i \leqslant r+1)$, 满足

$$F = \mathbb{F}_q(z) \quad \text{且} \quad (z)_\infty = P_1 + P_2 + \cdots + P_{r+1};$$

存在 E 上的元素 $w_0=1,w_1,\ldots,w_{r-1}$, 它们在 F 上是线性无关的;

(ii) 假设对于某个 $\ell\geqslant 1$, 当 $1\leqslant i\leqslant \ell$ 时, 存在两两互不相同的有理位 $P_{i,1},P_{i,2},\ldots,P_{i,r+1}$ 对应 F 中同一个有理位 Q_i, 使得

$$\{P_{i,1},P_{i,2},\ldots,P_{i,r+1}\}_{i=1}^{l}\cap \mathrm{supp}((z)_\infty)=\emptyset,$$

则对任意 $1\leqslant i\leqslant \ell$, 矩阵

$$M=\begin{pmatrix} 1 & w_1(P_{i,1}) & \cdots & w_{r-1}(P_{i,1}) \\ 1 & w_1(P_{i,2}) & \cdots & w_{r-1}(P_{i,2}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & w_1(P_{i,r+1}) & \cdots & w_{r-1}(P_{i,r+1}) \end{pmatrix}$$

的任意 $r\times r$ 子阵是可逆的.

选取 F 上的除子 $G=(t-1)(z)_\infty$ 满足 $\mathrm{supp}(G)\cap\{Q_1,\ldots,Q_\ell\}=\emptyset$. Riemann-Roch 空间 $\mathcal{L}(G)$ 是 $\mathbb{F}_q$ 上的有限维向量空间, 其维数为 $t=\ell(G)\geqslant \deg(G)-g(F)+1$, 且 $\{1,z,\ldots,z^{t-1}\}$ 是 Riemann-Roch 空间 $\mathcal{L}(G)$ 的一组 $\mathbb{F}_q$- 基. 事实上, 对每个 $0\leqslant i\leqslant r-1$, 选取的元素 w_i 满足条件

$$(w_i)_\infty=P_1+P_2+\cdots+P_{i+1}.$$

从而对所有 $1\leqslant i\leqslant \ell$, $1\leqslant j\leqslant r+1$ 和 $0\leqslant l\leqslant r-1$, 有 $\nu_{P_{i,j}}(w_l)\geqslant 0$. 考虑函数集合

$$V=\left\{\sum_{j=1}^{t}a_{0j}z^{j-1}+\sum_{i=1}^{r-1}\left(\sum_{j=1}^{t-1}a_{ij}z^{j-1}\right)w_i : a_{ij}\in\mathbb{F}_q\right\}.$$

由文献 [18, 性质 19] 可知, 代数几何码的子码

$$C(\mathcal{P},V)=\{(f(P))_{P\in\mathcal{P}} : f\in V\}$$

是一个具有局部恢复性 r 的 q 元 $[n,k,d]$ 线性码, 其中码长 $n=\ell(r+1)$, 维数 $k=rt-(r-1)$, 最小距离

$$d\geqslant n-\deg(G)=n-(t-1)(r+1).$$

由 Hurwitz 亏格公式可知, E 中在函数域扩张 E/F 分歧的有理位最多有 $r+3$ 个 (参见文献 [18, 引理 8]), 剩余的有理位在 E/F 中是完全分裂的. 因此, 对于局部修复性 $r=3,5,7,11,23$ 的情形, 我们可以通过椭圆曲线的自同构群得到以下最优局部修复码的构造 (参见文献 [18, 性质 21]).

命题 4.1 令 $\mathfrak{E}/\mathbb{F}_q$ 是由 Weierstrass 方程定义的含有 N 个有理点的椭圆曲线, 令 E 是其对应的函数域 $\mathbb{F}_q(\mathfrak{E})$. 设 $\mathcal{G}$ 是 $\mathrm{Aut}(\mathfrak{E}/\mathbb{F}_q)$ 的一个阶为 $|\mathcal{G}|=r+1=2s$ 的子群, 其中正整数 $s\geqslant 2$ 且 $r<q$. 更进一步, 假设集合 $\{\sigma(x):\ \sigma\in\mathcal{G}\}$ 大小为 s, 则对满足 $1\leqslant \ell\leqslant \lfloor\frac{N-2r-4}{r+1}\rfloor$ 和 $1\leqslant t\leqslant \ell$ 的正整数 l 和 t, 存在 q 元最优的具有局部恢复性 r 的

$$[n=\ell(r+1),k=rt-r+1,d=n-(t-1)(r+1)]$$

线性码.

通过使用在文献 [18, 第 3 节] 中构造的极大椭圆曲线, 可以得到以下结论 (参见文献 [18, 定理 2]).

定理 4.6 设 $q=p^a$, 其中 p 为素数且 $a>0$ 为偶数. 如果 p 和 r 满足下列条件中任意一条:

(i) $r=3$, $p=2$ 或 $p\equiv 3 \pmod 4$;

(ii) $r=5$, $p=3$ 或 $p\equiv 2 \pmod 3$;

(iii) $r=7$, $p=2$;

(iv) $r=11$, $p=2$ 或 $p=3$;

(v) $r=23$, $p=2$,

则对满足 $1\leqslant t<\ell\leqslant\lfloor\frac{q+2\sqrt{q}-r-2}{r+1}\rfloor$ 的正整数 t 和 ℓ, 存在 q 元最优的具有局部恢复性 r 的

$$[n=(r+1)\ell, k=r(t-1)+1, d=n-(t-1)(r+1)]$$

线性码.

5 码长的上界

在第 3 和 4 节中构造的最优局部修复码的码长至多是字母表大小的线性函数. 一个很自然的问题是, 在给定的有限域 $\mathbb{F}_q$ 上, 最优局部修复码的最大码长能有多大. 事实上, 对于最小距离是 3 和 4 的情形, 已经证明最优局部修复码的码长可以任意地大. 然而, 当最小距离 $d\geqslant 5$ 时, 最优局部可修复码的最大码长是有限制的, 即存在上界.

5.1 最小距离 $d=3,4$ 时码长无界

本小节介绍 Luo 等在文献 [21] 中的结果. 对于最小距离 $d=3,4$ 的情形, 他们给出了码长可以任意大的最优局部可修复码的构造. 他们的方法是通过特定的生成多项式和校验多项式构造出的循环码来实现的.

设 q 为素数方幂, n 为正整数, 满足 $\gcd(n,q)=1$. 设 $r\geqslant 2$, $\gcd(n,q-1)\equiv 0 \pmod{r+1}$. 设 $\beta\in\bar{\mathbb{F}}_q$ 是 n 次本原单位根, 令 $\alpha=\beta^{\frac{n}{r+1}}$, 则有 $\alpha\in\mathbb{F}_q\setminus\{0,1\}$ 且 $\alpha^{q-1}=1$. 设

$$g(x)=(x-1)(x^{\frac{n}{r+1}}-\alpha),$$

则很容易验证由 $g(x)$ 生成的循环码 C 是一个 q 元最优的具有局部修复性 r 的 $[n,k=n-(1+\frac{n}{r+1}),3]$ 线性码.

为了证明 C 具有局部修复性 r, 由文献 [21, 引理 2] 可知, 只需证明 C 的对偶码 $C^{\perp}$ 的 Hamming 距离至多为 $r+1$. 这是因为对偶码 $C^{\perp}$ 是由 $h(x)=(x^n-1)/g(x)$ 生成的循环码, 而 $C^{\perp}$ 中码字 $(x-1)h(x)=\sum_{i=0}^{r}\alpha^{r-i}x^{\frac{in}{r+1}}$ 的 Hamming 重量是 $r+1$. 由 $g(x)$ 的定义可知 1 和 β 是 $g(x)$ 的根, 再由 BCH 界可知 C 的最小距离至少是 3. 因此, $g(x)$ 生成的循环码 C 是满足 Singleton- 型上界的最优局部修复码.

类似地, 考虑最小距离 $d=4$ 的最优局部修复码. 假设 $r\geqslant 3$, $\gcd(\frac{n}{r+1},r+1)$ 整除 2, 则由 Bezout 等式可知, 存在整数 a 和 b, 使得

$$a\cdot\frac{n}{r+1}+b(r+1)=2.$$

令 $\gamma=\alpha^a$, 则可以证明由

$$g(x)=(x-1)(x-\gamma)(x^{\frac{n}{r+1}}-\alpha)$$

生成的循环码是 q 元最优的具有局部修复性 r 的 $[n, k = n - (2 + \frac{n}{r+1}), 4]$ 码. 因此, 可以得到如下结果 (参见文献 [21]).

定理 5.1 令 q 是一个素数方幂. 设 n 是满足 $\gcd(n, q) = 1$ 的正整数, 则

(i) 当 $r \geqslant 2$ 且 $\gcd(n, q-1) \equiv 0 \pmod{r+1}$ 时, 存在 q 元最优的具有局部修复性 r 的 $[n, k = n - (1 + \frac{n}{r+1}), 3]$ 循环码;

(ii) 当 $r \geqslant 3$、$\gcd(n, q-1) \equiv 0 \pmod{r+1}$ 且 $\gcd(\frac{n}{r+1}, r+1)$ 整除 2 时, 存在 q 元最优的具有局部修复性 r 的 $[n, k = n - (2 + \frac{n}{r+1}), 4]$ 循环码.

5.2 最小距离 $d \geqslant 5$ 的码长上界

本小节介绍文献 [20] 中给出的当最小距离 $d \geqslant 5$ 时 q 元域上最优局部修复码的码长上界.

首先, 利用码长 n、最小距离 d 和局部修复性 r 给出存在互不相交恢复集的一个判定准则 (参见文献 [20, 引理 8]).

引理 5.1 设 C 是一个 q 元最优的具有局部修复性 r 的 $[n, k, d]$ 码, 当

$$\frac{n}{r+1} \geqslant \left(d - 2 - \left\lfloor \frac{d-2}{r+1} \right\rfloor\right)(3r+2) + \left\lfloor \frac{d-2}{r+1} \right\rfloor + 1$$

时, C 存在 $m = \frac{n}{r+1}$ 个大小都为 $r+1$ 的互不相交的恢复集.

将 C 的校验矩阵经过一系列初等列变换后删掉前 m 个坐标, 可构造一个 $[n-m, \geqslant k, \geqslant \lfloor \frac{d-1}{2} \rfloor + 1]$ 线性码 C_2. 利用码 C_2 的参数必定满足 Hamming 界, 我们可以得到如下关于码长的上界 (参见文献 [20, 定理 10]).

定理 5.2 设 C 是一个 q 元最优的具有局部修复性 r 的 $[n, k, d]$ 码, 满足 $(r+1) \mid n$、$n = \Omega(dr^2)$ 及引理 5.1 的条件. 若 $d \geqslant 5$ 且 $d \equiv a \pmod 4$, 其中 $1 \leqslant a \leqslant 4$, 则有

$$n \leqslant \begin{cases} \dfrac{r+1}{r} \times \dfrac{d-a}{4(q-1)} \times q^{\frac{4(d-2)}{d-a}}, & \text{若 } a = 1, 2, \\ \dfrac{r+1}{r} \times \left(\dfrac{d-a}{4(q-1)} \times q^{\frac{4(d-3)}{d-a}} + 1\right), & \text{若 } a = 3, 4. \end{cases}$$

特别地, 当 d 不能被 4 整除时, $n \leqslant O(dq^3)$; 当 $4 \mid d$ 时, $n \leqslant O(dq^{3+4/(d-4)})$. 此外, 对于 $d = 5, 6, 7, 8, 9, 10$ 的情形, 分别有 $n = O(q^2), O(q^3), O(q^3), O(q^4), O(q^{5/2}), O(q^3)$.

上述关于最优局部修复码的码长上界考虑的都是固定最小距离的情形. 如果最小距离 d 与码长 n 成比例, 那么码长上界的情形如何呢? 针对这种情形, 利用文献 [11, 定理 1] 和 Plotkin 界, 得到最优局部修复码的最小距离不超过 $\frac{r^2+2r+3}{r} \cdot q$ (参见文献 [20, 定理 13]). 因此, 我们有如下上界.

命题 5.1 当最小距离 $d = O(n)$ 和局部修复性 r 为常数时, 最优局部修复码的码长 n 以 $O(q)$ 为上界.

6 码长的构造性下界

从前面的内容中很容易看出, 当 $d > 4$ 时, 第 4 节前面所有构造出的最优局部修复码的长度都满足 $n \leqslant O(q)$. 本节证明当 $d > 4$ 时, 存在码长达到 $n = \Omega(q^{1+\epsilon})$ 的最优局部修复码, 其中 $\epsilon > 0$ 为常数.

首先, 给出一个最优局部修复码码长的构造性的下界 [20]. 这种最优局部可修复码的校验矩阵 H 由下列算法构造出来. 对于 $i = 1, 2, \ldots, \frac{n}{r+1}$ 和 $j = 1, 2, \ldots, r+1$, 选取 $\boldsymbol{h}_{i,j} \in \mathbb{F}_q^{\frac{n}{r+1}+d-2}$ 作为 H 的列,

使得 $\boldsymbol{h}_{i,j}$ 的前 $\frac{n}{r+1}$ 个分量中的第 i 个分量非零, 且 $\boldsymbol{h}_{i,j}$ 与任意之前选择的最多 $d-2$ 列都线性无关. 更多细节参见文献 [20, 定理 16].

定理 6.1 对满足 $d \leqslant r+2$ 的正整数 r 和 $d>4$, 存在长度 $n=\Omega_{d,r}(q^{1+1/[(d-3)/2]})$ 的最优局部修复码. 特别地, 如果 $r \geqslant 3$ 且 $(r+1) \mid n$, 则存在最小距离为 5 的最优局部修复码使得其码长 $n=O(q^2)$ 达到最佳数量级.

上述关于最优局部修复码的码长下界在文献 [26,27] 中被进一步改进.

6.1 最小距离 $d=5,6$

本小节介绍 Jin[27] 利用二元常重码给出的距离为 5 和 6 的 q 元最优局部修复码的具体结构. 与文献 [20] 相比, Jin[27] 的构造更为明确, 可以得到具有更好下界的最优局部修复码.

设 $r \geqslant 4$ 为整数. 如果存在 $(q,m,2r;r+1)$ 二元常重码 A, 则存在一族 $\mathbb{F}_q$ 的子集 $\{I_i\}_1^m$, 使得对所有 $1 \leqslant i \neq j \leqslant m$, 有 $|I_i|=r+1$ 且 $|I_i \cap I_j| \leqslant 1$. 对任意 $1 \leqslant i \leqslant m$, 记

$$I_i=\{a_{i,1},a_{i,2},\ldots,a_{i,r+1}\},$$

定义矩阵

$$D_i=\begin{pmatrix} a_{i,1} & a_{i,2} & \cdots & a_{i,r+1} \\ a_{i,1}^2 & a_{i,2}^2 & \cdots & a_{i,r+1}^2 \\ a_{i,1}^3 & a_{i,2}^3 & \cdots & a_{i,r+1}^3 \end{pmatrix}.$$

从而, 可定义矩阵

$$H=\left(\begin{array}{c|c|c|c} \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{1} & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \\ \hline D_1 & D_2 & \cdots & D_m \end{array}\right), \tag{6.1}$$

其中 $\mathbf{1}$ 和 $\mathbf{0}$ 分别代表维数为 $r+1$ 的全 1 向量和零向量.

以 H 为校验矩阵的线性码 C 是一个 q 元最优的具有局部修复性 r 的 $[n,k,5]$ 码, 其中码长 $n=m(r+1)$, 维数 $k=n-m-3$ (参见文献 [27, 定理 9]). 事实上, C 具有局部修复性 r, 是因为由校验矩阵 H 的选取可知 C 中任何码字 $(c_1,c_2,\ldots,c_n) \in C$ 满足

$$\sum_{j=1}^{r+1} c_{(i-1)(r+1)+j}=0.$$

同时不难看出, H 的任何 4 列都线性无关 (参见文献 [27, 引理 8]), 因而, C 的最小距离至少为 5. 另一方面, 由 Singleton- 型上界可知, C 的最小距离 $d \leqslant n-k-m+2=5$.

根据上述构造可知, 如果明确给出一个二元常重码, 则可以明确地构造一个最优局部修复码. 通过结合引理 2.3 和 2.4 的二元常重码的界, 可以得到以下结果.

定理 6.2 我们可以构造出满足如下条件的最优局部修复码.

(1) 若 $r+1 \geqslant 5$ 是一个素数方幂, 则对任意 $t \geqslant 1$, 可具体构造出一族 q 元最优的具有局部修复性 r 的 $[n,k,5]$ 线性码, 其中

$$q=(r+1)^t, \quad m=\frac{q(q-1)}{r}, \quad n=m(r+1), \quad k=n-m-3.$$

(2) 若 $r+1 \geqslant 8$ 是 2 的方幂, 则对任意 $t \geqslant 1$, 可具体构造出一族 q 元最优的具有局部修复性 r 的 $[n,k,6]$ 线性码, 其中

$$q=(r+1)^t, \quad m=\frac{q(q-1)}{r}, \quad n=m(r+1), \quad k=n-m-4.$$

(3) 若 $4 \leqslant r \leqslant q-1$ 且 $(r+1) \mid n$, 则可具体构造出一族 q 元最优的具有局部修复性 r 的 $[n,k,5]$ 线性码, 其中

$$n \geqslant \frac{\binom{q}{r+1}}{q^{r-1}-1}, \quad k=n-\frac{n}{r+1}-3.$$

(4) 设 q 是 2 的方幂. 当 $4 \leqslant r \leqslant q-1$ 且 $(r+1) \mid n$ 时, 可具体构造出一族 q 元最优的具有局部修复性 r 的 $[n,k,6]$ 线性码, 其中

$$n \geqslant \frac{\binom{q}{r+1}}{q^{r-1}-1}, \quad k=n-\frac{n}{r+1}-4.$$

6.2 最小距离 $d \geqslant 7$

本小节利用文献 [26] 中给出的具有 Vandermonde 结构的校验矩阵来构造最优局部修复码.

设 $\{A_i\}_1^m$ 是 $\mathbb{F}_q$ 的子集族, 满足对 $1 \leqslant i \leqslant m$, 有 $|A_i|=r+1$. 记 $A_i=\{a_{i,1},a_{i,2},\ldots,a_{i,r+1}\}$. 定义矩阵

$$D_i=\begin{pmatrix} a_{i,1} & a_{i,2} & \cdots & a_{i,r+1} \\ a_{i,1}^2 & a_{i,2}^2 & \cdots & a_{i,r+1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{i,1}^{d-2} & a_{i,2}^{d-2} & \cdots & a_{i,r+1}^{d-2} \end{pmatrix}.$$

更进一步, 定义矩阵

$$H=\left(\begin{array}{c|c|c|c} \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{1} & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \\ \hline D_1 & D_2 & \cdots & D_m \end{array}\right), \tag{6.2}$$

其中 $\mathbf{1}$ 和 $\mathbf{0}$ 分别代表维数为 $r+1$ 的全 1 向量和零向量.

若 $d \geqslant 5$, 则 (6.2) 中定义的 H 中任意 $d-1$ 列线性无关当且仅当 $|\bigcup_{i\in S} A_i| \geqslant r|S|+1$ 对于任意 $S \subseteq [m]$, 满足 $|S| \leqslant \lfloor \frac{d-1}{2} \rfloor$ (参见文献 [26, 定理 3.1]). 因此, 我们立即得到如下结论.

定理 6.3 设 $t=\lfloor\frac{d-1}{2}\rfloor$ 且 $(r+1)\mid n$, 若存在 $m=n/(r+1)$ 个大小均为 $(r+1)$ 的集合 $A_1,A_2,\ldots,A_m\subseteq\mathbb{F}_q$, 使得

$$\left|\bigcup_{i\in S}A_i\right|\geqslant r|S|+1 \tag{6.3}$$

对任意 $S\subseteq[m]$, 满足 $|S|\leqslant t$, 则存在码长为 n、最小距离为 d 且局部修复性为 r 的 q 元最优局部修复码.

为了得到最优的局部修复码, 我们需要找到满足 (6.3) 的集合族, 这相当于构造一个避免出现小回路的 $(r+1)$- 均匀超图.

引理 6.1 存在 m 个满足条件 (6.3) 的集合当且仅当存在一个 $(r+1)$- 均匀超图 $H=([q],E)$ 满足 $|E|=m$, 且对任意 $\ell\leqslant t$, 图 H 中不含 Berge ℓ- 回路.

用 BC_k 表示 k- 回路的集合. 设 $\mathcal{B}_k=\{BC_2,\ldots,BC_k\}$. 用 $ex_{r+1}(n,\mathcal{B}_t)$ 表示不含 $\mathcal{B}_t$ 中任何子图的 $(r+1)$- 均匀超图的最大边数. Verstraete 在文献 [39] 中具体构造了不包含 $\mathcal{B}_3$ (或 $\mathcal{B}_4$) 中任意子图的 $(r+1)$- 均匀超图 $H=([q],E)$, 其中 $|E|=q^{2-o(1)}$ (或 $|E|=q^{\frac{3}{2}-o(1)}$). 此外, Verstraete [1)] 还得到了

$$ex_{r+1}(n,\mathcal{B}_t)=\Omega_{r,t}(n(n\log n)^{\frac{1}{t-1}}).$$

文献 [26, 定理 4.1 和 4.2] 证明了, 对于固定的 t, 可以通过多项式时间的确定性算法来随机构造出满足条件 (6.3) 的子集 $\{A_i\}_{i=1}^m$. 因此, 我们可以将以上所有结论总结如下.

定理 6.4 设 $r\geqslant d-2$ 且 $(r+1)\mid n$.

(1) 当 $d=7$ 或 $d=8$ 时, 可具体构造出码长为 $n=q^{2-o(1)}$、最小距离为 d 的具有局部修复性 r 的最优局部修复码;

(2) 当 $d=9$ 或 $d=10$ 时, 可具体构造出码长为 $n=q^{\frac{3}{2}-o(1)}$、最小距离为 d 的具有局部修复性 r 的最优局部修复码;

(3) 当 $d\geqslant 11$ 时, 存在码长为 $n=\Omega(q(q\log q)^{\frac{1}{\lfloor(d-3)/2\rfloor}})$、最小距离为 d 的具有局部修复性 r 的最优局部修复码;

(4) 当 $d\geqslant 11$ 为常数时, 可具体构造出码长为 $n=\Omega(q^{1+\frac{1}{\lfloor(d-3)/2\rfloor}})$、最小距离为 d 的具有局部修复性 r 的最优局部修复码.

参考文献

1 Huang C, Chen M, Li J. Pyramid codes: Flexible schemes to trade space for access efficiency in reliable data storage systems. In: Proceedings of Sixth IEEE International Symposium on Network Computing and Applications. Cambridge: IEEE, 2007, 79–86

2 Gopalan P, Huang C, Simitci H, et al. On the locality of codeword symbols. IEEE Trans Inform Theory, 2012, 58: 6925–6934

3 Jin L, Kan H, Zhang Y. Constructions of locally repairable codes with multiple recovering sets via rational function fields. IEEE Trans Inform Theory, 2020, 66: 202–209

4 Kamath G, Prakash N, Lalitha V, et al. Codes with local regeneration and erasure correction. IEEE Trans Inform Theory, 2014, 60: 4637–4660

5 Li X, Ma L, Xing C. Construction of asymptotically good locally repairable codes via automorphism groups of function fields. IEEE Trans Inform Theory, 2019, 65: 7087–7094

6 Ma L, Xing C. Constructive asymptotic bounds of locally repairable codes via function fields. IEEE Trans Inform Theory, 2020, 66: 5395–5403

1) Verstraete J. Personal communication. 2018

7 Papailiopoulos D S, Dimakis A G. Locally repairable codes. IEEE Trans Inform Theory, 2014, 60: 5843–5855
8 Barg A, Tamo I, Vlăduţ S. Locally recoverable codes on algebraic curves. IEEE Trans Inform Theory, 2017, 63: 4928–4939
9 Han J, Lastras-Montano L. Reliable memories with subline accesses. In: Proceedings of 2007 IEEE International Symposium on Information Theory (ISIT). Cambridge: IEEE, 2007, 2531–2535
10 Forbes M, Yekhanin S. On the locality of codeword symbols in non-linear codes. Discrete Math, 2014, 324: 78–84
11 Cadambe V, Mazumdar A. Bounds on the size of locally recoverable codes. IEEE Trans Inform Theory, 2015, 61: 5787–5794
12 Silberstein N, Rawat A S, Koyluoglu O O, et al. Optimal locally repairable codes via rank-metric codes. In: Proceedings of 2013 IEEE International Symposium on Information Theory (ISIT). Cambridge: IEEE, 2013, 1819–1823
13 Tamo I, Papailiopoulos D S, Dimakis A G. Optimal locally repairable codes and connections to matroid theory. IEEE Trans Inform Theory, 2016, 62: 6661–6671
14 Prakash N, Kamath G M, Lalitha V, et al. Optimal linear codes with a local-error-correction property. In: Proceedings of 2012 IEEE International Symposium on Information Theory (ISIT). Cambridge: IEEE, 2012, 2776–2780
15 Tamo I, Barg A. A family of optimal locally recoverable codes. IEEE Trans Inform Theory, 2014, 60: 4661–4676
16 Jin L, Ma L, Xing C. Construction of optimal locally repairable codes via automorphism groups of rational function fields. IEEE Trans Inform Theory, 2020, 66: 210–221
17 Barg A, Haymaker K, Howe E, et al. Locally recoverable codes from algebraic curves and surfaces. In: Algebraic Geometry for Coding Theory and Cryptography. New York: Springer, 2017, 95–126
18 Li X, Ma L, Xing C. Optimal locally repairable codes via elliptic curves. IEEE Trans Inform Theory, 2019, 65: 108–117
19 Zhang Z F, Xu J K, Liu M L. Constructions of optimal locally repairable codes over small fields (in Chinese). Sci Sin Math, 2017, 47: 1607–1614 [张志芳, 徐敬可, 刘木兰. 构造小域上的最优局部修复码. 中国科学: 数学, 2017, 47: 1607–1614]
20 Guruswami V, Xing C, Yuan C. How long can optimal locally repairable codes be? IEEE Trans Inform Theory, 2019, 65: 3662–3670
21 Luo Y, Xing C, Yuan C. Optimal locally repairable codes of distance 3 and 4 via cyclic codes. IEEE Trans Inform Theory, 2019, 65: 1048–1053
22 Cai H, Miao Y, Schwartz M, et al. On optimal locally repairable codes with super-linear length. IEEE Trans Inform Theory, 2020, 66: 4853–4868
23 Chen B, Fang W, Xia S, et al. Constructions of optimal (r, δ) locally repairable codes via constacyclic codes. IEEE Trans Commun, 2019, 67: 5253–5263
24 Chen B, Xia S, Hao J, et al. Constructions of optimal cyclic (r, δ) locally repairable codes. IEEE Trans Inform Theory, 2018, 64: 2499–2511
25 Fang W, Fu F. Optimal cyclic (r, δ) locally repairable codes with unbounded length. Finite Fields Appl, 2020, 63: 101650
26 Xing C, Yuan C. Construction of optimal locally recoverable codes and connection with hypergraph. arXiv:1811.09142, 2018
27 Jin L. Explicit construction of optimal locally recoverable codes of distance 5 and 6 via binary constant weight codes. IEEE Trans Inform Theory, 2019, 65: 4658–4663
28 Hirschfeld J, Korchmáros G, Torres F. Algebraic Curves over a Finite Field. Princeton Series in Applied Mathematics. Princeton: Princeton University Press, 2008
29 Silverman J H. The Arithmetic of Elliptic Curves. Graduate Texts in Mathematics, vol. 106. New York: Springer-Verlag, 2009
30 Stichtenoth H. Algebraic Function Fields and Codes. Graduate Texts in Mathematics, vol. 254. Berlin: Springer-Verlag, 2009
31 Ling S, Xing C. Coding Theory—A First Course. Cambridge: Cambridge University Press, 2004
32 Ma L, Xing C. Constructions of block codes from algebraic curves over finite fields. In: Surveys in Combinatorics. London Mathematical Society Lecture Note Series, vol. 424. Cambridge: Cambridge University Press, 2015, 295–324
33 Niederreiter H, Xing C. Rational Points on Curves over Finite Fields: Theory and Applications. London Mathematical Society Lecture Note Series, vol. 285. Cambridge: Cambridge University Press, 2001
34 Tsfasman M A, Vlăduţ S G. Algebraic-Geometric Codes. Dordrecht: Kluwer, 1991
35 Tsfasman M A, Vlăduţ S G, Nogin D. Algebraic Geometric Codes: Basic Notions. Providence: Amer Math Soc, 1990
36 MacWilliams F J, Sloane N. The Theory of Error-Correcting Codes. Amsterdam: North-Holland, 1977
37 Schonheim J. On maximal systems of k-tuples. Studia Sci Math Hungar, 1996: 1: 363–368

38 Xing C, Ling J. A construction of binary constant-weight codes from algebraic curves over finite fields. IEEE Trans Inform Theory, 2005, 51: 3674–3678
39 Verstraete J. Extremal Problems for Cycles in Graphs. Cham: Springer, 2016, 83–116

A survey on optimal locally repairable codes

Liming Ma & Chaoping Xing

Abstract Locally repairable codes, or locally recoverable codes (LRC for short) are designed for application in distributed and cloud storage systems. Similar to the classical block codes, there is an important bound called the Singleton-type bound for locally repairable codes. In this paper, an optimal locally repairable code refers to a block code achieving this Singleton-type bound. Like the classical MDS (maximum distance separable) codes, the optimal locally repairable codes have some very nice combinatorial structures. Since the introduction of Singleton-type bound for locally repairable codes, people have put tremendous effort on constructions of optimal locally repairable codes. In this paper, we give a survey on optimal locally repairable codes, including bounds and constructions.

Keywords **locally repairable code, function field, linear code, parity-check matrice**

MSC(2020) **94B05, 68P20, 11T71, 14G50, 94B65**

doi: 10.1360/SSM-2020-0293