第 19 卷 第 1 期 电 子 科 学 学 刊 Vol.19 No.1
1997 年 1 月 JOURNAL OF ELECTRONICS Jan. 1997

GF(q) 上非线性码的距离分布的均值和均方差 [1]

符方伟 沈世镒

(南开大学数学系 天津 300071)

摘 要 本文给出 GF(q) 上非线性码的距离分布的均值和均方差的下界和上界，并且举例说明这些下界和上界是可以达到的。

关键词 编码理论，距离分布， Althöfer-Sillke 不等式， MacWilliams-Delsarte 恒等式

中图号 TN911.2

1 引 言

在重传反馈差错控制通信中，码的不可检错误概率是衡量通信系统好坏的一个重要性能指标。由文献 [1-4] 知计算码的不可检错误概率等价于确定码的距离分布，这促使人们去研究码的距离分布的性质和计算问题。对于线性码来说，码的距离分布等于码的重量分布。由于线性码具有较好的代数结构，因此关于线性码的重量分布的研究成果很多。但是对于非线性码来说，码的距离分布不一定等于码的重量分布，此时码的距离分布的计算十分困难，这方面的研究成果不多。既然很难确定码的距离分布，人们转而研究码的距离分布的某些数字特征 (例如均值和均方差等) 的计算问题。这些数字特征反映了码的距离分布的某些性质和数字信息。Althöfer 和 Sillke[5] 利用数学归纳法和经过一些复杂的计算给出了二元非线性码的距离分布的均值的下界和上界。本文利用 q 元 MacWilliams-Delsarte 恒等式给出了 q 元非线性码的距离分布的均值和均方差的下界和上界。对于均值情形，本文结果是文献 [5] 结果的推广，并且避免了文献 [5] 某些烦琐的计算。对于均方差情形，本文结果表明， q 元非线性码的情形下的结果完全不同于二元非线性码情形下的结果，其上界和下界的表达形式要复杂得多。最后我们举例说明这些上界和下界均可达到。如何给出更精确的上界和下界仍是一个尚待研究的问题。

2 主要结果

记 $V(n,q)$ 为 GF(q) 上 n 长向量空间， $d_H(\cdot,\cdot)$ 表示向量之间的 Hamming 距离。设 $C \subseteq V(n,q)$ 为一个 q 元 n 长码，码字个数为 M， E_i 为 C 中距离为 i 的码字对数目，记 $B_i = E_i/M$， $\{B_i\}_{i=0}^{n}$ 称为 C 的距离分布。实际上 C 的距离分布对应一个离散随机变量 X，取值 $0,1,\cdots,n$，

[1] 1995-02-13 收到， 1996-04-08 定稿

国家自然科学基金、高等学校博士学科点专项科研基金、国家教委优秀青年教师基金和回国留学人员科研基金资助项目

$P\{X=i\}=B_i/M, i=0,1,\cdots,n$ 。 C 的距离分布的均值定义为

$$\exp[C]\triangleq EX=\frac{1}{M}\sum_{i=0}^{n}iB_i. \tag{1}$$

C 的距离分布的均方差定义为

$$\mathrm{Var}[C]\triangleq DX=\frac{1}{M}\sum_{i=0}^{n}B_i[i-\exp[C]]^2. \tag{2}$$

容易验证

$$\exp[C]=\frac{1}{M^2}\sum_{a\in C}\sum_{b\in C}d_H(a,b), \tag{3}$$

$$\mathrm{Var}[C]=\frac{1}{M^2}\sum_{a\in C}\sum_{b\in C}[d_H(a,b)-\exp[C]]^2. \tag{4}$$

本文的主要结果为

定理 1　$[1+n(q-1)]/q-q^{n-1}/M\le\exp[C]\le n(q-1)/q.$

注：定理 1 中 $\exp[C]$ 的下界只有当 $M\ge q^n/[1+n(q-1)]$ 时才有意义。

定理 2

$$\mathrm{Var}[C]\ge\begin{cases}n(q-1)/q^2, & q\ne 2\text{且}M\ge q^n/(q-1)\text{时};\\ (n-1)(q-1)/q^2+q^{n-1}/M-(q^{n-1}/M)^2, & M\le q^n/(q-1)\text{时};\end{cases}$$

$$\mathrm{Var}[C]\le\begin{cases}[n(q-1)-2]/q^2+2q^{n-2}/M, & 2\le q\le 4\text{时};\\ (n-1)(q-1)/q^2+q^{n-1}/M-(q^{n-1}/M)^2, & q>4\text{且}M\ge 2q^n/(q-2)\text{时};\\ [4n(q-1)-8+(q-4)^2]/(4q^2)+2q^{n-2}/M, & q>4\text{且}M\le 2q^n/(q-2)\text{时}.\end{cases}$$

注： $q=2$ 时，一定有 $M\le q^n/(q-1)$ 。

推论 1　(Althöfer-Sillke 不等式 [5])　C 是二元 n 长码，码字数为 M ，则

$$(n+1)/2-2^{n-1}/M\le\exp[C]\le n/2.$$

推论 2　C 是二元 n 长码，码字数为 M ，则

$$(n-1)/4+2^{n-1}/M-(2^{n-1}/M)^2\le\mathrm{Var}[C]\le(n-2)/4+2^{n-1}/M.$$

例 1　取 $C=V(n,q)$ ，此时 $M=q^n$ ，距离分布为 $B_i=\binom{n}{i}(q-1)^i, i=0,1,\cdots,n$ 。

$$\exp[V(n,q)]=\frac{1}{q^n}\sum_{i=0}^{n}i\binom{n}{i}(q-1)^i=\frac{n(q-1)}{q},$$

此时同时达到定理 1 中均值的上界和下界。

$$\mathrm{Var}[V(n,q)]=\frac{1}{q^n}\sum_{i=0}^{n}\binom{n}{i}(q-1)^i\left[i-\frac{n(q-1)}{q}\right]^2=\frac{n(q-1)}{q^2}.$$

当 $q=2$ 时， $M=2^n=2^n/(2-1)$ ；当 $q\neq 2$ 时， $M=q^n>q^n/(q-1)$ ，经过简单计算易知，此时定理 2 中均方差的下界均为 $n(q-1)/q^2$ ，故此时达到定理 2 中均方差的下界。当 $q>4$ 时， $M=q^n\geq 2q^n/(q-2)$ ，经过简单计算易知 $q=2,3,4$ 和 $q>4$ 情形下定理 2 中均方差的上界均为 $n(q-1)/q^2$ ，故此时达到定理 2 中均方差的上界。这个例子说明定理 1 和定理 2 中的下界和上界均可达到。

例 2 固定 $h\in \mathrm{GF}(q)$ ，取 $C=A_h(n,q)=\{(h,a)|a\in V(n-1,q)\}$ ，此时 $M=q^{n-1}$ ，距离分布为 $B_i=\binom{n-1}{i}(q-1)^i, i=0,1,\cdots,n-1, B_n=0$, $\exp[A_h(n,q)]=(n-1)(q-1)/q$ ，容易验证此时达到定理 1 中均值的下界。 $\mathrm{Var}[A_h(n,q)]=(n-1)(q-1)/q^2$ ，此时 $M=q^{n-1}$ 满足 $M=q^{n-1}\leq q^n/(q-1)$ ，容易验证此时达到定理 2 中均方差的下界。此时定理 2 中均方差的上界当 $q=2,3,4$ 时为 $(n-1)(q-1)/q^2+3(q-1)/q^2$ ；当 $q>4$ 时为 $(n-1)(q-1)/q^2+(q+2)^2/(4q^2)$ 。

例 3 取 C 为 $\mathrm{GF}(q)$ 上的极长码 $S_r(q)$ ，即 q 元 Hamming 码的对偶码，此时码长 $n=(q^r-1)/(q-1)$ ，码字数 $M=q^r$ ，任意两个不同码字的距离为 q^{r-1} ，距离分布为 $B_0=1$, $B_{q^{r-1}}=q^r-1$, $B_i=0, i\neq 0, q^{r-1}$ ，则 $\exp[S_r(q)]=(q^r-1)/q$ 。容易验证此时达到定理 1 中的均值的上界。此时定理 1 中均值的下界为 $(q^r-q^{n-r})/q$ 。

$\mathrm{Var}[S_r(q)]=(q^r-1)/q^2$ ，而此时定理 2 中均方差的下界为 $(q^r-1)/q^2-[q^{2(n-r-1)}-q^{n-r-1}+(q-1)/q^2]$ ，均方差的上界当 $q=2,3,4$ 时为 $(q^r-1)/q^2+2(q^{n-r}-1)/q^2$ ；当 $q>4$ 时为 $(q^r-1)/q^2+(q-4)^2/(4q^2)+2(q^{n-r}-1)/q^2$ 。

3 q 元 MacWilliams-Delsarte 恒等式和几个引理

设 C 是一个 q 元 n 长码，码字个数为 M ， $\{B_i\}_{i=0}^n$ 为它的距离分布，称 $D(s)=\sum_{i=0}^n B_i s^i$ 为 C 的距离分布函数。记 $w_H(\cdot)$ 为向量的 Hamming 重量， $<\cdot,\cdot>$ 为向量的内积。设 $\chi(\cdot)$ 为 $(\mathrm{GF}(q),+)$ 上的一个非退化特征 (定义见文献 [6]， 第 13 页或文献 [7]， 第 216 页) 。令

$$\hat{B}_i=\frac{1}{M}\sum_{\substack{u\in V(n,q)\\ w_H(u)=i}}\left|\sum_{c\in C}\chi(<u,c>)\right|^2, \quad i=0,1,\cdots,n,$$

由文献 [7]， 第 226~230 页知 $\{B_i\}_{i=0}^n$ 满足 $\hat{B}_i\geq 0$, $\hat{B}_0=M$, $\sum_{i=0}^n\hat{B}_i=q^n$ 。令 $\hat{D}(s)=\sum_{i=0}^n\hat{B}_i s^i$ ，则得 (q 元 MacWilliams-Delsarte 恒等式，文献 [7] 第 220~237 页) ：

$$\hat{D}(s)=[1+(q-1)s]^n D\left(\frac{1-s}{1+(q-1)s}\right), \tag{5}$$

$$D(s)=\frac{1}{q^n}[1+(q-1)s]^n\hat{D}\left(\frac{1-s}{1+(q-1)s}\right). \tag{6}$$

引理 1 $\exp[C]=n(q-1)/q-\hat{B}_1/(qM)$ 。

证明 (6) 式两边对 s 取导数后令 $s=1$ ，即得结论。

引理 2 $\mathrm{Var}[C]=n(q-1)/q^2+[(q-2)/q^2](\hat{B}_1/M)-(1/q^2)(\hat{B}_1/M)^2+(2/q^2)(\hat{B}_2/M)$ 。

证明　(6) 式两边对 s 取 2 次导数后令 $s=1$，经过简单计算后 (利用引理 1) 即知结论成立。

4　定理的证明

4.1　定理 1 的证明

由引理 1 和 $\hat{B}_1 \geq 0$ 知 $\exp[C] \leq n(q-1)/q$。由 $\{\hat{B}_i\}_{i=0}^n$ 的性质知 $\hat{B}_1 = q^n - M - \sum_{i=2}^n \hat{B}_i \leq q^n - M$。故由引理 1 知

$$\exp[C] \geq n(q-1)/q - (1/q)(q^n/M - 1) = [1 + n(q-1)]/q - q^{n-1}/M,$$

则定理 1 结论成立。

4.2　定理 2 的证明

设 $f(x) = [(q-2)/q]x - x^2, x > 0$，具有性质：(1) 当 $0 \leq x \leq (q-2)/q$ 时，$f(x) \geq 0$；当 $x \geq (q-2)/q$ 时，$f(x) \leq 0$；(2) 当 $0 \leq x \leq (q-2)/(2q)$ 时，$f(x)$ 严格单增；当 $x \geq (q-2)/(2q)$ 时，$f(x)$ 严格单降，且 $f(x)$ 在 $x = (q-2)/(2q)$ 达到最大值。

令 $x_0 = \hat{B}_1/(qM)$，因为 $0 \leq \hat{B}_1 \leq q^n - M$，则 $0 \leq x_0 \leq q^{n-1}/M - 1/q$。由引理 2 知 $\mathrm{Var}[C] \geq n(q-1)/q^2 + f(x_0)$。当 $q \neq 2$ 且 $q^{n-1}/M - 1/q \leq (q-2)/q$ 时，即 $q \neq 2$ 且 $M \geq q^n/(q-1)$ 时，$f(x_0) \geq 0$，则 $\mathrm{Var}[C] \geq n(q-1)/q^2$。当 $M < q^n/(q-1)$ 时，$f(x_0) \geq f(q^{n-1}/M - 1/q)$，则

$$\mathrm{Var}[C] \geq n(q-1)/q^2 + f(q^{n-1}/M - 1/q) = (n-1)(q-1)/q^2 + q^{n-1}/M - (q^{n-1}/M)^2。$$

综合以上所述知定理 2 中 $\mathrm{Var}[C]$ 的下界成立。

由 $\{\hat{B}_i\}_{i=0}^n$ 的性质知 $0 \leq \hat{B}_2 \leq q^n - M - \hat{B}_1$，则由引理 2 知

$$\begin{aligned}\mathrm{Var}[C] &\leq n(q-1)/q^2 + [(q-2)/q^2](\hat{B}_1/M) - (1/q^2)(\hat{B}_1/M)^2 + (2/q^2)(q^n - M - \hat{B}_1)/M \\ &= [n(q-1) - 2]/q^2 + 2q^{n-2}/M + [(q-4)/q^2](\hat{B}_1/M) - (1/q^2)(\hat{B}_1/M)^2. \qquad (7)\end{aligned}$$

当 $2 \leq q \leq 4$ 时，$\mathrm{Var}[C] \leq [n(q-1)-2]/q^2 + 2q^{n-2}/M$。当 $q > 4$ 时，令 $g(x) = [(q-4)/q]x - x^2$，$x > 0$，具有性质：(1) 当 $0 \leq x \leq (q-4)/q$ 时，$g(x) \geq 0$；当 $x \geq (q-4)/q$ 时，$g(x) \leq 0$。(2) 当 $0 \leq x \leq (q-4)/(2q)$ 时，$g(x)$ 严格单增；当 $x \geq (q-4)/(2q)$ 时，$g(x)$ 严格单降，且 $g(x)$ 在 $x = (q-4)/(2q)$ 达到最大值。

由 (7) 式知 $\mathrm{Var}[C] \leq [n(q-1)-2]/q^2 + 2q^{n-2}/M + g(x_0)$。因为 $0 \leq x_0 \leq q^{n-1}/M - 1/q$，则当 $q^{n-1}/M - 1/q \leq (q-4)/(2q)$ 时，即 $M \geq 2q^n/(q-2)$ 时，$g(x_0) \leq g(q^{n-1}/M - 1/q)$，故

$$\mathrm{Var}[C] \leq (n-1)(q-1)/q^2 + q^{n-1}/M - (q^{n-1}/M)^2。$$

当 $M < 2q^n/(q-2)$ 时，$g(x_0) \leq g[(q-4)/(2q)]$，故

$$\mathrm{Var}[C] \leq [4n(q-1) - 8 + (q-4)^2]/(4q^2) + 2q^{n-2}/M.$$

综合以上所述知定理 2 中 Var[C] 的上界成立。

参 考 文 献

[1] 王新梅. 最佳 $(n,2,w)$ 二进制等重纠错码的存在性和猜想. 中国科学 (A 辑), 1987, 17(11): 1225–1232.

[2] 王新梅. 非线性等重码的不可检错误概率. 电子学报, 1989, 17(1): 8–14.

[3] 王新梅. 非线性等重码检错性能的进一步分析. 通信学报, 1992, 13(4): 10–17.

[4] 杨义先. 截短 Hamming 码和截短 R-M 码的不可检错误概率. 通信学报, 1991, 12(4): 38–45.

[5] Althöfer I, Sillke T. An average distance inequality for large subsets of the cube. Journal of Combinatorial Theory, 1992, 56B(2): 296–301.

[6] Van Lint J H. Introduction to Coding Theory, Graduate Texts in Mathematics. Berlin: Springer-Verlag, 1982, 13–15.

[7] Roman S. Coding and Information Theory, Graduate Texts in Mathematics. Berlin: Springer-Verlag, 1991, 220–237.

ON THE MEAN VALUE AND VARIANCE OF DISTANCE DISTRIBUTION OF NON-LINEAR CODES IN GF(q)

Fu Fangwei Shen Shiyi

(*Department of Mathematics, Nankai University, Tianjin* 300071)

Abstract This paper presents lower bounds and upper bounds for the mean value and variance of distance distribution of non-linear codes in GF(q). By presenting several examples, it is shown that these bounds could be achieved.

Key words Coding theory, Distance distribution, Althöfer-Sillke inequality, MacWilliams-Delsarte identity

符方伟： 男， 1963 年生，教授，博士，主要从事信息论、编码理论和密码学理论的研究和教学工作.

沈世镒： 男， 1939 年生，教授，博士生导师，主要从事信息论、编码理论、密码学理论和神经网络的数学理论的研究和教学工作.