Reed-Muller Codes and Incidence Matrices

M. Karimi Behbahani G. B. Khosrovshahi[1] B. Tayfeh-Rezaie

Institute for Studies in Theoretical Physics and Mathematics (*IPM*)
P.O. Box 19395-5746*, Tehran, Iran*

Abstract

In this paper, we discuss Reed-Muller codes using a set-theoretic approach. We present a new basis of minimum weight codewords which correspond to subspaces of a fixed dimension of affine geometry over $GF(2)$. A generator matrix in standard form is also given.

Index Words: Reed-Muller codes, Basis of minimum weight codewords, Incidence matrix, Standard form
MR Subject Classification: 94B05, 05B20

1 Introduction

Reed-Muller codes are one of the most important families of codes studied so far. These codes are important from points of view of coding theory, design theory and finite geometries. Reed-Muller codes can be defined by several different approaches. One of the simplest ways is to define these codes by the use of boolean functions. Let $v_1, \ldots, v_m$ be m boolean variables. A binary vector of length 2^m is obtained for each boolean function $f(v_1, \ldots, v_m)$ when the variables take all 2^m possible values. The set of vectors of all boolean functions $f(v_1, \ldots, v_m)$ which are polynomials of degree at most r is a binary linear code called Reed-Muller code of order r and is denoted by $\mathcal{R}(r, m)$. Reed-Muller codes are related to the geometric codes. It is well known that the binary code of design of points and $(m-r)$-flats in affine geometry $AG_m(\mathbb{F}_2)$ is $\mathcal{R}(r, m)$. It is also known that the incidence vectors of $(m-r)$-flats are exactly the minimum weight codewords. A detailed discussion of many properties

[1]Corresponding author, email: rezagbk@ipm.ir.

of Reed-Muller codes can be found at [4]. Also, a complete account of these codes and generalized Reed-Muller codes is given in [2].

In this paper, we study Reed-Muller codes via a set-theoretic approach. We will define these codes in terms of incidence matrices of subsets of a given set. In [1], it is shown that the subspaces of dimension $m-r$ of $AG_m(\mathbb{F}_2)$ generate $\mathcal{R}(r,m)$. We present a basis of minimum weight codewords for $\mathcal{R}(r,m)$ which correspond to subspaces of dimension $m-r$ in $AG_m(\mathbb{F}_2)$. We also provide a generator matrix in standard form for this family of codes.

2 Definitions

For definitions and notations on binary linear codes, the reader is referred to [4].

Let t,k,v be integers such that $0 \leq t \leq k \leq v$. The *incidence matrix* of t-subsets versus k-subsets of a v-set V, denoted by $W_{tk}(v)$, is a (0,1)-matrix whose rows and columns are indexed by the t-subsets and k-subsets of V, respectively and $W_{tk}(v)(T,K) = 1$ if and only if $T \subseteq K$. If there is no confusion, we write W_{tk} instead of $W_{tk}(v)$.

Let m and r be integers such that $0 \leq r \leq m$ and let $X = \{1,2,\ldots,m\}$. Let $P_{r,m}$ be a (0,1)-matrix whose rows are indexed by the i-subsets of X for all $0 \leq i \leq r$ and whose columns are indexed by the power set of X and $P_{r,m}(R,S) = 1$ if and only if $R \subseteq S$. By convention, we suppose that the indices of rows and columns of $P_{r,m}$ are ordered lexicographically. It is clear that the entry (i,j) of $P_{r,m}$ is 0 for $i > j$ and so $P_{r,m}$ is full rank over any field. The binary code with the generator matrix $P_{r,m}$ is called *Reed-Muller code of order* r, and is denoted by $\mathcal{R}(r,m)$. It is not hard to see that this is equivalent with the traditional definition of Reed-Muller codes by the use of boolean functions. The next two theorems are well known facts about Reed-Muller codes.

Theorem 1. *$\mathcal{R}(r,m)$ is a $[2^m, \sum_{i=0}^{r}\binom{m}{i}, 2^{m-r}]$ code.*

Proof. The dimension of code is $\sum_{i=0}^{r}\binom{m}{i}$, since $P_{r,m}$ is a full rank matrix. For minimum distance d, we consider the following recursive structure of $P_{r,m}$:

$$P_{r,m} = \begin{bmatrix} P_{r-1,m-1} & 0 \\ P_{r,m-1} & P_{r,m-1} \end{bmatrix}.$$

We use an induction argument on $m+r$. The assertion holds trivially for $\mathcal{R}(0,m)$ and $\mathcal{R}(m,m)$. By the induction hypothesis, the minimum distances of $\mathcal{R}(r-1,m-1)$ and $\mathcal{R}(r,m-1)$ are $d_1 = 2^{m-r}$ and $d_2 = 2^{m-1-r}$, respectively. Therefore, we have

$$\begin{aligned} d &= \min(d_1, 2d_2) \\ &= 2^{m-r}. \end{aligned}$$

$\square$

Theorem 2. *$\mathcal{R}(r,m)^\perp = \mathcal{R}(m-1-r,m)$ for $0 \leq r < m$.*

Proof. The sum of dimensions of $\mathcal{R}(r,m)^\perp$ and $\mathcal{R}(m-1-r,m)$ is 2^m. Therefore, it suffices to show that the rows of $P_{r,m}$ are orthogonal to the rows of $P_{m-1-r,m}$. Consider two arbitrary rows I and J of $P_{r,m}$ and $P_{m-1-r,m}$, respectively. The number of subsets K of X containing both I and J is equal to $2^{m-|I\cup J|}$ and this number module 2 is the product of rows I and J. Since $|I \cup J| \leq m-1$, we find that the rows I and J are orthogonal to each other. $\square$

We now show that the minimum weight codewords generate Reed-Muller code by giving a basis which only consists of this type of codewords. We define a polynomial function for every subset S of X by

$$f_S = \prod_{i \in S} x_i.$$

For the empty set, define $f_\varnothing = 1$. Now, for each codeword c, we set

$$f_c = \sum_{c(S)=1} f_S.$$

Theorem 3. *For every subset $K = \{i_1, \ldots, i_k\}$ of X provided that $k \geq r$ and $i_1 < i_2 < \cdots < i_k$, let*

$$h_K = \left(\prod_{j=1}^{k-r+1} x_{i_j} + 1 \right) \prod_{j=k-r+2}^{k} (x_{i_j} + 1).$$

Then $\{h_K\}$ is a basis of minimum weight codewords for $\mathcal{R}(m-r,m)$.

Proof. The number of h_K is $\sum_{i=r}^{m} \binom{m}{i}$ which is equal to $\sum_{i=0}^{m-r} \binom{m}{i}$. Also, the weight of any h_K is clearly 2^r. We have to show that h_K are really codewords. Consider an specific h_K and suppose that c is its corresponding vector. By Theorem 2, we need to show that $P_{r-1,m}.c = 0$. Let I be a row of $P_{r-1,m}$. If $I \not\subseteq K$, then the product of row I and c is clearly 0. So let $I \subseteq K$. Assume that

$$|I \cap \{i_1, \ldots, i_{k-r+1}\}| = a,$$
$$|I \cap \{i_{k-r+2}, \ldots, i_k\}| = b.$$

First, suppose that $a = 0$ and $b = |I|$. Then the product is equal to $2.2^{r-1-b} = 2^{r-b} \equiv 0 \pmod 2$, since $|I| \leq r-1$. Now assume that $a \neq 0$, then $b < |I| \leq r-1$, and the product is equal to $2^{r-1-b} \equiv 0 \pmod 2$.

It is easy to see that $\sum h_{K_i}$ is always nonzero. Choose K_0 among K_i with the largest cardinality. Trivially, $\sum h_{K_i}$ has f_{K_0} as one of its components. $\square$

Note. The above basis of $\mathcal{R}(m-r, m)$ can be expressed by means of subspaces of dimension r of $AG_m(\mathbb{F}_2)$. For any h_K, the corresponding subspace is given by the following $m-r$ equations:

$$X_{i_j} = 0, \qquad j \notin \{1, \ldots, k\},$$
$$X_{i_1} = X_{i_j}, \qquad j \in \{2, \ldots, k-r+1\}.$$

Note that a basis of r-flats is given in [3]. For any K (as in Theorem 3), the corresponding r-flat is given by the following $m-r$ equations:

$$X_{i_j} = 0, \qquad j \notin \{1, \ldots, k\},$$
$$X_{i_j} = 1, \qquad j \in \{r+1, \ldots, k\}.$$

3 Standard Form

In this section, we find a generator matrix in standard form for Reed-Muller codes. We use the following simple and well known identity for incidence matrices:

$$W_{it}W_{tk} = \binom{k-i}{t-i} W_{ik}, \tag{1}$$

which holds for $0 \leq i \leq t \leq k$.

Using (1), we can transform the generator matrix $P_{r,m}$ to a standard form. We need to consider the block structure of $P_{r,m}$. In this presentation of $P_{r,m}$, we index the rows and columns respectively with the numbers i and j such that $0 \leq i \leq r$ and $0 \leq j \leq m$. Now it is seen that the entry (i,j) is W_{ij}. For example, the block structure of $P_{2,6}$ is shown in Figure 1.

$$\begin{bmatrix} I & W_{0,1} & W_{0,2} & W_{0,3} & W_{0,4} & W_{0,5} & W_{0,6} \\ 0 & I & W_{1,2} & W_{1,3} & W_{1,4} & W_{1,5} & W_{1,6} \\ 0 & 0 & I & W_{2,3} & W_{2,4} & W_{2,5} & W_{2,6} \end{bmatrix}.$$

Figure 1. The block structure of $P_{2,6}$.

We proceed inductively to change $P_{r,m}$ to a standard form. Consider the row s ($0 \le s < r$). We add the row s with the coefficient W_{ls} to the row l to get vanished the entry (l, s) for any $0 \le l < s$. Eventually, we find a block matrix in the form $[I \;\; A]$ in which the entry (i, j) is $b(i, j)W_{ij}$ where $0 \le i \le r < j \le m$. The coefficients $b(i, j)$ are identified by the following recursive relations:

$$b(i, j) \equiv 1 + \sum_{l=i+1}^{r} \binom{j-i}{l-i} b(l, j) \pmod 2, \tag{2}$$

and $b(r, j) = 1$ for $r < j \le m$. For example, the standard form of $P_{2,6}$ is given in Figure 2. We now solve (2).

Lemma 1. $b(i, j) \equiv \binom{f-r+i}{j-r-1} \pmod 2$ *for any integers* $i, r, j, m,$ *and* f *such that* $0 \le i \le r < j \le m \le f = 2^n - 1$.

Proof. By (2), we need to show that

$$\sum_{l=i}^{r} \binom{j-i}{l-i} \binom{f-r+l}{j-r-1} \equiv 1 \pmod 2. \tag{3}$$

Let $\mathrm{M}_m(i, j, r)$ denote the left hand of (3). For $0 \le i \le r < j \le m < f$, we have

$$\begin{aligned} \mathrm{M}_{m+1}(i+1, j+1, r+1) &= \sum_{p=i+1}^{r+1} \binom{j-i}{p-i-1} \binom{f-r-1+p}{j-r-1} \\ &= \sum_{l=i}^{r} \binom{j-i}{l-i} \binom{f-r+l}{j-r-1} \\ &= \mathrm{M}_m(i, j, r). \end{aligned}$$

The same argument shows that $\mathrm{M}_m(i, j, r) = \mathrm{M}_m(i+1, j+1, r+1)$ for $0 \le i \le r < j < m$. Therefore, by induction it suffices to show that $\mathrm{M}_f(i, f, r) \equiv 1 \pmod 2$ or equivalently

$$\sum_{l=i}^{r} \binom{f-i}{l-i} \binom{f-r+l}{f-r-1} \equiv 1 \pmod 2,$$

$$\begin{bmatrix} I & 0 & 0 & W_{0,3} & W_{0,4} & 0 & 0 \\ 0 & I & 0 & W_{1,3} & 0 & W_{1,5} & 0 \\ 0 & 0 & I & W_{2,3} & W_{2,4} & W_{2,5} & W_{2,6} \end{bmatrix}.$$

Figure 2. A generator matrix of $\mathcal{R}(2,6)$ in standard form.

in which $0 \le i \le r < f$. We try to determine where the coefficients $\binom{f-i}{l-i}$ and $\binom{f-r+l}{f-r-1}$ are both odd numbers. So suppose that they are odd. Let $y = l - i,\ x = f - r + l$, and $x' = x + r + 1 - f$. Therefore, we have

$$x' = y + i + 1, \qquad i + 1 \le x' \le r + 1. \tag{4}$$

Consider the binary expansions of $x' = (x'_n \cdots x'_1)$, $y = (y_n \cdots y_1)$, $i = (i_n \cdots i_1)$, and $r + 1 = (r'_n \cdots r'_1)$. By Lucas lemma, we have

$$\begin{aligned} i_j = 1 &\Rightarrow y_j = 0, \\ r'_j = 0 &\Rightarrow x'_j = 0. \end{aligned} \tag{5}$$

Let $Z_{r+1} = \{j |\ r'_j = 0\}$ and $O_i = \{j |\ i_j = 1\}$. Let k_0 and e be the smallest and largest integers such that $x'_{k_0} = 1$ and $r'_e = 1$, respectively. By (4) and (5), $k_0 \notin O_i \cup Z_{r+1}$. By (4) and (5), we also have

$$\begin{aligned} x'_k = 1, &\qquad k \in O_i \quad \& \quad k > k_0 \\ x'_k = 0, &\qquad k \in Z_{r+1} \quad \& \quad k > k_0. \end{aligned} \tag{6}$$

By (6), there is no $k > k_0$ such that $k \in O_i \cap Z_{r+1}$. Let h denote the number of $k > k_0$ such that $k \in O_i \cup Z_{r+1}$. Now fix a k_0. Then the number of solutions x' of (4) is 2^{e-k_0-h}. This is 1 only for $k_0 = e - h$. Therefore, the number of x' satisfying (4) is always odd. This completes the proof. $\square$

References

[1] E. F. Assmus, *On Berman's characterization of the Reed-Muller codes*, Special issue on orthogonal arrays and affine designs, Part I, J. Statist. Plann. Inference, 56 (1996), no. 1, 17–21.

[2] E. F. Assmus and J. D. Key, *Designs and their codes*, Cambridge University Press, Cambridge, 1992.

[3] S.Gao, J. D. Key, *Bases of minimum-weight vectors for codes from designs*, Finite Fields Appl., 4 (1998), no. 1, 1–15.

[4] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes I, II*, North-Holland Mathematical Library, Vol. 16, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977.