Math 2^8: The Theory of Error-Correcting Codes

"Linear programming" (LP) bounds I

We've noted already that a necessary condition for there to be an $[n, k, \geq d]$ code over a q-letter field is the existence of a candidate weight enumerator $W_C(X, Y) = X^n + O(Y^d)$ such that $W_C(1, 1) = q^k$ and both $W_C(X, Y)$ and $W_C(X + (q-1)Y, X - Y)$ have nonnegative coefficients (the latter because $q^{-k}W_C(X + (q-1)Y, X - Y)$ is the weight enumerator of $C^\perp$ by MacWilliams). This amounts to a collection of linear (in)equalities on the coefficients N_w of W_C, whose feasibility is a question of linear programming (LP). Equivalently, we can fix n and d and, instead of prescribing $|C| = W_C(1, 1)$, ask to maximize it subject to the other constraints, which again is an LP problem. Our final topic for Math 256 is studying the resulting upper bounds. In this handout, we first obtain the coefficients of these inequalities, and recognize them as Krawtchouk polynomials. We then generalize to codes that need not be linear. Finally we consider what happens in the rare cases that equality is attained.

As usual let $N_w = N_w(C)$ be the number of codewords of weight w. Then $N_0 = 1$, $N_w \geq 0$ for all w, and $N_w = 0$ for $0 < w < d$. Also N_w is the $X^{n-w}Y^w$ coefficient of $W_C(X, Y)$, so each coefficient $N_i^\perp$ of $W_C(X + (q-1)Y, X - Y)$ is a linear combination of the N_w, say $N_i^\perp = \sum_{w=0}^n \kappa_i(w)N_w$. The multiplier $\kappa_i(w)$ is the $X^{n-i}Y^i$ coefficient of

$$(X-Y)^w(X+(q-1)Y)^{n-w} = \left(\sum_{j=0}^{w}(-1)^j\binom{w}{j}X^{w-j}Y^j\right)\left(\sum_{j'=0}^{n-w}(q-1)^{j'}\binom{n-w}{j'}X^{n-w-j'}Y^{j'}\right).$$

Expanding and collecting the terms with $j + j' = i$ we find that

$$\kappa_i(w) = \sum_{j=0}^{i}(-1)^j(q-1)^{i-j}\binom{w}{j}\binom{n-w}{i-j} = K_i(w),$$

the value at w of the degree-i Krawtchouk polynomial. (This coincidence is not at all mysterious: remember that we obtained the MacWilliams identity via the discrete Fourier transform of the function $c \mapsto X^{n-\mathrm{wt}(c)}Y^{\mathrm{wt}(c)}$, and were led to K_i via the discrete Fourier transform of 1_{S_i}.) Notice that this also gives us the generating function

$$\sum_{i=0}^{n} K_i(x)X^{n-i}Y^i = (X + (q-1)Y)^{n-x}(X - Y)^x$$

for the K_i; we shall need this formula several times in the sequel.

Now suppose that C is any $(n, M, \geq d)$ code, not necessarily linear, over some alphabet F of $q > 1$ elements. We do not assume that q is a prime power, so F need not have the structure of a finite field; but we can always choose some identification of F with an abelian group so that we can do discrete Fourier analysis on F^n.[1] Then it is still true that $d(c, c') = \mathrm{wt}(c - c')$ for all words c, c'. We no longer assume that $c - c' \in C$ when $c, c' \in C$, but can still keep track of the distribution of Hamming distances in C by defining

$$W_C(X, Y) = \frac{1}{M}\sum_{c,c' \in C} X^{n-d(c,c')}\,Y^{d(c,c')},$$

[1] A more natural approach, but one that takes somewhat longer to develop, is to consider F^n as a homogeneous space for its full automorphism group $S_q^n \rtimes S_n$, which acts distance-transitively on F^n, and apply the theory of Gelfand pairs to this group and its point stabilizer.

with the factor $1/M$ making this definition agree with our previous definition of W_C in the case that C is linear. Thanks to that factor the identity $W_C(1,1) = M$ still holds as well. We next show that $W_C(X + (q-1)Y, X - Y)$ has nonnegative coefficients in this context too, and thus that any bound on $|C|$ obtained from linear programming holds equally for linear and nonlinear codes. As before, denote by $1_C \in A[F^n]$ the characteristic function of C, representing the formal sum $\sum_{c \in C} c$. Then W_C encodes the weight distribution of $M^{-1} 1_C 1_{-C}$ (which is the same as 1_C in the linear case but not in general). For any character $\chi : F^n \to \mathbf{C}^*$, extended as before to a ring homomorphism $\mathbf{C}[F^n] \to \mathbf{C}$, we have $\chi(1_{-C}) = \overline{\chi(1_C)}$, because $\chi(-c) = \overline{\chi(c)}$ for all $c \in F^n$; therefore

$$\chi\big(\frac{1}{M} 1_C 1_{-C}\big) = \frac{1}{M} \chi(1_C)\chi(1_{-C}) = \frac{1}{M} |\chi(1_C)|^2 \geq 0.$$

Thus $M^{-1} 1_C 1_{-C}$ has a nonnegative Fourier transform. We now apply Parseval to the inner product $W_C(X + (q-1)Y, X - Y)$ of $M^{-1} 1_C 1_{-C}$ with the function

$$w \mapsto (X + (q-1)Y)^{n - \mathrm{wt}(w)} (X - Y)^{\mathrm{wt}(w)}.$$

The Fourier transform of this function is $w \mapsto q^n X^{n-\mathrm{wt}(w)} Y^{\mathrm{wt}(w)}$, whose inner product with a nonnegative function is a homogeneous polynomial in X and Y with nonnegative coefficients, Q.E.D.

We can deduce an upper bound on M as follows. Suppose $P = \sum_{i=0}^n a_i K_i$ is a linear combination with nonnegative coefficients a_i, such that $a_0 > 0$ and $P(x) \leq 0$ for all $x \geq d$. Necessarily $P(0) > 0$ because $K_i(0) = (q-1)^i \binom{n}{i} > 0$ for each i. Then

$$P(0) = P(0) N_0 \geq \sum_{w=0}^n P(w) N_w = \sum_i a_i \left(\sum_w K_i(w) N_w \right) \geq a_0 \sum_w N_w = a_0 M,$$

whence $M \leq P(0)/a_0$. The problem of minimizing $P(0)/a_0$ subject to the linear inequalities $a_i \geq 0$, $P(x) \leq 0$ is again a LP problem (e.g. fix $a_0 = 1$ and minimize $P(0)$), in fact the dual of our original LP problem. Working with the dual problem has the advantage that if we need only check that the solution is feasible, not that it is optimal (which we can check for any particular choice of q, n, d but probably not in the asymptotic setting). If a code actually attains $M = P(0)/a_0$ then equality must hold throughout, so in particular $N_w = 0$ for all $w > 0$ where $P(w) \neq 0$, which is a strong constraint.

We illustrate with the extended Golay code $\mathcal{G}_{24}$, where equality is attained. Set $q = 2$, $n = 24$, and $d = 8$. Carrying out the LP computation we find $P(x) = (x-8)(x-12)^2(x-16)^2(x-24)$, which we might also have tried by guessing that P should vanish at the nonzero weights of $\mathcal{G}_{24}$. We find $(a_0, a_1, a_2, a_3, a_4, a_5, a_6) = (1728, 1728, 3447/4, 468, 168, 60, 45/4)$ and $a_i = 0$ for $i > 6$. Therefore any binary code of length 24 and minimal distance at least 8 has size at most $8 \ 12^2 \ 16^2 \ 24/1728 = 4096 = 2^{12}$, and any code C attaining this bound must have all distances in $\{0, 8, 12, 16, 24\}$. We can get further information about C from the condition that $\sum_w K_i(w) N_w = 0$ for $1 \leq i \leq 6$; for example, using $P = (x-8)(x-12)^2(x-16)$ we find that $a_0 = 9$ and $P(0)(N_0 + N_{24}) = a_0 |C|$, whence $N_0 + N_{24} = 2$, giving $N_{24} = 1$ since we know that $N_0 = 1$. That is, C is closed under ones' complement. Similarly we can obtain $N_8 = N_{16} = 759$ and $N_{12} = 2576$. In this case we don't even need this additional information to deduce that C is a translate of $\mathcal{G}_{24}$, because we can argue as follows. Translate C so that it contains zero. Then C consists of words of doubly even weight whose pairwise differences are also doubly even. Therefore the linear span $\langle C \rangle$ is a doubly even linear code, and thus self-orthogonal, so of size at most $2^{n/2} = 2^{12}$. But C already has size 2^{12}, so $C = \langle C \rangle$, i.e. C is a Type II $[24, 12, 8]$ code — and we have already proved that $\mathcal{G}_{24}$ is the unique such code up to coordinate permutation.

Exercise: For $(q, n, d) = (3, 12, 6)$, prove in the same way that $|C| \leq 3^6 = 729$, and if equality holds then C is a translate of a Type III code. (It follows that $C \cong \mathcal{G}_{12}$, though we haven't actually given this argument in class.)