A functional view of upper bounds on codes

Alexander Barg[1,*] and Dmitry Nogin[2]

[1] University of Maryland, College Park, MD, USA
and Dobrushin Mathematical Laboratory, IITP RAS, Moscow, Russia
abarg@umd.edu

[2] Dobrushin Mathematical Laboratory, IITP RAS
Moscow, Russia
nogin@iitp.ru

1. Introduction. In the problem of bounding the size of codes in compact homogeneous spaces, Delsarte's polynomial method gives rise to the most powerful universal bounds on codes. Many overviews of the method exist in the literature; see for instance Levenshtein (1998). The purpose of this talk is to present a functional perspective of this method and give some examples.

Let X be a compact metric space whose isometry group G acts transitively on it. The zonal polynomials associated with this action give rise to a family of orthogonal polynomials $\mathcal{P}(X) = \{P_\kappa\}$ where $\kappa = 0, 1, \dots$ is the total degree. These polynomials are univariate if G acts on X doubly transitively (the well-known examples include the Hamming and Johnson graphs and their q-analogs and other Q-polynomial distance-regular graphs; the sphere $S^{n-1} \in \mathbb{R}^n$) and are multivariate otherwise.

First consider the univariate case. Then for any given value of the degree $\kappa = i$ the family $\mathcal{P}(X)$ contains only one polynomial, denoted below by P_i. Suppose that the distance on X is measured in such a way that $d(x, x) = 1$ and the diameter of X equals -1 (to accomplish this, a change of variable is made in the natural distance function on X). We refer to the model case of $X = S^{n-1}$ although the arguments below apply to all spaces X with the above properties. Let $\langle f, g\rangle = \int_{-1}^{1} fg d\mu$ be the inner product in $L_2([-1, 1], d\mu)$ where $d\mu(x)$ is a distribution on $[-1, 1]$ induced by an invariant measure on G. We assume that $\langle 1, 1\rangle = 1$.

By Delsarte's fundamental theorem, the size of the code $C \subset X$ whose distances take values in $[-1, s]$ is bounded above by

$$|C| \le \inf_{f\in\Phi} f(1)/\hat{f}(0) \tag{1}$$

where

$$\Phi = \{f : f(x) \le 0, x \in [-1, s]; \quad \hat{f}(0) > 0,\ \hat{f}(i) \ge 0, i = 1, 2, \dots\}, \tag{2}$$

where $\hat{f}(i) = \langle f, P_i\rangle$ are the Fourier coefficients of f.

2. A functional approach. 2.1. Notation. Let V be the space of real square-integrable functions on $[-1, 1]$ and let V_k be the space of polynomials of degree k or less. Let $p_i = P_i/\langle P_i, P_i\rangle, i = 0, 1, \dots$ be the normalized polynomials. The polynomials $\{p_i\}$ satisfy a three-term recurrence of the form

$$\begin{aligned} xp_i &= a_i p_{i+1} + b_i p_i + a_{i-1} p_{i-1}, \\ &\qquad i = 1, 2, \dots; p_{-1} = 0, p_0 = 1;\ a_{-1} = 0. \end{aligned} \tag{3}$$

In other words, the matrix of the operator $x : V \to V$ (multiplication by the argument) in the orthonormal basis is a semi-infinite symmetric tridiagonal matrix. Let $X_k = E_k \circ x$ where $E_k = \mathrm{proj}_{V\to V_k}$. It is well known that for $k \ge 1$ the spectrum of X_k coincides with the set $\{x_{k+1,1}, \dots, x_{k+1,k+1}\}$ of zeros of p_{k+1}. Below we denote the largest of these zeros by x_{k+1}. Let

$$K_k(x, s) \triangleq \sum_{i=0}^{k} p_i(s)p_i(x)$$

be the k-th reproducing kernel. By the Christoffel-Darboux formula,

$$(x - s)K_k(x, s) = a_k(p_{k+1}(x)p_k(s) - p_{k+1}(s)p_k(x)). \tag{4}$$

[*] Research supported in part by NSF grants CCF0515124 and CCF0635271, and by NSA grant H98230-06-1-0044.

In particular, $X_k K_k(x, x_{k+1,i}) = x_{k+1,i} K_k(x, x_{k+1,i})$. Note that $K_k(x, y)$ acts on V_k as a delta-function at y:

$$\langle K_k(\cdot, y), f(\cdot)\rangle = f(y). \tag{5}$$

2.2. Remarks on the choice of polynomials. The choice of polynomials for problem (1)-(2) was studied extensively in the works of Levenshtein (1978-1998) and Sidelnikov (1980). For any given n, polynomials that attain the minimum in (1) and satisfy conditions (2) are known from these works. The purpose of this section is to observe how these polynomials arise under the functional approach developed here.

Without loss of generality let us assume that $f(1) = 1$. We need to maximize the linear functional

$$\mathcal{F}(f) := \hat{f}(0) = \langle f, 1\rangle.$$

Let us restrict the class of functions to V_n. By the Markov-Lucacs theorem, a polynomial $f(x)$ that is nonpositive on $[-1, s]$ can be written in the form

$$f_n(x) = (x-s)g^2 - (x+1)\phi_1^2 \quad \text{or} \quad f_n(x) = (x+1)(x-s)g^2 - \phi_2^2$$

according as its degree $n = 2k+1$ or $2k+2$ is odd or even. Here $g, \phi_1 \in V_k, \phi_2 \in V_{k+1}$ are some polynomials. Below the negative terms will be discarded. We use a generic notation c for multiplicative constants chosen to fulfill the condition $f(1) = 1$.

2.2.1. The MRRW polynomial. Restricting our attention to odd degrees $n = 2k+1$, let us seek $f(x)$ in the form $(x-s)g^2$. Let us write the Taylor expansion of $\mathcal{F}$ in the "neighborhood" of g. Let $h \in V_k$ be a function that satisfies $\|h\| \le \varepsilon$ for a small positive ε and the condition $h(1) = 0$. We obtain

$$\begin{aligned}\mathcal{F}((x-s)(g+h)^2) &= \mathcal{F}((x-s)g^2) + \langle (x-s)(g+h), g+h\rangle - \langle (x-s)g, g\rangle \\ &= \mathcal{F}(f) + \mathcal{F}'(h) + {}^1\!/\!_2\langle \mathcal{F}''h, h\rangle\end{aligned}$$

where $\mathcal{F}' = 2(x-s)g, \mathcal{F}'' = 2(x-s)$ are the Fréchet derivatives of $\mathcal{F}$. This relation shows that for f to be a stationary point of $\mathcal{F}$, the function g should satisfy $d\mathcal{F} = 2\langle g, (x-s)h\rangle = 0$ for any function h with the above properties. First assume that $s = x_{k+1}$. Then by (4), a stationary point of $\mathcal{F}$ is given by $g = K_k(x, s)$, and we obtain f in the form

$$f_n(x) = c(x-s)(K_k(x,s))^2.$$

Since $\hat{f}_n(0) = 0$, conditions (2) are not satisfied; however, it can be checked that they are satisfied if $x_k < s < x_{k+1}$. For all such s, the polynomial f_n is a valid choice for problem (1). This polynomial was used in the works of McEliece *et al.* and Kabatiansky and Levenshtein to derive the MRRW and KL bounds on codes.

2.2.2. Levenshtein polynomials, $n = 2k+1$. So far in our optimization we did not use the condition $h(1) = 0$. To use it, let us write $h = (1-x)h_1, h_1 \in V_{k-1}$ and repeat the above calculation. We find that stationary points of $\mathcal{F}$ should satisfy

$$d\mathcal{F}^{(-)} = 2\langle (x-s)g, (1-x)h_1\rangle = 0,$$

where $\mathcal{F}^{(-)}(\,.\,) = \int .\,(1-x)d\mu$ is the moment functional w.r.t. the distribution $d\mu^{(-)}(x) = (1-x)d\mu(x)$. The stationary point of $\mathcal{F}^{(-)}$ is given by the reproducing kernel $K_k^-(x,s)$ *with respect to this distribution*:

$$K_k^-(x,s) = \sum_{i=0}^{k} p_i^-(s)p_i^-(x) \tag{6}$$

where $\{p_i^-(x), i = 0, 1, \dots\}$ is the corresponding orthonormal system. To find the polynomials $p_i^-(x)$ observe that

$$\mathcal{F}^{(-)}(p_i^- p_j^-) = \mathcal{F}(p_i^-(x)p_j^-(x)(1-x)) = \delta_{i,j}$$

is satisfied for $p_i^-(x) = K_i(1,x)/(a_i p_{i+1}(1)p_i(1))^{1/2}$. Indeed, if $j < i$ then the function $(1-x)K_i(1,x)$ is in the subspace spanned by p_{i+1}, p_i and thus is orthogonal to $K_j(1,x)$. To conclude, the function sought can be taken in the form

$$f_n^-(x) = c(x-s)(K_k^-(x,s))^2.$$

2.2.3. Levenshtein polynomials, $n = 2k+2$. In this case we seek the polynomial in the form $f_n = (x-s)(x+1)g^2$. The necessary condition for the stationary point takes the form $\langle (x-s)(1+x)(1-x)g, h\rangle = 0$. From this, $g = K^{\pm}_{k-1}(x,s)$ where the kernel $K^{\pm}_{k-1}$ is taken with respect to the distribution $d\mu^{(\pm)}(x) = (1+x)(1-x)d\mu(x)$. The corresponding orthogonal polynomials $p_i^{\pm}(x)$ are also easily found: up to normalization they are equal

$$p_i^{\pm}(x) = K_i(x,-1)p_{i+1}(1) - K_i(x,1)p_{i+1}(-1).$$

Then $f_n^{\pm}(x) = c(x-s)(x+1)(K^{\pm}_{k-1}(x,s))^2$.

Remarks.

1. The polynomials $f_n^-, f_n^{\pm}$ were constructed and applied to coding theory in Levenshtein (1978). Polynomials closely related to them were studied in a more general context in the works of M. G. Krein *et al.*; see Krein and Nudelman (1974). The orthogonal systems $\{p_i^-\}, \{p_i^{\pm}\}$ are sometimes called *adjacent polynomials* of the original system $\{p_i\}$.

2. The above arguments do not imply optimality in Delsarte's problem of the polynomials constructed since the second differential of the functionals $\mathcal{F}, \mathcal{F}^{(-)}, \mathcal{F}^{(\pm)}$ above is undefined (e.g., $d^2\mathcal{F}(g) = 2\langle (x-s)h, h\rangle$). Optimality in V_n of the polynomials $f_n^-, f_n^{\pm}$ as well as the range of n, s in which conditions (2) hold were established by Sidelnikov and Levenshtein in the papers cited.

3. Asymptotic bounds derived from (1) relying upon the polynomials $f_n, f_n^-, f_n^{\pm}$ coincide. For the finite values of the parameters, better bounds are obtained from $f_n^-, f_n^{\pm}$.

3. Spectral method. The ideas for constructing polynomials in the problem (1)-(2) discussed below originate in the work of Bachoc (2006). They were elaborated upon in a previous work of the authors.

We develop the remark made after (4), namely that for any $i \ge 1$, $K_k(x, x_{k,i})$ is an eigenfunction of the operator X_k. Since $K_k(x,s)$ is a good choice for the polynomial in Delsarte's problem, it is possible to base the choice of the polynomial on the eigenvectors of X_k as opposed to the analytic arguments discussed above. In particular, $K_k(x,s)$ arises as an eigenfunction of the operator T_k that acts on V_k as $\phi \mapsto X_k\phi - c\hat{\phi}(k)p_k$ for some constant c. A minor difficulty arises in proving conditions (2) for the polynomials thus chosen since the coefficients $\{a_i, i = 0,1,\dots\}$ are negative for some orthogonal systems (for instance, for the Krawtchouk polynomials). It is resolved by replacing X_k with the operator $S_k = E_k \circ p_1$ since then the coefficients in the three-term expansion of p_1p_i are always nonnegative.

We note that the argument about the eigenfunctions does not depend on the choice of the L_2 space; in particular, the kernels $K_k^-, K_k^{\pm}$ arise if the operator X_k is written with respect to the basis of the corresponding adjacent polynomials ($\{p_i^-\}$ or $\{p_i^{\pm}\}$) and their generating distribution.

To further discuss this method, we first observe that for the basis $\{p_i\}$ and distribution $d\mu$, it leads to the estimate

$$M \le \frac{4a_kp_{k+1}(1)p_k(1)}{P_1(1) - \lambda_{\max}(S_k)} \tag{7}$$

(Barg and Nogin (2006)), which yields bounds very close to those obtained relying on the MRRW polynomial. By the remark in the previous paragraph, an appropriate modification of the spectral method can be used to derive Levenshtein-type bounds as well. For instance, for the Krawtchouk system (the Hamming space $\mathcal{H}_n$), adjacent polynomials p_i^- are simply the Krawtchouk polynomials associated with the Hamming space $\mathcal{H}_{n-1}$, so the operator X_k is computed from the recurrence (3) with $a_i = {}^1\!/_2\sqrt{(n-1-i)(i+1)}, i = 0,1,\dots,k-1; b_i = n/2, i = 0,1,\dots,k.$

The method outlined above has two advantages. First, it enables one to obtain simple estimates of the largest eigenvalue of S_k which is important in verifying the condition $f(x) \le 0, x \in [-1,s]$. The second advantage is a more substantial one: this method can be extended to the case of *multivariate zonal polynomials* when the analytic alternative is not readily available. This case arises when the space X is homogeneous but not 2-point homogeneous. Worked examples include the real Grassmann manifold $G_{k,n}$ (Bachoc (2006); the P_i are given by the generalized k-variate Jacobi polynomials) and the so-called ordered Hamming space (Barg and Purkayastha (2007)). We provide a few more details on the latter case in order to illustrate the general method.

Let $\mathcal{Q}$ be a finite alphabet of size q. Consider the set $\mathcal{Q}^{r,n}$ of vectors of dimension rn over $\mathcal{Q}$. A vector $\boldsymbol{x}$ will be written as a concatenation of n blocks of length r each, $\boldsymbol{x} = \{x_{11}, \dots, x_{1r}; \dots; x_{n1}, \dots, x_{nr}\}$.

For a given vector $\boldsymbol{x}$ let $e_i, i = 1, \dots, r$ be the number of r-blocks of $\boldsymbol{x}$ whose rightmost nonzero entry is in the ith position counting from the beginning of the block. The r-vector $e = (e_1, \dots, e_r)$ will be called the *shape* of $\boldsymbol{x}$. A shape vector $e = (e_1, \dots, e_r)$ defines a partition of a number $N \le n$ into a sum of r parts. Let $e_0 = n - \sum_i e_i$. Let $\Delta_{r,n} = \{e \in (\mathbb{Z}_+ \cup \{0\})^r : \sum_i e_i \le n\}$ be the set of all such partitions. The zonal polynomials associated to $\mathcal{Q}^{r,n}$ are r-variate polynomials $K_f(e), f, e \in \Delta_{r,n}$ of degree $\kappa = \sum_i f_i$. They are orthogonal on $\Delta_{r,n}$ according to the following inner product

$$\sum_{e \in \Delta_{r,n}} K_f(e) K_g(e) w(e) = 0 \quad (f \ne g).$$

The weight in this relation is given by the multinomial probability distribution

$$w(e_1, \dots, e_r) = n! \prod_{i=0}^{r} \frac{p_i^{e_i}}{e_i!} \qquad (p_i = q^{i-r-1}(q-1), i = 1, \dots, r; p_0 = q^{-r}),$$

so the polynomials $K_f(e)$ form a particular case of *r-variate Krawtchouk polynomials*.

Let $\boldsymbol{x} \in \mathcal{Q}^{r,n}$ be a vector of shape e. Define a weight function on $\mathcal{Q}^{r,n}$ by setting $\mathrm{w}(\boldsymbol{x}) = \sum_i ie_i$ and let $d_r(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{w}(\boldsymbol{x} - \boldsymbol{y})$ be the ordered Hamming metric (known also as the Niederreiter-Rosenbloom-Tsfasman metric). We note that in the multivariate case there is no direct link between the variables and the metric. For instance, for the space $\mathcal{Q}^{r,n}$ the polynomials (as well as relations in the corresponding association scheme) are naturally indexed by shape vectors e while the weight is some function e.

The Delsarte theorem in this case takes the following form: *The size of an (n, M, d) code $C \subset \mathcal{Q}^{r,n}$ is bounded above by $M \le \inf_{f \in \Phi} f(0)/f_0$*, where

$$\Phi = \{f(x) = f(x_1, \dots, x_r) = f_0 + \sum_{e \ne 0} f_e K_e(x) : f_0 > 0, f_e \ge 0 (e \ne 0); f(e) \le 0 \; \forall e \text{ s.t. } \sum_{i=1}^{r} ie_i \le d\}$$

The argument for the univariate case given in this section can be repeated once we establish a three-term relation for the polynomials $K_f(e)$. Let $\mathbb{K}_\kappa$ be the column vector of the normalized polynomials K_f ordered lexicographically with respect to all f that satisfy $\sum_i f_i = \kappa$ and let $P(e)$ be a suitably chosen linear polynomial. Then

$$P(e)\mathbb{K}_\kappa(e) = A_\kappa \mathbb{K}_{\kappa+1}(e) + B_\kappa \mathbb{K}_\kappa(e) + A_{\kappa-1}^T \mathbb{K}_{\kappa-1}(e)$$

where A_κ, B_κ are matrices of order $\binom{\kappa+r-1}{r-1} \times \binom{\kappa+s+r-1}{r-1}$ and $s = 1, 0$, respectively. The elements of these matrices can be computed explicitly from combinatorial considerations. This gives an explicit form of the operator $S_\kappa = E_\kappa \circ P(e)$ in the orthonormal basis. Relying on this, it is possible to derive a bound on codes in the NRT space of the form (7) and perform explicit calculations, both in the case of finite parameters and for asymptotics. The full details of the calculations are given in work cited above.

References

Levenshtein (1978-1998): V. I. Levenshtein, VIIth Conference on Coding Theory, Vilnius (1978); Problemy Kibernetiki **40** (1983), pp. 43-110; Acta Applicandae Mathematicae **29** (1992), 1–82; Handbook of Coding Theory vol. 1 (1998), pp. 499–648.

Sidelnikov (1980): V. M. Sidelnikov, Extremal polynomials used in bounding the code volume, Probl. Inform. Trans. **16** 3 (1980), 174–186.

Krein and Nudelman (1974): M. G. Krein and A. A. Nudelman, The Markov moment problem and extremal problems, Moscow, 1974; English translation: American Mathematical Society, Providence, R.I., 1977.

Bachoc (2006): C. Bachoc, Linear programming bounds for codes in Grassmannian spaces, IEEE Trans. Inform. Theory, **52** 5 (2006), 2111–2125.

Barg and Nogin (2006): A. Barg and D. Nogin, Spectral approach to linear programming bounds on codes, Probl. Inform. Trans., **42** 2 (2006), 77–89.

Barg and Purkayastha (2007): A. Barg and P. Purkayastha, Bounds on ordered codes and orthogonal arrays, arXiv:cs/0702033.