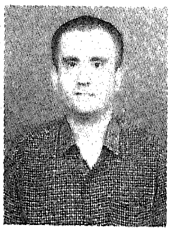
How to Stay Away from Each Other in a Spherical Universe

2. Orthogonal Polynomials and Spherical Codes

Bhaskar Bagchi



Bhaskar Bagchi is with the Indian Statistical Institute since 1971, first as a student and then as a member of the faculty. He is interested in diverse areas of mathematics like combinatorics, elementary and analytic number theory, functional analysis, combinatorial topology and statistics.

The first part of this general article appeared in September 1997.

In the first part we discussed Tammes' problem and its higher dimensional analogues. In this part we shall see how harmonic analysis and the theory of orthogonal polynomials can be used to settle some of the most interesting instances of this problem.

Inner Products and Gram Matrices

Recall that the inner product $\langle x, y\rangle$ between the points x and y of $I\!R^d$ is the number $\sum_{i=1}^{d} x_i y_i$. This useful notion combines in a single formula the notion of distance and angle. In fact, the distance $||x - y||$ between x and y is given by $||x - y||^2 = \langle x - y, x - y\rangle$, while the cosine of the angle subtended at the origin by x and y is $\langle x, y\rangle/(||x|| \cdot ||y||)$. In particular, the distance between two points x and y in a spherical code is $\sqrt{2 - 2\langle x, y\rangle}$. It follows that maximising the minimum distance is the same as minimising the maximum inner product (between distinct points of the code). For any spherical code X, we shall use $I(X)$ to denote the set $\{\langle x, y\rangle : x \neq y,\ x, y \in X\}$. Thus an optimal code minimises the maximum of $I(X)$ over all spherical codes X of the same size and rank. Let G denote the Gram matrix of X. By definition, it is the $n \times n$ matrix $G = (\langle x, y\rangle : x, y \in X)$. From standard linear algebra, one knows that G is a correlation matrix (i.e., a non negative definite matrix all whose diagonal entries equal 1) of rank $\leq d$ which determines X uniquely (upto rotation). Thus, our problem may also be phrased as: minimise the maximum off diagonal entry $\mu(G)$ of G as G runs over all $n \times n$ (real) correlation matrices of rank not exceeding d.

Optimality of the Simplex, Cross Polytope, and the in between Sizes.

We now present some elementary proofs of these results, involving nothing more than a bit of linear algebra. The regular simplex X is a code of size $d+1$ and rank d with $I(X) = \{-1/d\}$. To prove its optimality and uniqueness, it suffices to show that it is the only code Y with these parameters such that $\max I(Y) \leq -1/d$, that is, $I(Y) \subseteq [-1, -1/d]$. Define the polynomial function $f : [-1, 1] \longrightarrow I\!R$ by $f(t) = t + 1/d$. Note that $f \leq 0$ on $[-1, -1/d]$. In particular, $f(\langle y_1, y_2 \rangle) \leq 0$ for $y_1, y_2 \in Y$, $y_1 \neq y_2$. On the other hand, we have $\sum\{f(\langle y_1, y_2 \rangle) : y_1, y_2 \in Y, y_1 \neq y_2\} = (d+1) + \sum\{\langle y_1, y_2 \rangle : y_1, y_2 \in Y, y_1 \neq y_2\} = \sum\{\langle y_1, y_2 \rangle : y_1, y_2 \in Y\} = ||\sum\{y : y \in Y\}||^2 \geq 0$. Thus we have a non negative sum of non positive terms, so that the terms $f(\langle y_1, y_2 \rangle)$ must equal zero for all $y_1 \neq y_2$ in Y. Thus $I(Y) = \{-1/d\}$, so that the $d+1$ points in Y are equidistant and hence Y is the regular simplex.

Next let $d+1 < n \leq 2d$. If Y is a spherical code with these parameters, then we shall show that its minimum distance is $\leq \sqrt{2}$, or, equivalently, its maximum inner product is ≥ 0. Since the cross polytope has minimum distance $= \sqrt{2}$, any of its subsets has minimum distance $\geq \sqrt{2}$, so that we shall hereby establish the optimality of any subset of size $n > d+1$ of the (set of vertices of the) cross polytope.

So let Y be a spherical code of rank d and size n in the range indicated. Suppose, if possible, that its maximum inner product is strictly negative. That is, the off diagonal entries of its Gram matrix G are strictly negative. Put $A = I - G$, where I is the $n \times n$ identity matrix. Then A is a square matrix with zero diagonal entries and strictly positive off diagonal entries. By Peron Frobenius theory, the largest eigen value of such a matrix is simple (i.e., has geometric multiplicity $= 1$). But G is nonnegative definite with rank $\leq d$ and so nullity $\geq n - d$. Therefore the largest eigen value of A $(= 1)$ has multiplicity $\geq n - d$. So we must have $n - d \leq 1$, i.e., $n \leq d + 1$, contrary to our assumption.

While this argument shows that, in particular, the cross polytope is an optimal spherical code of size $2d$ and rank d, it does not prove the uniqueness of the optimal code in this case. To prove the uniqueness, let Y be any spherical code with these parameters such that $I(Y) \subseteq [-1,0]$. We have to show that Y must be the cross polytope. We have already noted that $S_1 := \sum\{\langle y_1, y_2\rangle : y_1, y_2 \in Y\} \geq 0$. Let us also note that $S_2 := \sum\{(\langle y_1, y_2\rangle^2 - 1/d) : y_1, y_2 \in Y\} \geq 0$. Indeed, letting λ_j, $1 \leq j \leq d$, denote the (possibly) non zero eigenvalues of the Gram matrix G of Y, we find that S_2 equals $\text{trace}(G^2) - \frac{1}{d}(\text{trace}(G))^2 = \sum_{i=1}^{d}(\lambda_i - 2)^2 \geq 0$. (Since the sum of the d eigenvalues in question is $\text{trace}(G) = 2d$, their average is 2.) Now let f be the polynomial function on $[-1,1]$ given by $f(t) = (t+1)t$. Thus $f \leq 0$ on $I(Y)$. On the other hand, we have $\sum\{f(\langle y_1, y_2\rangle) : y_1, y_2 \in Y, y_1 \neq y_2\} = S_1 + S_2 \geq 0$. Therefore f vanishes on $I(Y)$, so that $I(Y) \subseteq \{-1, 0\}$. Thus any two of the points in Y are either orthogonal or they are negatives of each other. Since Y has size $2d$, it follows that Y consists of the elements of an orthonormal basis and their negatives. Thus Y is the set of vertices of a cross polytope.

Gegenbauer Polynomials and Optimality of the Icosahedron.

Let us say that a polynomial function g on $[-1,1]$ is *positive* of rank d if for every spherical code Y of rank d (and whatever size), we have $\sum\{g(\langle y_1, y_2\rangle) : y_1, y_2 \in Y\} \geq 0$. Clearly, the proofs of optimality and uniqueness given above succeeded largely because of our ability to locate appropriate positive polynomials. One natural class of positive polynomials are the Gegenbauer polynomials Q_k, $k = 0, 1, 2, \ldots$. These are recursively given by the formulae

$$\begin{aligned} Q_0(t) &= 1, \\ Q_1(t) &= dt, \\ tQ_k(t) &= (1 - a_{k-1})Q_{k-1}(t) + a_{k+1}Q_{k+1}(t), \end{aligned}$$

for $k = 1, 2, \ldots$. Here $a_0 = 0$ and $a_k = k/(d + 2k - 2)$ for $k = 1, 2, \ldots$.

Let $d\omega$ denote integration with respect to the unique rotation invariant Borel probability ω on the sphere S^{d-1}. Then we have the formula

$$\int Q_k(\langle x,y\rangle)Q_k(\langle y,z\rangle)\,d\omega(y) = Q_k(\langle x,z\rangle).$$

(Of course, this strange formula did not come from thin air. Its source is harmonic analysis on the sphere. If the vector space Harm(k) of homogeneous harmonic polynomials of degree k on the sphere is equipped with the L^2 inner product, then we get a functional Hilbert space whose reproducing kernel is $(x,y) \mapsto Q_k(\langle x,y\rangle)$. The integral formula pops out of this theory.) Using this formula, we find that for any spherical code Y of rank d,

$$\sum\{Q_k(\langle y_1,y_2\rangle):\ y_1,y_2 \in Y\} =$$

$$\int (\sum\{Q_k(\langle x,y\rangle):\ y \in Y\})^2 d\omega(x).$$

This shows that the Gegenbauer polynomials are indeed positive of rank d.

With this tool in hand, we are now ready to prove that the icosahedron is the unique optimal spherical code of size 12 and rank 3. Using the coordinates presented in the section on the platonic solids in Part 1, one sees that for the icosahedron X, we have $I(X) = \{-1, \pm 1/\sqrt{5}\}$. In particular, the maximum inner product of this code is $1/\sqrt{5}$. Let f be the polynomial $f(t) = (t+1)(t+1/\sqrt{5})^2(t-1/\sqrt{5})$. Then f vanishes precisely on $I(X)$ and it is non positive on $I(Y)$ for any spherical code Y (with the same parameters as X) which is at least as good as X. But, in terms of the Gegenbauer polynomials (with $d=3$) f has the expansion $f = \sum_{j=0}^{4} c_j Q_j$ where a calculation shows that $c_0 = f(1)/12$ and the remaining three coefficients c_j are strictly positive (Verify this!). Hence the sum $\sum\{f(\langle y_1,y_2\rangle):\ y_1,y_2 \in Y,\ y_1 \neq y_2\}$ equals $-12f(1) + \sum\{f(\langle y_1,y_2\rangle):\ y_1,y_2 \in Y\} = -12f(1) + 144c_0 + \sum_{j=1}^{4} c_j \sum\{Q_j(\langle y_1,y_2\rangle):\ y_1,y_2 \in Y\} = \sum_{j=1}^{4} c_j \sum\{Q_j(\langle y_1,y_2\rangle):\ y_1,y_2 \in Y\} \geq 0$. (In the second step, we have used the fact that Q_0 is the constant function

Box 1. Invariant Probability on the Sphere.

On S^2, this is nothing but surface area, normalised to make total area of the unit sphere to be one. The usual coordinate transformation from cartesian to polar coordinates shows that, integration with respect to the invariant probability ω on S^{d-1} is given (for any 'nice' function f on S^{d-1}) by the formula

$$\int f \, d\omega = c \cdot \int_0^{\pi} \int_0^{2\pi} \cdots \int_0^{2\pi} \sin\theta_2 \sin^2\theta_3 \cdots \sin^{d-2}\theta_{d-1} \, f(x_1, \cdots, x_d) \, d\theta_1 \cdots d\theta_{d-1},$$

where $x_j = \cos\theta_{j-1} \cdot \prod_{i=j}^{d-1} \sin\theta_i$ for $1 \le j \le d$ and c is the constant (depending only on d) chosen to make $\int 1 \, d\omega$ equal 1. (In this formula $\theta_0 = 0$ and empty products stand for 1.)

1. The inequality in the end follows from the positivity of the Q_j's.) Therefore, as before f vanishes on $I(Y)$ so that $I(Y) \subseteq I(X)$. Also, equality in the last inequality forces $\sum\{Q_j(\langle y_1, y_2\rangle) : y_1, y_2 \in Y\} = 0$ for $1 \le j \le 4$. In particular, for $j = 1$ we get $\|\sum\{y : y \in Y\}\|^2 = \frac{1}{d}\sum\{Q_1(\langle y_1, y_2\rangle) : y_1, y_2 \in Y\} = 0$. Hence $\sum\{y : y \in Y\} = 0$. Taking inner product with any fixed $y_0 \in Y$, we get $\sum\{\langle y, y_0\rangle : y \in Y, y \ne y_0\} = -1$. Thus, if a, b, c denote the number of y in Y such that $\langle y, y_0\rangle = -1, +1/\sqrt{5}$ and $-1/\sqrt{5}$ respectively, then $a+b+c = 11$ and $a\cdot(-1)+b\cdot(1/\sqrt{5})+c\cdot(-1/\sqrt{5}) = -1$. Hence $a = 1$, $b = c = 5$. This shows that the antipode (or negative) of each point of Y is again in Y, and for any two points of Y which are not antipodes, the inner product is $\pm 1/\sqrt{5}$. It is easy to conclude from here that Y must be the icosahedron.

Tight and Quasi-Tight Spherical Codes

A spherical code Y is called a (spherical) t-design (for a non negative integer t) if for all polynomial functions f (on the sphere) of degree $\le t$, the average value of f over Y equals the surface integral (with respect to ω) of f over the sphere. This notion was introduced by P Delsarte, J M Goethals and J J Seidel in a famous paper in the sixth volume of *Geometriae Dedicata* (1977). The *strength* σ of Y is defined to be the largest t for which Y is a t-design. The usefulness of spherical codes of large strength in numerical integration is pretty obvious. In their paper, Delsarte and others characterised σ as the largest integer for which $\sum\{Q_j(\langle y_1, y_2\rangle : y_1, y_2 \in Y\} = 0$ for $1 \leq j \leq \sigma$. They also proved the inequality $\sigma \leq 2\delta$ where the *degree* δ of Y is defined to be the size of the set $I(Y)$. A spherical code is called *tight* if $\sigma = 2\delta$ and it is *quasi-tight* if $\sigma = 2\delta - 1$. Generalising the arguments presented above, it can be shown that: *Any tight or quasi-tight spherical code is optimal. Further, when a tight or quasi-tight code of a given size and rank exists, all the optimal codes in that case are tight or quasi-tight.* The proof exploits the connection with spherical harmonics and some deep properties of the Gegenbauer polynomials. The details will appear later (hopefully) in an article.

Examples

The regular polygons of odd size are the only tight codes of rank $d = 2$. E Bannai and R M Damerell proved in 1979 that all the tight codes of higher rank have degree $\delta \leq 2$. The regular simplexes (one for each rank) are the only tight codes of degree 1. Thus all the remaining tight codes must have $d \geq 3$, $\delta = 2$. By a formula of Delsarte and others, the size n of such a spherical code is given by $n = d(d+3)/2$. Only two such tight spherical codes are known. These are the Schlafle polytope with parameters $(n, d, \delta) = (27, 6, 2)$ and the tight code with parameters $(275, 22, 2)$. The latter admits the sporadic simple group of McLaughlin as its automorphism group.

The regular polygons of even size are the only quasi-tight spherical codes of rank two. The cross polytopes are quasi-tight of degree 2, while the icosahedron is quasi-tight of degree 3. The E_8 root system and the Leech code are quasi-tight of degree 4 and 5 respectively. There are two quasi-tight codes with parameters $(n, d, \delta) = (56, 7, 3)$ & $(552, 23, 3)$ which are, in some definite sense, double covers of the Schlafle code and the Mclaughlin code, respectively. There is one with parameters $(4600, 23, 4)$, which may be viewed as a 'contraction' of the Leech code. It admits the second sporadic simple group of Conway as an automorphism group. All the quasi-tight codes mentioned here are antipodal (i.e. with each point, the antipode also occurs; note that any antipodal code has odd strength, so that it could not be tight). Bannai and Damerell proved that, with the sole exception of the Leech code (which has $\delta = 6$), all antipodal quasi-tight codes of rank $d \geq 3$ must have degree $\delta \leq 4$. The size n of a quasi-tight antipodal code is given by $n = 2\binom{d+\delta-2}{\delta-1}$.

Now we pass to examples of quasi tight spherical codes which are not antipodal. We know only one such code of degree $\delta \geq 3$; it has parameters $(891, 22, 3)$. There is an infinite series of quasi-tight codes of degree 2 with parameters $n = (q+1)(q^3+1)$, $d = q(q^2-q+1)$ for any prime power $q \geq 3$ - these are intimately related to the extremal generalised quadrangles. Apart from these, we only have four sporadic examples with $\delta = 2$. They have parameters $(n, d) = (16, 5)$, $(100, 22)$, $(162, 21)$. The second of these admits the sporadic simple group of Higman and Sims as automorphism group.

This completes the list of the known tight and quasi tight spherical codes. Further details on these examples may be found in the paper by Delsarte and others. The theorem quoted above applies to all these examples. In fact, with the exception of the larger members of the infinite series, all of them yield unique optimal codes. However, the theorem does not apply to the 600-cell (though it is uniquely optimal) since its degree is 8 and the strength is 'only' 11.

To conclude, we should mention that the set $I(X)$ of inner products of a tight or quasi tight spherical code X of rank $\geq$ 3 is determined by the parameters n, d and δ, as follows. For $k = 0, 1, 2, \ldots$, define the polynomial R_k by $R_k = \sum_{j=0}^{k} Q_j$. Then, $I(X)$ is the zero set of the polynomial Ψ, where $\Psi = R_\delta$ if X is tight, and if X is quasi tight then

$$\Psi = t \cdot R_{\delta-1} + (1-t) \cdot R_\delta,$$

where

$$t = \frac{R_\delta(1) - n}{R_\delta(1) - R_{\delta-1}(1)}.$$

Using this formula, the reader may easily compute the minimum distances of the optimal spherical codes mentioned above.

Suggested Reading

- H S M Coxeter. *Regular Polytopes*. Dover. New York, 1973.
- J H Conway and N J A Sloane. *Sphere Packings, Lattices and Groups*. Springer Verlag. New York, 1988.

Address for Correspondence
Bhaskar Bagchi
Statistics & Mathematics Unit
Indian Statistical Institute
R.V. College Post
Bangalore 560 059, India
bbagchi@isibang.ernet.in

Here it is in Genesis: 'He took one of Adam's ribs, and made the rib into a woman.' Cloning if ever I heard it.

From: Gene Antics