THE NUCLEUS OF A SET

Gilbert Strang (*)

(received July 28, 1967)

Consider the subset $\mathcal{K} \subseteq \mathcal{C}$ [0,1] containing those functions for which

(1) $\quad f(0) = 0, \quad |f(x) - f(y)| \leq |x - y| \quad \text{for} \quad 0 \leq x, y \leq 1$

One never attempts to visualize $\mathcal{K}$; it is just a compact blur in the infinite-dimensional space $\mathcal{C}$. Nevertheless, we want to establish that it shares with several other sets an odd but rather remarkable "geometric" property: it is overwhelmingly concentrated around a single element. This element we call the nucleus of $\mathcal{K}$.

To give a precise definition of the nucleus, we adopt Kolmogorov's measure $N_\epsilon(\mathcal{A})$ of the size of an arbitrary compact set $\mathcal{A}$ in a metric space. N_ϵ is the minimal number of subsets of diameter $\leq 2\epsilon$ required to cover $\mathcal{A}$; its logarithm is called the ϵ-entropy of $\mathcal{A}$. For p to be the nucleus of $\mathcal{A}$, we require that every closed sphere S(p, r) around p, of arbitrarily small radius r, should contain more than half the set $\mathcal{A}$. Thus we make the following

DEFINITION: The element p is the nucleus of a compact set $\mathcal{A}$ if for every $r > 0$ there exists an $\epsilon(r) > 0$ such that

(2) $$N_\epsilon(\mathcal{A} \cap S(p, r)) > \frac{1}{2} N_\epsilon(\mathcal{A}) \quad \text{for} \quad \epsilon < \epsilon(r).$$

*This research was partially supported by the Sloan Foundation, the Office of Naval Research, and the National Science Foundation (GP 7477).

Obviously there cannot be two nuclei p and p': if r is small enough to make the spheres $S(p, r)$ and $S(p', r)$ disjoint, then (2) cannot hold for both spheres.

A simple example of a set of reals having a nucleus is a convergent sequence together with its limit, say the set

$$\mathcal{A} = \{0, 1, 1/2, 1/3, \dots\}.$$

The limit point $p = 0$ is the nucleus, since for any sphere (interval) $S(0, r)$ we have

$$\left.\begin{array}{l} N_\epsilon(\mathcal{A} \cap S) \to \infty \\ N_\epsilon(\mathcal{A} - S) \leq 1/r \end{array}\right. \quad \text{as} \quad \epsilon \to 0.$$

The nucleus is not, however, just an arbitrary limit point in disguise; every point of $\mathcal{K}$ is a limit point, but there is a unique nucleus. The nucleus may well depend on the choice of norm; if $\mathcal{A}$ is made up of two convergent sequences, the nucleus can become either (or neither) of the corresponding limit points, by an alteration in the norm.

Our definition may be of value in certain improperly posed problems, such as the problem of interpolation:

Find $f \in \mathcal{C}[0, 1]$, given $f(x_i) = y_i$, $1 \leq i \leq N$.

Of course this problem has a unique solution in case f is known to be a polynomial of degree $N - 1$, or to belong to some other manifold of a suitable dimension. If one is given only qualitative information, such as the condition $f \in \mathcal{K}$, there may be infinitely many candidates f. When this set of candidates has a nucleus, that seems to us a reasonable choice for the interpolating f. It is interesting that under certain natural assumptions - a bound on some derivative $f^{(n)}$, for example - the nucleus appears to be one of the spline functions now in fashion.

Naturally there are interesting compact sets which do not possess a nucleus, such as the Hilbert cube. Furthermore, there are sets which almost certainly have a nucleus, but for which the computations of N_ϵ are forbidding; alternative definitions which require less counting would be happily entertained. Estimates of N_ϵ have been carried out for a variety of function classes, with sufficient accuracy to distinguish the smoothness or the number of variables involved in the construction of $\mathcal{A}$. Most of these estimates are Russian, ingenious, and asymptotic; they provide the leading term in N_ϵ as $\epsilon \to 0$, which is often insufficient to verify the condition (2) for a nucleus.

We conjecture, for example, that 0 is the nucleus of the ellipse in Hilbert space defined by

$$\sum_{1}^{\infty} \frac{x_i^2}{\lambda_i^2} \leq M^2, \qquad \lambda_i \to 0.$$

For this particular compact set a decent proof should be possible

In this note we are concerned first of all with the nucleus of $\mathcal{K}$. At the same time, however, we calculate by heuristic arguments the nucleus of every subset of $\mathcal{K}$ which is defined by imposing a finite number of linear constraints. The constraints in the interpolation problem are the simplest possible; a more exciting constraint is

$$\int_0^1 f(x)dx = \alpha, \quad \alpha < \frac{1}{2}.$$

In this case the nucleus is almost certainly the function

$$f(x) = \log \frac{\cosh \lambda}{\cosh \lambda(1-x)},$$

where λ is chosen to satisfy the constraint.

The problem of nuclei in subsets of $\mathcal{K}$ has links with the theory of random walk, and with Brownian motion. It might be appropriate to call it "Bernoulli walk", since it corresponds to steps of size $\Delta f = \pm \Delta x$, rather than the steps of order $\Delta x^{1/2}$ in Brownian motion. (This explains why the walk has to be conditioned by constraints such as $\int f = \alpha$, in order to make something non-trivial happen in the limiting walk, $\Delta x \to 0$). This question will be developed at greater length by Dr. Jay Israel, who is able to replace our heuristic arguments by rigorous ones, at least for certain subsets.

There remain many interesting classes - of Hölder continuous functions, analytic functions, functions of several variables, and so on - for which the existence of nuclei remains to be examined.

2. Kolmogorov [1] has pointed out the simple structure of the set $\mathcal{K}$; our first step towards finding the nucleus is to reproduce his calculation of N_ϵ. Recall that the metric in $\mathcal{C}$ (and $\mathcal{K}$) is the sup norm

$$||f - g|| = \sup_{0 \leq x \leq 1} |f(x) - g(x)|.$$

Let the integer M be defined by $M < 1/\epsilon \leq M + 1$. Then the result we want is $N_\epsilon(\mathcal{K}) = 2^M$.

Set $\epsilon' = 1/(M+1)$, and consider those functions g which are linear between the nodes $j\epsilon'$ and satisfy

(3) $$g(0) = 0 = g(\epsilon'), \quad g((j+1)\epsilon') = g(j\epsilon') \pm \epsilon'$$

for $1 \leq j \leq M$. It is not hard to verify that every element f of $\mathcal{K}$ lies within $\epsilon' \leq \epsilon$ of one of these piecewise-linear functions g. Briefly, if $|f(x) - g_0(x)| \leq \epsilon'$ for $x \leq j\epsilon'$, then this inequality continues to hold up to the next node $(j+1)\epsilon'$ for at least one of the two extensions of g_0. Thus $N_\epsilon(\mathcal{K}) \leq 2^M$, since $\mathcal{K}$ is covered by the ϵ-spheres around the 2^M functions g. (Corollary: $\mathcal{K}$ is compact).

Now let $\epsilon'' = 1/M$, and consider another collection of 2^M piecewise-linear functions, this time satisfying

(4) $$h(0) = 0, \quad h((j+1)\epsilon'') = h(j\epsilon'') \pm \epsilon'', \quad 0 \leq j \leq M - 1.$$

Each pair of these functions is separated by $2\epsilon'' > 2\epsilon$ at the node of first divergence. They all belong to $\mathcal{K}$, and must be in different subsets whenever $\mathcal{K}$ is covered by sets of diameter 2ϵ. Therefore N_ϵ is not less than 2^M, and must be exactly 2^M.

Now we claim that the function $f \equiv 0$ is the nucleus of $\mathcal{K}$. For this we need a second device due to Kolmogorov, in order to count those of the functions h which lie <u>outside</u> the sphere $S(0, r)$. For such an h, let $j\epsilon''$ be the last node at which $|h(j\epsilon'')| > r$. Then for $x \geq j\epsilon''$, reflect the graph of h across the horizontal line through $h(j\epsilon'')$; the result h' is still one of the piecewise-linear functions constructed by (4), and it has the property that $|h'(1)| > r$. Furthermore, this map is never more than two to one; h' is the image of itself and at most one other h. Therefore

$$N_\epsilon(\mathcal{K} \cap S(0, r)) \geq 2^M - 2L,$$

where L is the number of our functions h for which $|h(1)| > r$.

This number L is easily estimated. If $h(1) = s\epsilon''$, then h must ascend $(M+s)/2$ times and descend $(M-s)/2$ times. The number of such functions is given by the binomial coefficient $\binom{M}{(M+s)/2}$. Therefore

$$L = \sum_{\substack{|s| > Mr \\ M+s \text{ even}}} \binom{M}{(M+s)/2}.$$

Detailed estimates of such sums are well-known; we use only the fact that $L = o(2^M)$ as $M \to \infty$. Thus

$$\frac{N_\epsilon(\mathcal{K} \cap S(0, r))}{N_\epsilon(\mathcal{K})} \to 1 \text{ as } \epsilon \to 0,$$

and the zero function is indeed the nucleus of $\mathcal{K}$.

3. In this section we consider the subsets of $\mathcal{K}$ formed by imposing linear constraints on the elements f. The zero function will violate the constraints, in general; the question is whether these subsets have a nucleus, and how to find it.

We start by dividing the interval [0, 1] into n equal sub-intervals. We retain the convention $\epsilon'' = 1/M$, and assume that $M/n = p$ is an integer. Then those h-functions for which $h(1/n) = s\epsilon''$ must ascend $(p+s)/2$ times and descend $(p-s)/2$ times in the first subinterval $0 \leq x \leq 1/n$. Let us introduce

$w = s/p$ for the slope of the chord between the initial point $(0, 0)$ and the common point $(1/n, s\epsilon'')$, and apply Stirling's formula to

(6) $$\binom{p}{\frac{p+s}{2}} = \frac{p!}{(\frac{p+s}{2})!\,(\frac{p-s}{2})!}$$

$$\sim \frac{\sqrt{2\pi p}\; p^p\, e^{-p}}{\pi\sqrt{p^2 - s^2}\;(\frac{p+s}{2})^{(p+s)/2}\;(\frac{p-s}{2})^{(p-s)/2}\; e^{-p}}$$

$$= \left(\frac{2}{\pi p(1-w)^2}\right)^{1/2}\left[\left(\frac{1+w}{2}\right)^{(1+w)/2}\left(\frac{1-w}{2}\right)^{(1-w)/2}\right]^{-p}$$

$$= Q(w,\ p), \text{ say}$$

Now consider all the h-functions for which

(7) $$h(\tfrac{j+1}{n}) - h(\tfrac{j}{n}) = \frac{w_j}{n}, \quad 1 \le j \le n;$$

that is, the chords over successive sub-intervals have slopes $w_1, \ldots, w_n$. Since (6) counts the number of distinct possibilities in a sub-interval, the total number of h-functions satisfying (7) is approximately

$$C = \prod_1^n Q(w_j, p).$$

Note that the leading term in log C is proportional to

(8) $$\sum_1^n \left[\frac{1+w_j}{2} \log \frac{1+w_j}{2} + \frac{1-w_j}{2} \log \frac{1-w_j}{2}\right].$$

As $n \to \infty$, the corresponding functional approaches

$$I(f') = \int_0^1 \left(\frac{1+f'}{2} \log \frac{1+f'}{2} + \frac{1-f'}{2} \log \frac{1-f'}{2}\right) dx.$$

Now come the heuristics. Let $\mathcal{A}$ be a (non-empty) subset of $\mathcal{K}$, defined by a finite number of continuous linear constraints

(9) $$\int_0^1 f(x)d\mu_i(x) = \alpha_i \quad .$$

We conjecture that there is a unique $f \in \mathcal{A}$ which minimizes $I(f')$, and that this minimizing function is the nucleus of $\mathcal{A}$. In short, we believe $I(f')$ to be decisive in estimating the density of the sets $\mathcal{A}$ near the element f.

Let us illustrate the computation of this element by considering the constraint

$$\int_0^1 f(x)dx = \alpha, \quad \alpha < 1/2 \quad .$$

An integration by parts converts this into

(10) $$\int_0^1 (1-x)f'(x)dx = \alpha.$$

Introducing a Lagrange multiplier λ, we want the minimum of

$$I(f') - \lambda(\int (1-x)f'(x)dx - \alpha).$$

Differentiating formally with respect to $f'(x)$,

(11) $$\frac{1}{2} \log \frac{1 + f'(x)}{1 - f'(x)} - \lambda(1-x) = 0.$$

This leads directly to

$$f'(x) = \tanh \lambda(1-x),$$

and thus to the result stated in the introduction.

For the general set of constraints (9), the calculation goes in the same way, and it would not be hard to prove uniqueness for the minimizing f. The proof that this f is the nucleus is another matter; even in the simplest cases one needs the technique of steepest descent for the sums which arise in estimating $N_\epsilon(\mathcal{A} \cap S(f, r))$. Roughly speaking, it is $I(f')$ which controls these estimates, and the precise values of ϵ and r play a minor role. We spare the reader the details, at least for the present.

REFERENCES

1. A.N. Kolmogorov and V.M. Tihomirov, ϵ-entropy and ϵ-capacity of sets in functional spaces, Uspehi Mat. 14 (1959) 3-86; American Math. Soc. Translations 17 (1961) 277-364.

Massachusetts Institute of Technology

https://doi.org/10.4153/CMB-1968-009-8 Published online by Cambridge University Press