SUBDIVISION OF SIMPLEXES - IS BISECTION BEST?

Martin Stynes

Problem 1

A closed and bounded interval in $\mathbb{R}$ is to be subdivided into 2 intervals by insertion of a single point. These 2 intervals in turn are to be subdivided into 4 intervals by insertion of a point in each. Continue this process. Let d_n be the length of the longest interval at the n th stage. How should the insertion points be chosen so as to minimize d_n

This is not very difficult! Obviously points should be inserted at the midpoints of intervals, i.e. the optimal polic is to bisect intervals at each stage.

Why is Problem 1 of interest? (No doubt for many reade this is a much harder question than Problem 1 itself!). Wel if we wish to solve $f(x) = 0$, $f:[a,b] \to \mathbb{R}$ with $f(a)f(b) < 0$, and we want after some fixed number of function evaluations t find an interval of minimum length that is guaranteed to cont- ain a root of f, then Problem 1 shows that the classical bise tion method is best. (For any other algorithm there is some function f for which the interval found is longer). In this article we shall examine the relevance of bisection to an n- dimensional generalization of Problem 1 which we'll call Problem n. This problem has as yet no complete solution. I arises in the comparison of methods used to solve $f(x)=(0,0..$ for $f : \mathbb{R}^n \to \mathbb{R}^n$, but we shall discuss it purely from a geomet- ric viewpoint.

Generalizing Problem 1 to $\mathbb{R}^n$

We first replace closed and bounded intervals by n-simp- lexes (triangles when n=2, tetrahedra when n=3). The reason that we generalize intervals to triangles and not rectangles is that if $g : \mathbb{R}^2 \to \mathbb{R}$, then on each triangle in $\mathbb{R}^2$ there is a unique affine function which interpolates g at the vertices

of that triangle; this is very useful when approximating a root of g, and rectangles do not have the same property. Similarly in higher dimensions.

Definitions: For $n \geq 1$, an n-simplex $S^n = (a_0a_1...a_n)$ is the closed convex hull of n+1 points $a_0, a_1, ..., a_n$ in $\mathbb{R}^q$, $q \geq n$, such that the vectors $a_1-a_0, a_2-a_0, ..., a_n-a_0$ are linearly independent. The points $a_0, a_1, ..., a_n$ are called the vertices of S^n. Any m-simplex $(1 \leq m \leq n)$ formed by taking the closed convex hull of any m+1 vertices of S^n is called a face of S^n. The one-dimensional faces (a_ia_j), $0 \leq i < j \leq n$, are called the edges of S^n. The diameter of S^n, $d(S^n)$, is the length of the longest edge of S^n in the Euclidean norm.

We subdivide any n-simplex $S^n = (a_0a_1...a_n)$ as follows. Choose a point $y \in S^n$. Form all n-simplexes $(a_0a_1... a_{i-1}y$ $a_{i+1} ... a_n)$. Note that if m is the minimum dimension of a face of S^n containing y, then the subdivision yields m+1 n-simplexes.

Problem n

Given an n-simplex S^n subdivide it as just described. This is the first stage. Similarly subdivide the resulting n-simplexes by inserting a point in each. This is the second stage. Continue thus. Let A_k denote the set of all n-simplexes T^n obtained at the k th stage. Define

$$d_k = \max_{T^n \in A_k} d(T^n)$$

We find an algorithm for inserting points which will yield

$$d_{kn} \leq Cr^k \text{ for } k = 1,2,3...$$

where $r > 0$ (independent of S^n) is as small as possible and C depends on S^n only. (We consider d_{kn} rather than d_k as exp- erience shows it's a more natural measure).

Example: For the n=1 case with $S^1 = [a,b]$ the bisection method yields an equality:

$$d_k = (b-a)(1/2)^k, \; k = 1,2,3...$$

Only partial results have been obtained for Problem n. The principal reference is [7]. There it is shown (essentially) that for any n and any algorithm one must have $r \geq \frac{1}{4}$, but it is also conjectured that in fact one must have $r \geq \frac{1}{2}$. An algorithm for which $r = \frac{1}{2}$ is exhibited in [7]; it is based on Whitney's simplical subdivision [8, pp 358-360]. This algorithm may be fairly described as a generalization of the one-dimensional bisection method. Nevertheless a different generalization has become established as the "n-dimensional bisection method" [1,2,3,4,5,6]. We shall concentrate on this latter algorithm as it is simple to describe, it is clearly a generalization of the one-dimensional method, and yet it has not been satisfactorily analysed up to now.

The n-Dimensional Bisection Method

For $n > 1$, given an n-simplex T^n choose any edge $(a_i a_j)$ of T^n whose length is $d(T^n)$. Let b be the midpoint of this edge. Bisect T^n into $(a_0 .. a_{i-1}\ b\ a_{i+1}\ ..\ a_n)$ and $(a_0 ..\ a_{j-1}\ b\ a_{j+1} ..\ a_n)$. That is, b is the point inserted into T^n to subdivide it.

In attacking Problem n this method is intuitively attractive. To decrease the diameter of an n-simplex one must divide edges, and the bisection method bisects the longest edge. It's intuitively reasonable that the method will yield n-simplexes of diameters shrinking to zero, and this fact was implicitly assumed in [4]; however a proof did not appear until later [1].

To demonstrate the elementary nature of the arguments which can be used in relation to Problem n, we shall give a new proof of (a slightly stronger result than) the main theorem of [1].

Lemma. Given a triangle (2-simplex) of diameter d, the length of the median obtained by joining the midpoint of the longest edge to the opposite vertex is at most $\sqrt{3}\, d/2$.

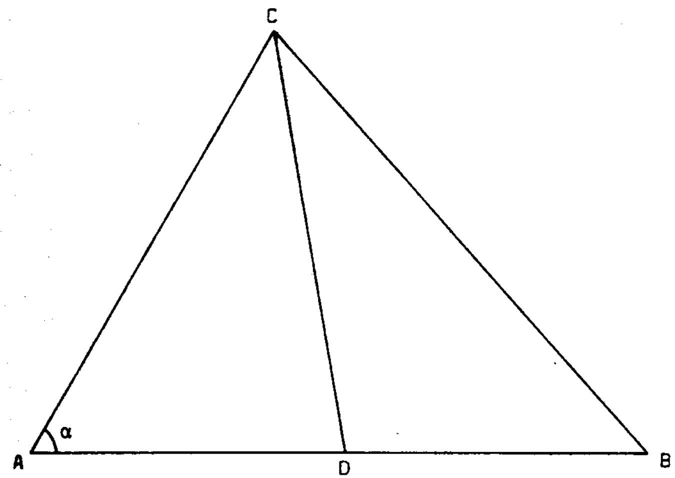

FIGURE 1

of. See Figure 1. There $AB \geq AC$, $AB \geq BC$, D is the mid-nt of AB. Now

$$BC^2 = AC^2 + AB^2 - 2AC.AB \cos\alpha$$

so $\quad AB \geq BC$ implies $\cos\alpha \geq AC/(2AB)$

Hence
$$\begin{aligned} CD^2 &= AC^2 + AD^2 - 2AC.AD\cos\alpha \\ &\leq AC^2 + AB^2/4 - AC.AB.AC/(2AB) \\ &\leq 3AB^2/4 \text{ as required.} \end{aligned}$$

Now any new edge $(a_k b)$ say formed by the bisection method n n-simplex T^n lies in the triangle $(a_i a_j a_k)$ just as CD in ABC above. From the Lemma it follows that the length new edge is at most $3\sqrt{d}(T^n)/2$.

Theorem. Let S^n be an n-simplex having exactly m+1 vertices as endpoints of edges of length greater than $\sqrt{3}d(S^n)/2$. Then after m iterations of the bisection method the diameter of any resulting n-simplex is at most $\sqrt{3}d(S^n)/2$.

Proof. Bisected edges have length at most $d(S^n)/2$. New edges have length at most $\sqrt{3}d(S^n)/2$ by the Lemma. So we need only show that after m iterations any edge of S^n whose length exceeds $\sqrt{3}d(S^n)/2$ has been bisected.

Let $S^n=(a_0a_1..a_m.a_n)$ where without loss of generality we assume that among all the a_k only $a_0, a_1,...,a_m$ are endpoints of edges whose lengths exceed $\sqrt{3}d(S^n)/2$. At the first bisection S^n becomes

$$S_1^n = (a_0..a_i..b..a_m..a_n)$$

and

$$S_2^n = (a_0..b..a_j..a_m..a_n)$$

where b is the midpoint of (a_ia_j) and $0 \leq i < j \leq m$. Consider S_1^n. By the first paragraph of the proof only the vertices $a_0,...,a_{j-1}, a_{j+1},...,a_m$ can be endpoints of edges exceeding $\sqrt{3}d(S^n)/2$ in length. That is, S_1^n has at most m vertices with this property. Similarly for S_2^n.

At the next iteration we will obtain 4 n-simplexes each having at most m-1 vertices which are endpoints of edges exceeding $\sqrt{3}d(S^n)/2$ in length. Repeating this argument n times in all proves the theorem.

Corollary. (Notation as in Problem n). For the n-dimensional bisection method we have

$$d_{kn} \leq C(\sqrt{3}/2)^k, \quad k = 1,2,3...$$

Proof. An n-simplex has n+1 vertices so in the Theorem we have m=n at most. Thus the Theorem implies that

$$d_n \leq d(S^n)\sqrt{3}/2.$$

Applying the Theorem again to each of the 2^n n-simplexes present after n iterations gives

$$d_{2n} \leq (d(S^n)\sqrt{3}/2).(\sqrt{3}/2) = d(S^n)\ (\sqrt{3}/2)^2.$$

Repeating this argument yields the Corollary.

In the notation of Problem n we have shown $r \leq \sqrt{3}/2$ for the bisection method. However for n=1 we clearly have $r = \frac{1}{2}$, and in fact for n > 1 all the computational evidence is that $r = \frac{1}{2}$ also. For n=2 it has been proven that $r = \frac{1}{2}$ [6] but the proof relies on a case by case analysis which is unlikely to extend to n > 2. Proving that $r = \frac{1}{2}$ for the bisection method when n > 2 is an open problem which I feel should not be too difficult - if one can find the right approach!

We close by pointing out that in fact $r \geq \frac{1}{2}$ for the bisection method.

Proposition. In the notation of Problem n we have $r \geq \frac{1}{2}$ for the bisection method.

Proof. For every n-simplex T^n let $V(T^n)$ denote the n-dimensional volume of T^n. Note that

$$V(T^n) < d^n(T^n) \qquad .... (*)$$

When an n-simplex is bisected it's not difficult to show that its n-dimensional volume is halved. Thus after kn iterations (k = 1,2,3,...) of the bisection method applied to S^n (say) the volume of any n-simplex T_k^n obtained is $V(S^n)/2^{kn}$.

Suppose now that the bisection method yields $r < \frac{1}{2}$ in Problem n. Then choosing a sequence of n-simplexes T_k^n, k = 1,2,3,... (notation as above) we have

$$\frac{V(T_k^n)}{d^n(T_k^n)} \geq \frac{V(S^n)}{C(2r)^{kn}} \to \infty \text{ as } k \to \infty.$$

This contradicts (*). Hence $r < \frac{1}{2}$ is impossible.

References

1. B. Kearfott. A Proof of Convergence and an Error Bound f the Method of Bisection in R^n. *Math. Comp.* 32 (1978), pp 1147-1153.
2. B. Kearfott. An Efficient Degree-Computation Method for Generalized Method of Bisection. *Numer. Math.* 32 (1979), pp. 109-127.
3. K. Sikorski. A Three-Dimensional Analogue to the Method Bisections for Solving Nonlinear Equations. *Math. Comp.* 33 (1979), pp. 722-738.
4. F. Stenger. Computing the Topological Degree of a Mappin in R^n. *Numer. Math.* 25 (1975), pp. 23-38.
5. M. Stynes. An n-Dimensional Bisection Method for Solving Systems of n Equations in n Unknowns. in Numerical Solut of Highly Nonlinear Problems. W. Foster (Ed.), North-Holland, Amsterdam, 1980, pp. 93-111.
6. M. Stynes. On Faster Convergence of the Bisection Method for all Triangles. *Math. Comp.* 35 (1980), pp. 1195-1201.
7. M.J. Todd. Optimal Dissection of Simplices. *SIAM J. Appl Math.* 34 (1978), pp. 792-803.
8. H. Whitney. Geometric Integration Theory, Princeton University Press, Princeton, N.J., 1957.

Department of Physical and Quantitative Sciences,
Waterford Regional Technical College.

A BIOGRAPHICAL GLIMPSE OF WILLIAM SEALY GOSSET

Philip J. Boland

William Sealy Gosset, alias 'Student' was an immensely talented scientist of diverse interests, but who will be primarily remembered for his contributions to the development of modern statistics. Born in Canterbury in 1876, he was educated at Winchester and New College, Oxford, where he studied chemistry and mathematics.

At the turn of the 19th century, Arthur Guinness, Son & Co. became interested in hiring scientists to analyse data concerned with various aspects of its brewing process. Gosset was to be one of the first of these scientists, and so it was that in 1899 he moved to Dublin to take up a job as a 'brewer' at St, James' Gate. In 1935 he left Dublin to become 'head brewer' in London but died soon thereafter at the young age of 61 in 1937.

After initially finding his feet at the brewery in Dublin, Gosset wrote a report for Guinness in 1904 on "The Application of the Law of Error to work of the Brewery". The report emphasised the importance of probability theory in setting an exact value on the results of brewery experiments, many of which were probable but not certain. Most of the report was the classic theory of errors (Airy and Merriman) being applied to brewery analysis, but it also showed signs of a curious mind at work exploring new statistical horizons. The report concluded that a mathematician should be consulted about special problems with small samples in the brewery. This led to Gosset's first meeting with Karl Pearson in 1905.

Karl Pearson (1857-1936) headed at University College London an industrious biometric laboratory which was very much concerned with large sample statistical analysis. Pearson had developed an extensive family of distribution curves,