JOURNAL OF APPROXIMATION THEORY 56, 41–47 (1989)

Piecewise Linear Approximations of Set-Valued Maps

ZVI ARTSTEIN

Department of Theoretical Mathematics,
The Weizmann Institute of Science,
Rehovot 76100, Israel

Communicated by Oved Shisha

Received March 19, 1986; revised May 24, 1988

1. INTRODUCTION

We consider in this note a set-valued function, say F, defined on $[0, 1]$, with values, $F(t)$, which are compact subsets of the Euclidean (real) n-space R^n. We wish to examine the possibility of approximating such a map F by one which is piecewise linear. Linear operations on sets are usually understood in the Minkowski sense; i.e., if α is a nonnegative real number and A, B are subsets of R^n, then $\alpha A = \{\alpha a: a \in A\}$ and $A + B = \{a + b: a \in A, b \in B\}$. These operations arise in the areas of stereology, integral geometry, optimization, control, and convex analysis; the approximation and interpolation of sets are natural problems in these areas.

If all the values $F(t)$ are convex sets, then the standard linear interpolation formula yields a piecewise linear set-valued map, which furnishes a good approximation, in parallel to the vector-valued case. Approximations along this line are examined in Vitale [8]. If, however, the values of F are not necessarily convex, then the standard interpolation may fail to supply an approximation, even if F is constant; see [8]. We comment on these phenomena in the next section.

To overcome the difficulty we approach the problem indirectly. The ensemble of compact sets can be regarded as a metric space; then the desired approximations are interpreted as abstract lines in the metric space. The approach is outlined in Section 3, where the basic questions and some answers are given.

A construction is offered in Section 4. It yields a solution to the previously posed abstract problems. It has also a natural meaning in terms of the linear structure in R^n. More properties, along with examples, remarks, and comments on other possibilities, are given in the closing section.

0021-9045/89 $3.00

2. The Straightforward Interpolation

Given a continuous function f defined on $[0, 1]$, into a linear topological vector space, a piecewise linear approximation is obtained as follows. A partition $\pi = \{0 = t_0 < t_1 < \cdots < t_k = 1\}$ is specified, then for $t_j \leqslant s \leqslant t_{j+1}$, the value of the approximation f_π is given by

$$f_\pi(s) = \frac{t_{j+1} - s}{t_{j+1} - t_j} f(t_j) + \frac{s - t_j}{t_{j+1} - t_j} f(t_{j+1}). \tag{2.1}$$

The two appealing properties of this linear interpolation are that f_π is indeed piecewise linear, i.e., its graph consists of a number of segments in the linear space, and that f_π approximates f uniformly as $\max(t_{j+1} - t_j)$ gets small.

The interpolation formula (2.1) may be applied to a set-valued function F, the operations being explained in the introduction. If the values of F are all compact and convex, and they vary continuously, then (2.1) yields a piecewise linear approximation, exactly as in the topological vector space case. Indeed it is possible to linearly embed the convex compact sets as a convex cone in a linear space; see Radstrom [5]. If, however, nonconvex values are allowed, then (2.1) may fail to provide an approximation at all. As an example take the constant set-valued map F with $F(t) = \{0, 1\}$ for all t. Then for any partition π, every interval in π contains a point s_0 such that $F_\pi(s_0) = \{0, \frac{1}{2}, 1\}$; the latter set cannot be regarded as a good approximation for $\{0, 1\}$.

3. An Abstract Approach

Since the direct formula does not work, we suggest another approach, namely to look for abstract lines, or pseudo-lines, connecting $F(t_j)$ and $F(t_{j+1})$ in the metric space of compact subsets of R^n. These lines should furnish good approximations, in a sense to be defined. But first we ought to commit ourselves to a specific metric. In this work we choose the Hausdorff metric which we recall now (see, e.g., Nadler [4, Definition (0, 1)]). The Hausdorff distance $h(A, B)$ between the two compact subsets A and B of R^n is given by

$$h(A, B) = \max\left(\max_{b \in B} \min_{a \in A} |a - b|, \max_{a \in A} \min_{b \in B} |a - b|\right), \tag{3.1}$$

where $|a - b|$ is the Euclidean distance between a and b. The Hausdorff metric arises naturally, and is used extensively in the applications and theory of set-valued maps. We ought, however, to point out that much of the analysis in this note depends heavily on the choice of the metric.

A rationale for using the linear interpolation (2.1) is that $\frac{1}{2}(f(t_j)+f(t_{j+1}))$ is the average of $f(t_j)$ and $f(t_{j+1})$, namely the solution to the least-squares problem. Likewise, $\alpha f(t_j)+(1-\alpha)f(t_{j+1})$ is the weighted average, namely, it minimizes, among all x, the value of $\alpha|f(t_j)-x|^2+(1-\alpha)|f(t_{j+1})-x|^2$. Looking for an abstract average we may adhere to this variational interpretation and, following Frechet [3], define the mean of the two compact sets A and B to be a set C such that

$$h(A, C)^2+h(B, C)^2=\min(h(A, X)^2+h(B, X)^2), \tag{3.2}$$

the *min* beeing taken over all compact sets X. The weighted average would be defined in a similar way. The piecewise linear approximation problem can then be phrased as follows.

(I) Given a partition π, find a continuous set-valued function F_π, such that $F_\pi(t_j)=F(t_j)$ for all j, and whenever $t_j \leqslant s_1 < s_2 < s_3 \leqslant t_{j+1}$, the value $F_\pi(s_2)$ is a weighted average of $F_\pi(s_1)$ and $F_\pi(s_3)$, with respective weights $(s_3-s_2)/(s_3-s_1)$ and $(s_2-s_1)/(s_3-s_1)$. (We may settle for less and demand that the average property holds for only $s_1=t_j$ and $s_3=t_{j+1}$.)

A solution to (I) is automatically a good uniform approximation of F, if $t_{j+1}-t_j$ are all small. A stronger property which we may wish to have is that the line joining $F(t_j)$ and $F(t_{j+1})$ in the metric space, namely the image of $[t_j, t_{j+1}]$, is by itself a segment, in the sense introduced by K. Menger (see the discussion in [7, Sect. 4]); namely it is an isometry of the interval $[0, h(F(t_j), F(t_{j+1}))]$. In our framework the condition is as follows.

(II) Given a partition π, find a continuous set-valued function F_π, such that $F_\pi(t_j)=F(t_j)$ for all j, and whenever $t_j \leqslant s_1 < s_2 \leqslant t_{j+1}$ then $h(F_\pi(s_1), F_\pi(s_2))=((s_2-s_1)/(t_{j+1}-t_j))\, h(F(t_j), F(t_{j+1}))$. (We may settle for less and demand this only in the case where either $s_1=t_j$ or $s_2=t_{j+1}$.)

The condition in (II) implies that in (I). This is true in a general metric space. (Indeed, for $r>0$ and $0<\alpha<1$ the pair of numbers (x_1, x_2) which minimizes $\alpha x_1^2+(1-\alpha)\,x_2^2$ subject to $x_1+x_2 \geqslant r$ is $x_1=(1-\alpha)\,r$ and $x_2=\alpha r$. If this is translated into distances, then together with the triangle inequality it shows that a line as defined in (II) is a solution to (I).)

Existence of the segments demanded in (II), namely an isometry $l(r)$ defined on $[0, h(A, B)]$ with $l(0)=A$ and $l(h(A, B))=B$, is not obvious, and depends strongly on the choice of the metric. Shephard and Webster [7, Sect. 4], studied this problem for four natural metrics on the family of convex bodies in R^n. Although the four metrics generate the same topology, only one of them, the Hausdorff metric, has the property that line segments are available. This sensitivity to the choice of the metric arises in other similar geometrical problems, e.g., the problem of metric selection (see Deutsch [2]) where linear selection of the metric projection is sought.

Another problem arises when one tries to follow these suggested guidelines. The Hausdorff distance, (3.1), is a sup type metric, and in sup metrics averages are not determined uniquely (as can easily be seen in R^2 with the max norm). This nonuniqueness, obvious for (I), is carried over to the stronger piecewise linear requirement (II). Counterexamples can be constructed even within the ensemble of convex sets; here is one.

EXAMPLE 3.1. Let π be the trivial partition of $[0, 1]$ and let $n=1$, i.e., the values of F are subsets of the real line. Let $F(0)=[0, 1]$ and $F(1)=[0, 2]$. Define $F_\pi(t)=[r(t), 1+t]$, with $r(t)$ continuous and $r(0)=r(1)=0$. If only r is differentiable, and $|r'(t)| \leqslant 1$, then F_π is a solution to (II).

Verification of the last claim is rather easy, as it is for the following modification, which shows that a solution to (II) may not preserve convexity or connectedness of the sets which are interpolated.

EXAMPLE 3.2. We modify the previous example by letting $F_\pi(t)$ be the union of two intervals $[r_1(t), r_2(t)]$ and $[r_3(t), 1+t]$, with $r_1(t) \leqslant r_3(t)$ and $r_i(0)=r_i(1)=0$ for $i=1, 2, 3$. If only each of the r_i is differentiable and $|r_i'(t)| \leqslant 1$, then F_π is a solution to (II).

Even though uniqueness does not hold, we may be interested in existence, and in solutions that preserve topological or linear properties of the original set-valued map. (If the values of F are convex sets then (2.1) provides a, but not necessarily the, solution to (II), with convex values.) Some answers are provided using a constructive approach in the next section.

4. A CONSTRUCTION

For simplicity of notation we start with the two values $A=F(0)$ and $B=F(1)$, and construct the interpolation $C(t)=F_\pi(t)$ on $[0, 1]$. The modification to a general partition is straightforward.

Given $a \in A$, we denote by $B(a)$ the set of points $b \in B$ such that $|a-b|$ is minimal. (The set $B(a)$ is nonempty, due to the compactness, and if B is convex then $B(a)$ is a singleton.) Similarly, for $b \in B$ the set $A(b)$ consists of the points in A closest to b in the Euclidean distance.

DEFINITION. We choose the set $C(t)$ to consist of all the points $(1-t)\,a+tb$, with $(a, b) \in A \times B$ and either $a \in A(b)$ or $b \in B(a)$.

Claim 4.1. For each t the set $C(t)$ is nonempty and compact. This follows from standard subsequencing arguments.

Claim 4.2. $C(0)=A$, $C(1)=B$. Trivial.

Clain 4.3. For every $0 \leqslant s_1 < s_2 \leqslant 1$, the distance $H(C(s_1), C(s_2))$ is equal to $(s_2-s_1)\,h(A, B)$; in particular C is a continuous set-valued map. To see this let us first verify that $h(C(s_1), C(s_2)) \leqslant (s_2-s_1)\,h(A, B)$. Let $c_1 \in C(s_1)$, then $c_1=(1-s_1)\,a+s_1 b$ for some (a, b) with either $a \in A(b)$ or $b \in B(a)$. In particular the point c_2 defined by $c_2=(1-s_2)\,a+s_2 b$, belongs to $C(s_2)$. A simple arithmetic shows that $|c_1-c_2|=(s_2-s_1)|a-b|$. Since either $a \in A(b)$ or $b \in B(a)$ it follows from (3.1) that $|a-b| \leqslant h(A, B)$ and the inequality is verified. Since the inequality was verified for all $0 \leqslant s_1 < s_2 \leqslant 1$, equality follows now from Claim 4.2 and the triangle inequality for the Hausdorff metric.

The properties just verified are those sought for in (II). The construction has also an interpretation in terms of the linear structure of R^n, as follows.

Claim 4.4. The graph of $C(t)$ is the union of affine functions on $[0, 1]$. Indeed, the graph of $C(t)$ is the union of the graphs of $(1-t)\,a+tb$ for eligible pairs (a, b). (Recall that for a point-valued function the interpolation would give one affine function.)

There are other interpolations with the previous property, and ours is not even the smallest one. For instance, if $r(t)$ in Example 3.1 is chosen such that $r''(t) \leqslant 0$, then the graph of the resulting set-valued map is a union of lines.

5. COMMENTS, EXAMPLES, MODIFICATIONS

Some comments on the convex case. If A and B in the construction are intervals in R^1, then the construction yields a convex-valued interpolation. This is not true in general for $n \geqslant 2$. As a counterexample consider the two segments $A=\{(p, r)\colon p=0, 0 \leqslant r \leqslant 1\}$ and $B=\{(p, r)\colon 0 \leqslant p \leqslant 1,\ r=0\}$. Then $C(t)$ is the union of the two segments $\{(p, r)\colon p=0,\ 0 \leqslant r \leqslant 1-t\}$ and $\{(p, r)\colon 0 \leqslant p \leqslant t, r=0\}$; and it is not convex. Preservation of convexity can be achieved by taking the convex hull $\mathrm{co}C(t)$ of the set $C(t)$ in the construction. The conditions in (II) will still be satisfied, since it is a general property of the Hausdorff distance that the convex hull operation is non-expansive, i.e., $h(\mathrm{co}C_1, \mathrm{co}C_2) \leqslant h(C_1, C_2)$. (Although taking the convex hull is a natural operation, I do not see any other reason for performing it. A good justification would be if the convex hull of a set were the closest

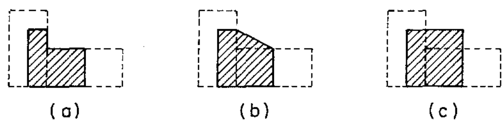

FIGURE 1

among the convex sets, but this is not correct.) By taking the convex hull of $C(t)$ we do not, in general, get back the solution to (II) furnished by (2.1) (we do in R^1). The example in the previous paragraph shows that; the convex hull of $C(\frac{1}{2})$ is $\{(p, r): p \geqslant 0,\ r \geqslant 0,\ p+r \leqslant \frac{1}{2}\}$ while $\frac{1}{2}A+\frac{1}{2}B=\{(p, r): 0 \leqslant p \leqslant \frac{1}{2}, 0 \leqslant r \leqslant \frac{1}{2}\}$.

In Fig. 1 we draw a variation on this example, The two sets, A and B are indicated by broken lines, the possible averages are the shaded areas. The average suggested by the construction is drawn in (a), its convex hull in (b), and $\frac{1}{2}A+\frac{1}{2}B$ in (c). Which is the "proper" average? This should probably be decided according to criteria supplied by the consumers of the average operation, e.g., the stereologists or the pattern analysts. See Serra [6] for an extensive discussion of the meaning and use of several geometrical operations. We do not dwell upon this problem here.

A remark on connectedness. Connectedness of A and B is not inherited by $C(t)$ in the construction. As an example consider in R^2 the sets $A=\{(p, r): r \geqslant 0, p^2+r^2=1\}$ and $B=\{(p, r): p \leqslant 0, p^2+r^2=\frac{1}{2}\}$. Both are connected, yet $C(\frac{1}{2})$ contains an isolated point $(\frac{1}{2}, \frac{1}{4})$. I do not see a natural operation which would, in the general case, change $C(t)$ into a connected set, maintaining the property in (II) (or in (I)).

A remark on empty values. Our set-valued maps always have nonempty values. (This is implied by the continuity of F, once F has one nonempty value; indeed the empty set is an isolated point in the Hausdorff metric.) Set-valued maps with empty values do, however, occur in the aforementioned applications. The construction, in general, can be applied to set-valued maps with empty values. The approximation will not be uniform, yet quite good visual similarities will be maintained. The same is true for semicontinuous set-valued maps. We leave out the details.

On lines and planes. The same formula that defines $C(t)$ in Section 4 for t in $[0, 1]$ is valid for all $-\infty < t < \infty$, and the resulting mapping satisfies Claim 4 for all $s_1 < s_2$. In particular, the mapping $C_1(t)=C(\alpha^{-1}t)$, with $\alpha=h(A, B)$, is an isometry of the entire real line into the space of compact sets, with $C(0)=A$, $C(\alpha)=B$. I do not know if, given A, B and D, there exists an isometry of the plane into the space of compact sets which passes through A, B and D; likewise for higher dimensions. If the answer is positive it may lead to a piecewise linear approximation of set-valued maps defined over planar or higher dimensional regions. I do not see a way to extend the construction of this paper to planar domains. Interpolations for higher dimensional domains can be obtained using the construction in Antosiewicz and Cellina [1], but these are not piecewise linear.

On the modification of the metric. As we noted before, the particular choice of the Hausdorff metric plays a crucial role in the analysis, in particular the drawback induced by the sup norm. For convex compact sets, Vitale [9] offers metrics which do not suffer from this drawback. The idea is to use the L_2 (or L_p in general) norm on the support function, instead of the max norm which characterizes the Hausdorff metric. The topology is unchanged but now all the boundary points of the convex set "participate" in the determination of the distance from another set. If Vitale's L_2 norm is adopted then there is a unique solution to (II), and it is the one given by the linear interpolation (2.1). Indeed, the convex compact sets are in this way linearly embedded in a Hilbert space. It would be nice to have analogous tools for large classes of not necessarily convex sets.

Acknowledgment

I thank the anonymous referee for thoughtful comments and valuable references.

Note added in proof. I found out that D. Levin in his paper "Multidimensional reconstruction by set-valued approximations," *IMA J. Numer. Anal.* **6** (1986), 173–184, deals with the same problem that the present paper does, employing, however, a different approach.

References

1. H. A. Antosiewicz and A. Cellina, Continuous extensions: Their construction and their application in the theory of differential equations, *in* "International Conference on Differential Equations" (H. A. Antosiewicz, Ed.), pp. 1–8, Academic Press, New York, 1975.
2. F. Deutsch, A survey of metric selections, *in* "Fixed Points and Nonexpansive Mappings" (R. Sine, Ed.), pp. 49–71, Contemporary Mathematics, Vol. 18, Amer. Math. Soc., Providence, RI, 1983.
3. M. Frechet, Les élements aléatoires de nature quelconque dans un espace distancie, *Ann. Inst. Henri Poincaré* **10** (1948), 215–230.
4. S. B. Nadler, Jr., "Hyperspaces of Sets," Dekker, New York, 1978.
5. H. Radstrom, An embedding theorem for spaces of convex sets, *Proc. Amer. Math. Soc.* **3** (1952), 165–169.
6. J. Serra, "Image Analysis and Mathematical Morphology" Academic Press, New York/London, 1982.
7. G. S. Shephard and R. J. Webster, Metrics for sets of convex bodies, *Mathematika* **12** (1965), 73–88.
8. R. A. Vitale, Approximations of convex set-valued functions, *J. Approx. Theory* **26** (1979), 301–316.
9. R. A. Vitale, L_p metrics for compact convex sets, *J. Approx. Theory* **45** (1985), 280–287.