

Uniform Approximation of Vector-Valued Functions with a Constraint

By Geneva G. Belford

Abstract. This paper deals with existence and characterization of best approximations to vector-valued functions. The approximations are themselves vector-valued functions with components taken from a linear space, but the constraint is imposed that certain of the approximation parameters should be identical for all components.

1. Introduction. The following situation often occurs in the analysis of experimental data. A set of experimental curves, all of which should roughly fit the same theoretical formula, has been obtained, and one desires to determine certain parameters in the theoretical formula. The experimental conditions have been such that, although some of the parameters should be the same for all of the given curves, others vary with the particular experiment. One would like, therefore, to be able to find the set of theoretical curves which simultaneously best approximates the set of given curves, under the constraint that some of the theoretical parameters should take on the same value for all members of the approximating set. As a simple example, one might want to fit a set of experimental curves by a set of straight lines, all with the same slope but with varying axis-intercepts.

In the next section, we give a precise formulation of this problem for approximation from linear families. Following this, we provide existence and characterization theorems and discuss some examples.

2. Formulation of the Problem. Let g_1, \dots, g_l be a given set of real functions continuous on a closed interval I of the real line. For brevity, let g denote the l -dimensional vector-valued function with components $\{g_i\}$. Let V be an $n + m$ dimensional linear subspace of $C(I)$, and suppose that $V = V_1 \oplus V_2$, where V_1 and V_2 are n - and m -dimensional subspaces, respectively. We then define the family of approximating vectors

$$F = \{(\varphi + \psi_1, \varphi + \psi_2, \dots, \varphi + \psi_l) : \varphi \in V_1, \psi_i \in V_2, i = 1, 2, \dots, l\}.$$

The uniform norm $N(f)$ of any vector-valued function f with components f_i in $C(I)$ is defined by

$$N(f) = \max_i \|f_i\|,$$

where $\|f_i\| = \max_{x \in I} |f_i(x)|$.

Let $\rho(g) = \inf_{f \in F} N(g - f)$. An element f in F is then called a best approximation to g from F if

Received June 4, 1971.

AMS 1970 subject classifications. Primary 41A50, 41A10.

Key words and phrases. Uniform approximation, vector-valued approximation, linear approximation, characterization of best approximation, polynomial approximation.

Copyright © 1972, American Mathematical Society

$$N(g - f) = \rho(g).$$

If $n = 0$, a best approximation f is obtained by constructing each component f_i as a best approximation from V_2 to g_i . If $m = 0$, we are seeking a best approximation φ from V_1 to the set of functions $\{g_i\}$. This problem has been discussed in a paper by Dunham [1]. The new cases are therefore those for which neither n nor m is zero. If one thinks of the subspace V as fixed, it is clear that the various $n > 0$ cases may be thought of as constrained versions of the simple $n = 0$ situation.

3. Existence of a Best Approximation. Existence is most easily proved by introducing a basis into V and considering the coefficient space. Let the functions v_1, \dots, v_n form a basis for V_1 and v_{n+1}, \dots, v_{n+m} form a basis for V_2 . Then the i th component $f_i \equiv \varphi + \psi_i$ of any element f in F may be written

$$f_i = \sum_{j=1}^n \alpha_j v_j + \sum_{j=1}^m \beta_j v_{n+j}.$$

Thus, any f is characterized by its associated coefficient vector

$$a = (\alpha_1, \dots, \alpha_n, \beta_{11}, \dots, \beta_{1m}, \beta_{21}, \dots, \beta_{im}).$$

We can consider a as a vector in R^d , the space of real-valued vectors of dimension $d = n + ml$. We then refer to R^d as the "coefficient space." To emphasize the dependence of f on a , we may write $f \equiv f(a, x)$. Note that the dependence of $f(a, x)$ on a is linear; hence, a standard argument suffices to show that a best approximation from V always exists.

The convexity of the set of best approximations may be proved by the usual argument. With respect to the question of uniqueness, it is almost intuitively clear that (if neither V_1 nor V_2 is empty) a best approximation will be unique only in rare cases, even when V is assumed to be a Haar subspace. This point will be discussed later, when some examples are considered. We first give some characterization results.

4. Characterization of a Best Approximation. The key to characterizing best approximations lies in properly defining "extremals" of the error. Since l error functions must be simultaneously considered, we use a definition of extremal analogous to that used by Moursund [5] in studying simultaneous approximation of a function and its derivatives. A generalized version of this definition has also been used by Johnson [2], [3] to study approximation of vector-valued functions.

Definition. The pair (x, k) , where $x \in I$ and k is an integer, is called an extremal of the approximation f to g if

$$|f_k(x) - g_k(x)| = N(f - g).$$

The following theorems then follow the same scheme as the discussion of linear Chebyshev theory in the book by Meinardus [4, p. 13 ff.].

THEOREM 1. *Let $f(a, x)$ be a fixed element in F and suppose that subsets $D_k \subset I$ ($k = 1, \dots, l$) have the property that $g_k(x) - f_k(a, x) \neq 0$ for all x in D_k . If there exists no $f(b, x) \in F$ such that, for all k ,*

$$(1) \quad [g_k(x) - f_k(a, x)]f_k(b, x) > 0$$

for all $x \in D_k$, then

$$(2) \quad \rho(g) \geq \inf_k \inf_{x \in D_k} |g_k(x) - f_k(a, x)| \equiv \delta.$$

Proof. If $\delta = 0$, the result is trivial; so take $\delta > 0$. The proof is by contradiction; assume that $\rho(g) < \delta$. Then, for some $f(c, x)$ in F ,

$$N(g - f(c, \cdot)) < \delta \quad \text{or} \quad \max_k \max_{x \in I} |g_k(x) - f_k(c, x)| < \delta.$$

But then it follows that, for all k and all $x \in D_k$,

$$(3) \quad |g_k(x) - f_k(c, x)| < |g_k(x) - f_k(a, x)|.$$

Now, let $f(b, x) = f(c, x) - f(a, x)$. It may then be readily proved from (3) that, for any k and any $x \in D_k$, $[g_k(x) - f_k(a, x)]f_k(b, x) > 0$, which contradicts the hypothesis and so proves the theorem.

Any of the sets D_k (but not all of them) may be taken to be empty. For any such k , (1) is assumed satisfied (for any b) and the quantity $\inf_{x \in D_k} |g_k(x) - f_k(a, x)|$ is taken to be $+\infty$ (so that this k does not enter into the computation of δ).

Theorem 1 provides a way of obtaining a lower bound on $\rho(g)$. In addition, by choosing particular sets D_k , we obtain a sufficient condition for a best approximation. To be precise, let $E_k(a) = \{x: (x, k) \text{ is an extremal of the approximation } f(a, x) \text{ to } g\}$. Then, the following corollary holds.

COROLLARY 1. *If there exists no $f(b, x) \in F$ such that for all k the inequality $[g_k(x) - f_k(a, x)]f_k(b, x) > 0$ is satisfied for all $x \in E_k(a)$, then $f(a, x)$ is a best approximation to g .*

The proof is immediate, since the sets $E_k(a)$ satisfy the conditions on the D_k of Theorem 1, and (2) then becomes $\rho(g) \geq N(g - f(a, \cdot))$. A converse of this corollary is also obtainable, as follows.

THEOREM 2. *If there exists an $f(b, x) \in F$ such that, for each extremal (x, k) of the approximation $f(a, x)$ to $g(x)$, $[g_k(x) - f_k(a, x)]f_k(b, x) > 0$, then there exists an $f(c, x) \in F$ that is a better approximation to $g(x)$ than $f(a, x)$ is.*

Proof. Let sets $E_k(a)$ be defined as before. Then, for each k such that $E_k(a)$ is nonempty, let

$$\min_{x \in E_k(a)} \{[g_k(x) - f_k(a, x)]f_k(b, x)\} \equiv a_k > 0,$$

and let $\alpha = \min \{a_k\}$. (Values of k for which $E_k(a)$ is empty are ignored in determining this minimum.)

Now, for each k for which $E_k(a)$ is nonempty, choose open subsets $U_k \subset I$ such that $E_k(a) \subset U_k$ and

$$f_k(b, x)[g_k(x) - f_k(a, x)] > \alpha/2$$

for all x in U_k . (If $E_k(a) = \emptyset$, let $U_k = \emptyset$.) Furthermore, for each k let

$$c_k \equiv N(g - f(a, \cdot)) - \max_{I - U_k} |g_k(x) - f_k(a, x)| > 0,$$

and let $\gamma = \min \{c_k\}$.

We shall show that there exists a better approximation of the form

$$f(c, x) = f(a, x) + \epsilon f(b, x),$$

where ϵ is some real, positive scalar. First, note that

$$(4) \quad [g_k(x) - f_k(c, x)]^2 = [g_k(x) - f_k(a, x)]^2 - 2\epsilon[g_k(x) - f_k(a, x)]f_k(b, x) + \epsilon^2[f_k(b, x)]^2.$$

Letting $N(f(b, \cdot)) = \beta$, we see that (4) implies that, for all k and all $x \in U_k$,

$$(5) \quad [g_k(x) - f_k(c, x)]^2 \leq [g_k(x) - f_k(a, x)]^2 + \epsilon^2\beta^2 - 2\epsilon\alpha/2.$$

On the other hand, for all k and $x \in I - U_k$,

$$(6) \quad |g_k - f_k(c, x)| \leq |g_k(x) - f_k(a, x)| + \epsilon |f_k(b, x)| \leq N(g - f(a, \cdot)) - \gamma + \epsilon\beta.$$

By choosing

$$\epsilon \leq \min\{\gamma/2\beta, \alpha/2\beta^2\},$$

we therefore obtain an $f(c, x)$ for which (from (6) and (5))

$$|g_k(x) - f_k(c, x)| < N(g - f(a, \cdot))$$

for all k and all $x \in I$. Thus,

$$N(g - f(c, \cdot)) < N(g - f(a, \cdot)),$$

and the theorem is proved.

Combining Corollary 1 with Theorem 2, we immediately obtain a characterization theorem analogous to the well-known Kolmogorov Criterion.

THEOREM 3. *A necessary and sufficient condition for $f(a, x)$ to be a best approximation to $g(x)$ from F is that there exist no $f(b, x) \in F$ such that*

$$(7) \quad [g_k(x) - f_k(a, x)]f_k(b, x) > 0$$

for every extremal (x, k) of the approximation $f(a, x)$ to $g(x)$.

The theorems of this section provide the basis for a standard type of algorithm for the construction of best approximations. Specifically, (7) is a system of linear inequalities which may be solved for a vector b if $f(a, x)$ is not a best approximation. The proof of Theorem 2 then gives a method of constructing a better approximation than $f(a, x)$. (In practice, one would probably optimize the parameter ϵ [4, p. 128].) The process is then repeated if necessary. For more details, the reader is referred to discussions of the virtually identical method for the single-function case. (For example, see [4, p. 128].)

5. Examples and Discussion.

Example 1. Suppose that on $I = [0, 1]$ the pair of functions $g_1 = x^2$, $g_2 = 2x^2$ are to be approximated by straight lines with identical intercepts (i.e., $V = P_1$, the set of polynomials of degree one or less, and $V_1 = P_0$). A best approximation from F is provided by the function pair $f_1 = -\frac{1}{4} + x$; $f_2 = -\frac{1}{4} + 2x$. To verify this, note that $N(g - f) = \frac{1}{4}$ and extremals (x, k) are $\{(0, 1), (1, 1), (0, 2), (\frac{1}{2}, 2), (1, 2)\}$ with associated error signs $\{+, +, +, -, +\}$, respectively. The condition of Theorem 3 is clearly satisfied—in fact f_2 is the best unconstrained approximation from P_1 to g_2 .

The above example serves to illustrate the fact that even in simple cases of polynomial approximation the best approximation may not be unique. The situation

here is that, although f_2 is uniquely determined, f_1 is not; for example, $f_1 = -\frac{1}{4} + 9x/8$ also gives $\|f_1 - g_1\| = \frac{1}{4}$.

This example also suggests that it may be generally true that for some k , f_k should be a best unconstrained approximation to g_k from V . That this is not the case is shown by the following example.

Example 2. Suppose that the functions to be approximated are $g_1 = x^2$ and $g_2 = x^3$, $I = [-1, 1]$, $V = P_1$ and $V_1 = P_0$. The best unconstrained approximations from P_1 to g_1 and g_2 are $\frac{1}{2}$ and $3x/4$, respectively, and the error norms are $\frac{1}{2}$ and $\frac{1}{4}$. The best approximation from F , however, has components $f_1 = \frac{3}{8}$, $f_2 = \frac{3}{8} + 3x/4$, with $N(g - f) = \frac{5}{8}$. The extremals are $\{(-1, 1), (1, 1), (-1, 2), (\frac{1}{2}, 2)\}$ with associated error signs $\{+, +, -, -\}$. To verify that this approximation is best, one must show that there exists no $f(b, x) = (\alpha + \beta_1x, \alpha + \beta_2x)$ such that

$$\begin{aligned} \alpha - \beta_1 &> 0, & \alpha + \beta_1 &> 0, \\ \alpha - \beta_2 &< 0, & \alpha + \beta_2/2 &< 0. \end{aligned}$$

This set of inequalities is readily seen to be inconsistent; thus no such $f(b, x)$ exists. It is also not difficult to convince oneself from elementary considerations that in this example the best approximation is unique.

As is generally the case for constrained or nonlinear approximation problems, whether or not the best approximation is unique will depend upon the particular function being approximated. This is readily seen by making a change in the g of Example 2.

Example 3. In Example 2, let $g_1 = x$, but leave everything else the same. The set \bar{F} of best approximations from F is now

$$\bar{F} = \{f = (\beta x, 3x/4): 3/4 \leq \beta \leq 5/4\}.$$

In view of this nonuniqueness, it is important to look more closely at the properties of the set of best approximations and the associated extremals. In this connection, one may obtain the following result.

THEOREM 4. *Let $f(a, x)$ and $f(b, x)$ both be best approximations from F to g . Let*

$$f(c, x) = \lambda f(a, x) + (1 - \lambda)f(b, x) \quad (0 < \lambda < 1).$$

Then, if (\bar{x}, k) is an extremal of the approximation $f(c, x)$ to g , it is also an extremal of the approximations $f(a, x)$ and $f(b, x)$ to g . Furthermore, $f_k(a, \bar{x}) = f_k(b, \bar{x}) = f_k(c, \bar{x})$.

Proof. By the convexity of the set of best approximations, $f(c, x)$ is also best. Then,

$$\begin{aligned} \rho(g) &= |f_k(c, \bar{x}) - g_k(\bar{x})| \\ &= |\lambda(f_k(a, \bar{x}) - g_k(\bar{x})) + (1 - \lambda)(f_k(b, \bar{x}) - g_k(\bar{x}))|. \end{aligned}$$

If the conclusions of the theorem were not true, the right side of this equality would be strictly less than $\rho(g)$. Hence, the theorem is proved by contradiction.

COROLLARY. *Suppose the best approximation from F to g is not unique and let $f(c, x)$ be any point in the interior of the set \bar{F} of best approximations. Then, for every $f(b, x) \in \bar{F}$, if (\bar{x}, k) is an extremal of the approximation $f(c, x)$ to g , it is also an extremal of the approximation $f(b, x)$ to g . Furthermore, $f_k(c, \bar{x}) = f_k(b, \bar{x})$.*

The proof is immediate. Simply consider the line segment from $f(b, x)$ through $f(c, x)$ to some third point $f(a, x)$ in \bar{F} , and then apply Theorem 4.

One sees from this corollary that all interior points of \tilde{F} have identical sets of extremals. Furthermore, maximal sets of extremals must be associated with boundary points of \tilde{F} . Example 3 illustrates these remarks. In that case, \tilde{F} is one-dimensional, with boundary points for $\beta = \frac{3}{4}$ and $\frac{5}{4}$. For any interior point of \tilde{F} , the set of extremals is

$$\{(-1, 2), (1, 2), (1/2, 2), (-1/2, 2)\}.$$

At the boundary points, this set is augmented by the additional extremals

$$\{(-1, 1), (1, 1)\}.$$

Of course, one would not in general expect the extremals to be identical for all boundary points. Looking again at Example 1, we see that there

$$\tilde{F} = \{f = (-1/4 + \beta x, -1/4 + 2x): 1 \leq \beta \leq \sqrt{2}\}.$$

For the boundary point at $\beta = 1$, we found the set of extremals

$$\{(0, 1), (1, 1), (0, 2), (1/2, 2), (1, 2)\},$$

while at $\beta = \sqrt{2}$ the set of extremals is

$$\{(0, 1), (1/\sqrt{2}, 1), (0, 2), (1/2, 2), (1, 2)\}.$$

The fact that neither of these sets contains the other is sufficient to show that, if they are extremal sets for best approximations, the corresponding best approximations lie on the boundary of \tilde{F} . In this example, the extremal set for any interior point of \tilde{F} is just the intersection of the extremal sets for the boundary points, although, in view of Example 3, such a result will not hold in general.

Department of Mathematics
University of Illinois
Urbana, Illinois 61801

1. C. B. DUNHAM, "Simultaneous Chebyshev approximation of functions on an interval," *Proc. Amer. Math. Soc.*, v. 18, 1967, pp. 472-477. MR 35 #3334.
2. L. W. JOHNSON, "Uniform approximation of vector-valued functions," *Numer. Math.*, v. 13, 1969, pp. 238-244. MR 40 #610.
3. L. W. JOHNSON, "Unicity in the uniform approximation of vector-valued functions," *Bull. Austral. Math. Soc.*, v. 3, 1970, pp. 193-198.
4. G. MEINARDUS, *Approximation of Functions: Theory and Numerical Methods*, Springer Tracts in Natural Philosophy, vol. 13, Springer-Verlag, New York, 1967. MR 36 #571.
5. D. G. MOURSUND, "Chebyshev approximations of a function and its derivatives," *Math. Comp.*, v. 18, 1964, pp. 382-389. MR 29 #3804.