

# Vector-Valued Approximation and its Application to Fitting Exponential Decay Curves\*

By Geneva G. Belford

**Abstract.** This paper deals with characterization of best approximations to vector-valued functions. The approximations are themselves vector-valued functions with components depending nonlinearly on the approximation parameters. The constraint is imposed that certain of the parameters should be identical for all components. An application to exponential approximation is discussed in some detail.

**1. Introduction.** The work reported in this paper was motivated by the following problem: Suppose a set of experimentally determined exponential decay curves is given. It is desired to approximate the curves by functions of the form  $\alpha \exp(\beta x)$ , where  $\beta$  should be the same for the entire set of curves and  $\alpha$  may vary from curve to curve. The problem is to determine how such approximation might best be made. This problem arises in a number of physical situations. In chemical kinetics, for example, monitoring of a chemical reaction which obeys a first-order rate law leads to just such exponential data, from which one wishes to extract a best  $\beta$  although the initial amount of material ( $\alpha$ ) varies from experiment to experiment.

In a previous paper [1], this type of constrained vector-valued approximation was studied for the simpler situation where the approximating functions depend linearly on the parameters. In this paper, results for nonlinear approximation are presented. Section 2 contains a precise formulation of the problem and a characterization theorem applicable to the construction of best approximations from general classes of nonlinear families. In Section 3, the particular problem discussed in the preceding paragraph is taken up. A very simple alternation theorem is obtained, as well as an interesting theorem on uniqueness.

**2. Formulation of the Problem and a Characterization Theorem.** Let  $g_1, g_2, \dots, g_l$  be a given set of real functions continuous on a closed interval  $I$  of the real line and let  $g$  denote the  $l$ -dimensional vector-valued function with components  $\{g_i\}$ . Let  $V$  be an  $n$ -parameter family of functions in  $C(I)$ . Denote an arbitrary element of  $V$  by  $\phi(\alpha_1, \dots, \alpha_n; x)$ , where  $(\alpha_1, \dots, \alpha_n) \in R^n$  is the parameter vector, and assume that  $\phi$  depends continuously on the  $\alpha_i$  as well as on  $x$ . For any integer  $m$  ( $0 \leq m \leq n$ ), we then define the family of approximating vectors

---

Received August 17, 1971, revised February 9, 1973.

*AMS (MOS) subject classifications* (1970). Primary 41A50.

*Key words and phrases.* Uniform approximation, vector-valued approximation, exponential approximation, nonlinear approximation.

\* This work was supported in part by the Advanced Research Projects Agency of the Department of Defense and was monitored by the U. S. Army Research Office-Durham under Contract No. DAHC04-72-C-0001.

Copyright © 1974, American Mathematical Society

$$F = \{f_1(a, x), f_2(a, x), \dots, f_l(a, x)\}:$$

$$a = (\alpha_{11}, \alpha_{12}, \dots, \alpha_{1m}, \alpha_{21}, \dots, \alpha_{2m}, \dots, \alpha_{lm}, \alpha_{m+1}, \dots, \alpha_n) \in R^q,$$

$$q = m(l - 1) + n, \text{ and } f_i(a, x) = \phi(\alpha_{i1}, \dots, \alpha_{im}, \alpha_{m+1}, \dots, \alpha_n; x)\}.$$

For example, the approximating family for a set of exponential decay curves (as described in the introduction) would be defined by  $\phi = \alpha_1 \exp(\alpha_2 x)$  and  $m = 1$ . In order to avoid the double-subscript notation, we will henceforth write  $a = (\alpha_1, \alpha_2, \dots, \alpha_q)$ .

The norm used in this paper is the usual uniform norm; that is, if  $f$  is any vector-valued function with components  $f_i$  in  $C(I)$ , the norm  $N(f)$  is defined by

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

where  $\|f_i\| = \max_{x \in I} |f_i(x)|$ . An element  $f$  in  $F$  is then called a best approximation to  $g$  from  $F$  if

$$N(g - f) = \inf_{f \in F} N(g - f) \equiv \rho(g).$$

One does not in general expect the existence of such a best approximation, since best approximations to a single function from nonlinear families often fail to exist. Uniqueness of the best approximation is also the exception rather than the rule, as is the case for the simpler situation when  $F$  is a linear subspace [1]. Comments on existence and uniqueness can be provided in particular cases, however, as will be seen in the next section.

In order to obtain characterization results, we first define the notion of "extremal."

**Definition 1.** The pair  $(x, k)$  is called an extremal of the approximation  $f(a, \cdot)$  to  $g$  if

$$|g_k(x) - f_k(a, x)| = N(g - f(a, \cdot)).$$

Let  $E_k(a) \equiv \{x : (x, k) \text{ is an extremal}\}$ .

Assuming that the partial derivatives  $\partial f_k(a, x)/\partial \alpha_i$  all exist and are continuous for  $(a, x) \in R^q \times I$ , we then extend the representation condition of Krabs [4] to the vectorial case as follows.

**Definition 2.** The family  $F$  is said to satisfy the representation condition if, for every pair of functions  $f(a, x), f(b, x)$  in  $F$ , there exist real numbers  $c_i(a, b)$  ( $j = 1, \dots, q$ ) and functions  $\psi_k(a, b; x)$  ( $k = 1, \dots, l$ ) positive on  $I$  such that

$$(1) \quad f_k(a, x) - f_k(b, x) = \psi_k(a, b; x) \sum_{i=1}^q c_i(a, b) \frac{\partial f_k(a, x)}{\partial \alpha_i} \text{ for } k = 1, \dots, l.$$

Extending familiar arguments given by Meinardus [5] and Krabs [4] for single function approximation, one then obtains the characterization theorem:

**THEOREM 1.** Let  $F$  satisfy the representation condition and suppose that  $\rho(g) > 0$ . Then  $f(a, x)$  is a best approximation to  $g$  if and only if there exists no  $b = (\beta_1, \dots, \beta_q) \in R^q$  such that, for all  $k$  and all  $x \in E_k(a)$ ,

$$(2) \quad [g_k(x) - f_k(a, x)] \left[ \sum_{i=1}^q \beta_i \partial f_k(a, x)/\partial \alpha_i \right] > 0.$$

**3. Application to Constrained Exponential Approximation.** In this section, we consider the particular case of approximation by one-term exponentials (elements of  $E_1 = \{ \alpha \exp(\beta x) : (\alpha, \beta) \in R^2 \}$ ), with the exponential factor  $\beta$  constrained to be the same for all components. That is, as noted before, we take  $f_k = \alpha_k \exp(\alpha_q x)$  for  $k = 1, \dots, l = q - 1$ ; the resulting family of vector-valued approximating functions will be denoted by  $F_{exp}$ . For simplicity, we shall take the interval  $I$  to be  $[0, 1]$ . The existence of best approximations from  $F_{exp}$  is readily demonstrated. The compactness of any bounded set  $\{f : f \in F_{exp}; N(f) \leq M\}$  is easily deduced from the known compactness result for  $l = 1$  [6]. The usual existence argument then goes through.

In order to apply Theorem 1, we must first verify that the representation condition holds for  $F_{exp}$ . (The smoothness condition prefacing Definition 2 is clearly satisfied.) The representation condition is known to hold for exponential functions [4], but extension from scalar to vector-valued functions is certainly not obvious, because of the requirement that the coefficients  $c_i$  not vary with  $k$ . Letting  $a = (\alpha_1, \dots, \alpha_q)$  and  $b = (\beta_1, \dots, \beta_q)$ , we need (from (1))

$$(3) \quad \alpha_k \exp(\alpha_q x) - \beta_k \exp(\beta_q x) = \exp(\alpha_q x) \psi_k(a, b; x) \{c_k(a, b) + \alpha_k c_q(a, b)x\}$$

for  $k = 1, 2, \dots, q - 1$ .

For  $\psi_k$  to be nonvanishing on  $[0, 1]$  as required, the linear factor on the right (in braces) must have a zero at the same point as does the left side. This condition leads to

$$(4) \quad c_k/c_q = -\alpha_k \log(\alpha_k/\beta_k)/(\beta_q - \alpha_q) \quad (k = 1, 2, \dots, q - 1).$$

(Equation (4) is obtained under the assumptions that  $\alpha_k \beta_k > 0$  ( $k = 1, \dots, q - 1$ ) and  $\beta_q - \alpha_q \neq 0$ . The other cases are readily handled by similar arguments.) Furthermore, the positivity of  $\psi_k$  requires also that the signs of both sides of (3) should match at any point  $\bar{x}$ . Taking  $\bar{x} = 0$ , we arrive at the condition

$$(5) \quad \text{sgn}(\alpha_k - \beta_k) = \text{sgn}(c_k) \quad (k = 1, \dots, q - 1).$$

Now since  $\text{sgn}(\alpha_k - \beta_k) = \text{sgn}[\alpha_k \log(\alpha_k/\beta_k)]$ , it is clear that by choosing any  $c_q$  such that  $\text{sgn } c_q = -\text{sgn}(\beta_q - \alpha_q)$  and then solving (4) for  $c_1, \dots, c_{q-1}$ , a set of scalars  $c_1, \dots, c_q$  satisfying (4) and (5) may be found. The functions  $\psi_k(a, b)$  are then defined to be

$$\psi_k \equiv \frac{\alpha_k \exp(\alpha_q x) - \beta_k \exp(\beta_q x)}{\exp(\alpha_q x) \{c_k + \alpha_k c_q x\}}, \quad k = 1, \dots, q - 1,$$

and the representation condition is verified.

Theorem 1 therefore is applicable. In order to deduce an alternation theorem from it, we first note the following familiar facts. Firstly, consistency or inconsistency of the inequalities (2) does not depend on the magnitude of the approximation error ( $N(g - f(a, \cdot))$ ) but only on the signs  $\text{sgn}[g_k(x) - f_k(a, x)]$ , which we shall denote by  $\sigma_k(x)$ . Secondly, a necessary and sufficient condition for the inconsistency of the set of inequalities (2) (where  $(x, k)$  runs over all extremals) is that the origin  $\theta = (0, \dots, 0)$  of  $R^q$  should lie in the convex hull of the set of  $q$ -vectors

$$(6) \quad S = \left\{ \sigma_k(x) \left( \frac{\partial f_k(a, x)}{\partial \alpha_1}, \dots, \frac{\partial f_k(a, x)}{\partial \alpha_q} \right) : (x, k) \text{ is an extremal} \right\}.$$

(This result on linear inequalities may be found in Cheney’s book [3, p. 19].) Finally, note that by Carathéodory’s theorem [3, p. 17] the condition that the origin  $\theta$  of  $R^q$  should be in the convex hull of  $S$  may be replaced by the condition that  $\theta$  should be a convex linear combination of some  $q + 1$  (or fewer) elements of  $S$ . For our exponential approximation, then, the condition is that there exist extremals  $(x_{ki}, k)$  (with  $k = 1, \dots, l, i = 1, \dots, \nu_k$ , and  $\sum_k \nu_k \leq q + 1$ ) and nonnegative constants  $\lambda_{ki}$  satisfying  $\sum_k \sum_i \lambda_{ki} = 1$  such that

$$(7) \quad \sum_{i=1}^{\nu_k} \lambda_{ki} \sigma_{ki} \exp(\alpha_q x_{ki}) = 0 \quad (k = 1, \dots, q - 1),$$

$$\sum_{k=1}^{q-1} \sum_{i=1}^{\nu_k} \sigma_{ki} \lambda_{ki} \alpha_k x_{ki} \exp(\alpha_q x_{ki}) = 0.$$

(Here we have used  $\sigma_{ki}$  to denote  $\sigma_k(x_{ki})$ .) Of course, only those extremals  $(x_{ki}, k)$  for which  $\lambda_{ki}$  is nonzero play a role in actually characterizing a best approximation. Thus, one immediately sees from (7) that any  $k$  for which  $\nu_k = 1$  does not enter into the characterization. Considering the various possibilities involving indices  $k$  for which  $\nu_k > 1$ , one quickly arrives at the following alternation theorem.

**THEOREM 2.** *The vector-valued function  $f$  is a best approximation from  $F_{\text{exp}}$  to  $g$  on  $[0, 1]$  if and only if one of the following conditions holds.*

(I) *For some index  $k$ ,  $f_k$  is a best unconstrained approximation to  $g_k$  from  $E_1$  (i.e., there are three points of alternation if  $\alpha_k \neq 0$  and two points of alternation if  $\alpha_k = 0$  [5, p. 178]) and  $\|g_k - f_k\| = N(g - f)$ .*

(II) *There exist two indices (say  $k = 1, 2$ ) with four associated extremals  $(x_{11}, 1), (x_{12}, 1), (x_{21}, 2), (x_{22}, 2)$  such that  $\alpha_1 \neq 0, \alpha_2 \neq 0, x_{11} < x_{12}, x_{21} < x_{22}$ , and*

$$\sigma_{11}\sigma_{12} = -1,$$

$$\sigma_{21}\sigma_{22} = -1,$$

$$\sigma_{11}\sigma_{21} = -\text{sgn}(\alpha_1 \alpha_2).$$

*Example.* Let  $g = (1, x)$ . The best approximation from  $F_{\text{exp}}$  to  $g$  on  $[0, 1]$  is given by  $f_1 = \frac{2}{3}e^{\beta x}, f_2 = \frac{1}{3}e^{\beta x}$  with  $\beta = \log 2$ . One readily verifies that  $N(f - g) = \frac{1}{3}$ , extremals are  $(0, 1), (1, 1), (0, 2), (1, 2)$ , and the alternation requirements of condition (II) of Theorem 2 are satisfied.

The problem of finding a best approximation to a general  $g$  is simplified enormously by the knowledge (from Theorem 2) that one need not consider more than two of the component functions simultaneously. Notice that the key problem is to determine a best  $\alpha_q$ , since then a set of  $\alpha_k$ ’s may be determined from the condition that  $\alpha_k \exp(\alpha_q x)$  should be the best approximation to  $g_k$  from the linear family  $\{\alpha \exp(\alpha_q x) : \alpha \in R\}$ . Thus one would proceed by constructing best unconstrained approximations from  $E_1$  to each  $g_k$ . If none of the exponential factors obtained in this way serves as a best  $\alpha_q$ , the next step is to examine all pairs  $g_{k_1}, g_{k_2}$ . For many of these pairs, the best approximation is characterized by condition (I) of Theorem 2 and is therefore of no further interest. If best approximations  $f_{k_1}, f_{k_2}$  characterized by condition (II) are then constructed for the remaining pairs, one of these necessarily yields a best  $\alpha_q$ .

This last step—construction of best approximations to a pair of functions  $\{g_1, g_2\}$ —deserves further discussion. We have first tried the straightforward approach based

on the fact that if (2) has a solution, a better approximation may be constructed from that solution. That is, with any extremal  $(x, k)$  of an approximation  $f(a, x)$  there is associated a linear inequality (from (2)) of the form

$$(8) \quad \sigma_k(x)\{\beta_k + \beta_3\alpha_k x\} > 0.$$

If the set of all such linear inequalities (associated with all extremals) has a solution  $\beta_1, \beta_2, \beta_3$ , then, for some  $\epsilon > 0$ ,  $f(a', x)$  provides a better approximation, where  $a' = (\alpha_1 + \epsilon\beta_1, \alpha_2 + \epsilon\beta_2, \alpha_3 + \epsilon\beta_3)$ . Thus by iteratively searching for extremals, solving inequalities (8), and correcting the approximation, one may hope to arrive eventually at a best approximation. In trials, this method has never failed to converge. A Remez-type algorithm has also been coded and limited trials to date show rapid convergence.

Further computational details and test results are available in [2]. One test may be worth mentioning here. We artificially generated "experimental" data by adding random errors  $\epsilon$  ( $|\epsilon| \leq .01$ ) to a set of three exponential curves of the form  $ae^{-x}$  with  $\alpha = 0.5, 1.0, 1.5$ . ("Data" points were computed for 20 equally spaced  $x$ -values on  $[0, 2]$ .) Our program, which identified extremals with a tolerance of  $10^{-4}$ , recovered the exponential factor  $\beta = -1$  as  $-1.0000$ . The traditional way of analyzing exponential data (least-squares straight-line fit to the logarithms of the function values) led to  $\beta = -0.9988$ . The difference is largely ascribable to the weighting induced by taking logarithms; direct least-squares fitting of exponentials is, however, a troublesome nonlinear problem even for a single function. Simultaneous uniform approximation appears to be a very practicable alternative.

Finally, we take up the question of uniqueness. It is obvious that the factors  $\alpha_k$  ( $k = 1, 2, \dots, q - 1$ ) are not in general unique. However, it is likely to be the parameter  $\alpha_q$  that is of principal interest, and, as the following theorem shows, this parameter is (with a trivial exception) uniquely determined.

**THEOREM 3.** *Let  $f(a, x)$  be a best approximation from  $F_{\text{exp}}$  to  $g$ . Let  $a = (\alpha_1, \dots, \alpha_q)$ . Suppose that either (i) condition (I) of Theorem 2 holds for some  $k$  such that  $\alpha_k \neq 0$ , or (ii) condition (II) of Theorem 2 holds for some pair of indices  $k_1, k_2$  such that  $\alpha_{k_1} \neq 0, \alpha_{k_2} \neq 0$ . Then if  $f(b, x)$ , with  $b = (\beta_1, \dots, \beta_q)$  is any other best approximation,  $\alpha_q = \beta_q$ .*

*Proof.* If (i) holds, the result is an immediate consequence of the uniqueness of best approximants from  $E_1$  [5, p. 178]. Now suppose that (ii) holds and for brevity let  $k_1 = 1, k_2 = 2$ . Assuming that  $\alpha_1\alpha_2 > 0$ , either the following set of inequalities or the set with all inequalities reversed must hold.

$$(9) \quad \beta_1 \exp(\beta_q x_{11}) - \alpha_1 \exp(\alpha_q x_{11}) \geq 0,$$

$$(10) \quad \beta_1 \exp(\beta_q x_{12}) - \alpha_1 \exp(\alpha_q x_{12}) \leq 0,$$

$$(11) \quad \beta_2 \exp(\beta_q x_{21}) - \alpha_2 \exp(\alpha_q x_{21}) \leq 0,$$

$$(12) \quad \beta_2 \exp(\beta_q x_{22}) - \alpha_2 \exp(\alpha_q x_{22}) \geq 0.$$

If equality holds in all four cases, then clearly  $\alpha_q = \beta_q$ . Hence, assume that one of these, say (9), is a strict inequality. Under the assumption that  $\alpha_1 > 0$ , (9) and (10) can only be compatible if

$$\exp[(\alpha_q - \beta_q)x_{11}] < \exp[(\alpha_q - \beta_q)x_{12}],$$

or, since  $x_{11} < x_{12}$ ,  $\alpha_q - \beta_q > 0$ . But if  $\alpha_q - \beta_q > 0$ , (11) and (12) are inconsistent.

Reversing all inequalities or making the other assumptions on the signs of  $\alpha_1, \alpha_2$  leads in the same way to inconsistency if at least one inequality is strict. Hence, we conclude that the equalities hold, and  $\alpha_q = \beta_q$ .

The argument above also shows that  $\alpha_1 = \beta_1$  and  $\alpha_2 = \beta_2$ . Therefore although the coefficients  $\alpha_k$  ( $k = 1, \dots, q - 1$ ) are in general not all unique, certain of the  $\alpha_k$ 's, in particular those entering into the characterization of a best approximation, are unique.

The theorem of Section 2 should also be found useful in constructing constrained vector approximations based on nonlinear families other than  $E_1$ . The especially simple alternation theorem for  $F_{\exp}$  depends heavily, however, on the nice properties of the exponential function and the fact that each  $f_k$  depends on only two parameters.

**4. Acknowledgement.** Programming and testing of a computer implementation of the method was carried out by Dr. Joseph Garber.

Center for Advanced Computation  
University of Illinois  
Urbana, Illinois 61801

1. G. G. BELFORD, "Uniform approximation of vector-valued functions with a constraint," *Math. Comp.*, v. 26, 1972, pp. 487–492.

2. G. G. BELFORD, *Simultaneous Fitting of Exponential Decay Curves*, CAC Document No. 61, Center for Advanced Computation, University of Illinois at Urbana-Champaign, 1973.

3. E. W. CHENEY, *Introduction to Approximation Theory*, McGraw-Hill, New York, 1966. MR 36 #5568.

4. W. KRABS, "Über differenzierbare asymptotisch konvexe Funktionenfamilien bei der nicht-linearen gleichmässigen Approximation," *Arch. Rational Mech. Anal.*, v. 27, 1967, pp. 275–288. MR 36 #4225.

5. G. MEINARDUS, *Approximation of Functions: Theory and Numerical Methods*, Springer, Berlin, 1964; English transl., Springer Tracts in Natural Philosophy, vol. 13, Springer-Verlag, New York, 1967. MR 31 #547; MR 36 #571.

6. E. SCHMIDT, "Zur Kompaktheit bei Exponentialsummen." *J. Approximation Theory*, v. 3, 1970, pp. 445–454. MR 42 #6471.