JOURNAL OF APPROXIMATION THEORY **27**, 127–134 (1979)

Necessary Conditions for Pareto Optimality in Simultaneous Chebyshev Best Approximation

YAIR CENSOR*

Medical Image Processing Group, Department of Computer Science, State University of New York at Buffalo, Amherst, New York 14226

Communicated by Oved Shisha

Received March 15, 1978

Necessary conditions for Pareto optimality in constrained simultaneous Chebyshev best approximation, derived from an abstract characterization theory of Pareto optimality, are presented. The generality of the formulation of the approximation problem dealt with here makes the results applicable to a large variety of concrete simultaneous best approximation problems. Some open problems are briefly described.

1. INTRODUCTION

Let $\{\psi_i^{(k)}(\alpha)\}_{i=1}^n$, $k = 1, 2,..., s$, be s given families of real-valued continuous functions on the interval $[a, b]$. Let $\psi_0^{(k)}(\alpha)$, $k = 1, 2,..., s$, be s given real-valued continuous functions on $[a, b]$. For $k = 1, 2,..., s$ define the following *nondifferentiable* but *convex* functions f_k: $R^n \to R$ by

$$f_k(\mathbf{x}) \equiv \left\| \psi_0^{(k)} - \sum_{i=1}^n x_i \psi_i^{(k)} \right\|_\infty = \max_{\alpha \in [a,b]} \left| \psi_0^{(k)}(\alpha) - \sum_{i=1}^n x_i \psi_i^{(k)}(\alpha) \right|.$$

Let us designate by *SBA* the problem of *constrained finite simultaneous Chebyshev best approximation* which is to characterize and/or find points $\mathbf{x} = (x_i) \in R^n$ that will solve the *multiobjective optimization problem*

$$\text{“min”}\{f_k(\mathbf{x})\}_{k=1}^s \qquad \text{such that} \qquad \mathbf{x} \in Q \subset R^n, \tag{P}$$

where Q stands for the feasible set of the problem and $s > 1$.

Solution (or solutions) of problem (P) depends, a priori, on which *solution concept* is chosen, i.e., on what meaning is attached to the symbol “min.” *Single dimensionality* solution concepts are those in which some real-valued function u: $Y \to R$, defined on $Y \equiv \{\mathbf{y} \in R^s \mid \mathbf{y} = f(\mathbf{x}),\ \mathbf{x} \in Q\}$, where $\mathbf{y} = f(\mathbf{x}) \equiv (f_1(\mathbf{x}), f_2(\mathbf{x}),..., f_s(\mathbf{x}))$, is first constructed and then extremal values of $u(\mathbf{y})$ on Y are sought.

* Present address: Department of Mathematics, University of Haifa, Mount Carmel, Haifa, Israel.

0021-9045/79/100127–08$02.00/0

The simultaneous approximation problems, dealt with in the papers mentioned below, differ in their formulations (some are Chebyshev while others are L_1 problems) and different methods are used to approach them. But, when it comes to the point of multiobjective minimization, they all fall under the heading of single dimensionality solution concepts. In fact, $u(y) \equiv \| \mathbf{y} \|_{l_1}$ is used in [6, 5], whereas in [3–5, 9, 10, 15–17] we encounter $u(\mathbf{y}) \equiv \| \mathbf{y} \|_{l_\infty}$.

The choice of the function $u(\mathbf{y})$ affects the nature of the simultaneous best approximation problem at hand and, therefore, one can quite reasonably ask what results can be obtained with other choices of $u(\mathbf{y})$, like, for example, $u(\mathbf{y}) \equiv \| \mathbf{y} \|_{l_p}$ with some $1 < p < \infty$. (In this respect see also [21, 11], where *compromise solutions* to the multiobjective problem are discussed.)

Our goal in this paper is to investigate the simultaneous best approximation problem for *Pareto optimality*, which is a *multidimensional solution concept* defined by

DEFINITION 1.1. A point $\mathbf{x}_0 \in R^n$ is called *Pareto optimal* for problem (P) if $\mathbf{x}_0 \in Q$ and there is no other $\mathbf{x} \in Q$ with $f_i(\mathbf{x}) \leqslant f_i(\mathbf{x}_0)$ for $i = 1, 2,..., s$, with at least one inequality strict.

To this end we put to work our abstract theory of Pareto optimality in multiobjective problems, formulated in [7] and published in [8], and derive necessary conditions for Pareto optimality in constrained finite simultaneous best approximation problems.

Pareto optimality in best approximation was discussed previously. While considering approximating functions which are *unisolvent of variable degree* and constitute a family which satisfies a condition called *the density condition*, Bacopoulos deals, in [1], solely with the *unconstrained* problem. In Gehner's work [12], constraints are allowed but only approximating functions of the form

$$\psi_i^{(k)}(\alpha) = \phi_i(\alpha)\, W_k(\alpha), \qquad i = 1, 2,..., n, \quad k = 1, 2,..., s,$$

are explicitly treated, where $\{\phi_i(\alpha)\}_{i=1}^n$ is any given set of functions which is *a Haar set* on $[a, b]$ and $\{W_k(\alpha)\}_{k=1}^s$ are positive weight functions. This family of approximating functions is quite large but might, in some cases, not be large enough. For example, if one wishes to approximate simultaneously a function and its derivatives by some given family of functions $\mathscr{F}$ and the families obtained from it by taking increasing order derivatives of the functions in $\mathscr{F}$, then cases which will not fit into Gehner's scheme may occur.

Here, a general framework allows us to handle a variety of constraints (including those appearing in [12]) and enables us to approximate by any linear combinations of continuous functions. Related to the subject matter of this paper are also [2] and some of the references mentioned therein.

Section 2 contains a brief review of relevant results from [8]. In Section 3 the problem is reformulated and subdifferential set calculations are carried out and in Section 4 the necessary conditions are presented. Section 5 concludes the paper with directions for further research and some other remarks.

Remark. Concepts, definitions, and notation, not explicitly explained here, are from the theory of optimization and from convex analysis. As our desk references for these we use [19, 20].

2. Some Results on Pareto Optimality in Multiobjective Problems

In [7, 8] a theory of Pareto optimality in multiobjective problems was proposed. Here is a condensed (no proofs) overview of relevant results from there.

Definition 2.1. (a) Let $f\colon R^n \to R$ and $\mathbf{x}_0 \in \operatorname{dom} f$. A $\mathbf{z} \in R^n$ *is called a direction of decrease* (*nonincrease*) *of* $f(\mathbf{x})$ *at* $\mathbf{x}_0$ if there exists a neighborhood of $\mathbf{z}$, U, and a real $\bar{\alpha} > 0$ such that $f(\mathbf{x}_0 + \alpha\mathbf{y}) < f(\mathbf{x}_0)$ ($f(\mathbf{x}_0 + \alpha\mathbf{y}) \leqslant f(\mathbf{x}_0)$) for every $\mathbf{y} \in U$ and every $\alpha \in (0, \bar{\alpha})$.

(b) Let $Q \subset R^n$ be a set with $\operatorname{int} Q \neq \varnothing$ and $\mathbf{x}_0 \in R^n$. A $\mathbf{z} \in R^n$ is called *a feasible direction for* Q *at* $\mathbf{x}_0$ if there exists a neighborhood of $\mathbf{z}$, U, and a real $\bar{\alpha} > 0$ such that $\mathbf{x}_0 + \alpha\mathbf{y} \in Q$ for every $\mathbf{y} \in U$ and every $\alpha \in (0, \bar{\alpha})$.

(c) Let $Q \subset R^n$ be a set with $\operatorname{int} Q = \varnothing$ and $\mathbf{x}_0 \in R^n$. A $\mathbf{z} \in R^n$ is called *a tangent direction for* Q *at* $\mathbf{x}_0$ if for *every* neighborhood of $\mathbf{z}$, U, and *every* real $\bar{\alpha} > 0$, there exist $\mathbf{y} \in U$ and $\alpha \in (0, \bar{\alpha})$ such that $\mathbf{x}_0 + \alpha\mathbf{y} \in Q$.

Lemma 2.2. (a) *Each of the following sets of directions generates an* open *cone with apex at the origin*: (i) *the directions of decrease*, (ii) *the directions of nonincrease*, (iii) *the feasible directions.*

(b) *The tangent directions generate a* closed *cone with apex at the origin.*

Definition 2.3. The *dual cone* K^* to the cone $K \subset R^n$ is the set of all continuous linear functionals which are nonnegative on K, i.e.,

$$K^* = \{\mathbf{y} \in R^n \mid \langle \mathbf{y}, \mathbf{x} \rangle \geqslant 0,\ \mathbf{x} \in K\}.$$

Next we give a characterization theorem for dual cones to cones of directions.

THEOREM 2.4. *Let $f: R^n \to R$ be a proper convex function and assume that $\mathbf{x}_0 \in \operatorname{int}(\operatorname{dom} f)$ and $\{\mathbf{x} \mid f(\mathbf{x}) < f(\mathbf{x}_0)\} \neq \varnothing$. If f is closed* or *has a relatively open effective domain then*

$$K^* = L^* = \{\lambda \mathbf{x} \mid \lambda \leqslant 0, \mathbf{x} \in \partial f(\mathbf{x}_0)\},$$

where K and L are the cones of directions of decrease, respectively nonincrease, of f at $\mathbf{x}_0$ and $\partial f(\mathbf{x}_0)$ stands for the subdifferential set *of f at $\mathbf{x}_0$ (see, e.g.,* [19, *p.* 215]).

With all these at hand, we can state the result on necessary conditions for Pareto optimality in multiobjective problems.

THEOREM 2.5. *Let $f_i: R^n \to R$, $i = 1, 2,..., s$, be proper convex functions with $\mathbf{x}_0 \in \operatorname{int}(\operatorname{dom} f_i)$ so that the assumptions of Theorem 2.4 hold for each of them. Denote by $K_i(L_i)$ the cones of directions of decrease (nonincrease) of f_i at $\mathbf{x}_0$. For $i = 1, 2,..., p$, let Q_i represent constraints sets with* int $Q_i \neq \varnothing$ *and M_i their cones of feasible directions at $\mathbf{x}_0$. Let Q_{p+1} be a constraint set with* int $Q_{p+1} = \varnothing$ *and M_{p+1} its cone of tangent directions at $\mathbf{x}_0$. Assume also that Q_i are convex for $i = 1, 2,..., p, p + 1$.*

Conclusion. A necessary condition for $\mathbf{x}_0$ to be a Pareto minimum for the problem

$$\begin{cases} \textit{Pareto-}\min\{f_i(\mathbf{x})\}_{i=1}^{s} \\ \textit{s.t. } \mathbf{x} \in Q = \bigcap_{i=1}^{p+1} Q_i \end{cases}$$

is that there exist $\beta_i \geqslant 0$ and $\mathbf{y}_i \in \partial f_i(\mathbf{x}_0)$, $i = 1, 2,..., s$, and linear functionals $l_k \in M_k^*$, $k = 1, 2,..., p + 1$, such that the $s + p + 1$ vectors $\beta_i \mathbf{y}_i$ $(i = 1, 2,..., s)$ and l_k $(k = 1, 2,..., p + 1)$ are not all identical zero and such that the following equation is satisfied:

$$\sum_{i=1}^{s} \beta_i \mathbf{y}_i = \sum_{k=1}^{p+1} l_k. \qquad (*)$$

The proof of this theorem follows from Theorem 3.8, Corollary 3.10, Theorem 4.12, and Theorem 5.1 of [8].

3. A REFORMULATION OF THE PROBLEM AND SOME SUBDIFFERENTIAL SET CALCULATIONS

The approach taken here resembles the way in which Pshenichnyi [18] treats the problem of characterizing a point of best approximation for a scalar (single-objective) problem.

Let $f_k(\mathbf{x}, \alpha)$, $k = 1, 2,..., s$, be real-valued functions with $\mathbf{x} \in Q \subset R^n$ and $\alpha \in A \subset R^m$, A a compact set. Assume that for every k both $f_k(\mathbf{x}, \alpha)$ and its gradient with respect to $\mathbf{x}$, $\nabla_{\mathbf{x}} f_k(\mathbf{x}, \alpha)$, are continuous, and that $f_k(\mathbf{x}, \alpha)$ is convex in $\mathbf{x}$ for all $\alpha \in A$. Define

$$f_k(\mathbf{x}) \equiv \max_{\alpha \in A} f_k(\mathbf{x}, \alpha), \qquad k = 1, 2,..., s,$$

and consider the multiobjective problem

$$\text{Pareto-min } \{f_k(\mathbf{x})\}_{k=1}^{s} \qquad \text{such that} \quad \mathbf{x} \in Q.$$

If we wish to determine conditions for optimality in some specific problem, on the grounds of the theory of [8], we must construct the dual cones to the cones of directions of decrease and to the cones of directions of non-increase of the objective functions. With the characterization theorem (Theorem 2.4, above), this problem "reduces" to that of calculating sub-differentials. In the present case, this is done with the aid of the following theorem of Valadier.

THEOREM 3.1 [14, Theorem 6.4.9]. *Let $\{f_\alpha\}_{\alpha \in A}$ be a family of functions such that $f_\alpha \in \text{conv}(X)$ (i.e., for every $\alpha \in A$, f_α: $X \to R$ is a real convex functional on the linear topological locally convex space X), and A is some compact set. Define $f \equiv \sup_{\alpha \in A} f_\alpha$ and assume that there exists an open set $U \subset X$ such that the mapping $A \times U \to f_\alpha(\mathbf{x})$ is finite and continuous on $A \times U$. Then for every $\mathbf{x}_0 \in U$,*

(i) *f is a continuous functional, and*

(ii) $\partial f(\mathbf{x}_0) = \text{cl co}[\bigcup_{\alpha \in F(\mathbf{x}_0)} \partial f_\alpha(\mathbf{x}_0)]$, *where* cl *denotes closure and* co *stands for convex hull and the set $F(\mathbf{x}_0)$ of α's over which the union is taken is given by*

$$F(\mathbf{x}_0) \equiv \{\alpha \in A \mid f_\alpha(\mathbf{x}_0) = f(\mathbf{x}_0)\}.$$

Now we calculate the subdifferential of $f_k(\mathbf{x})$ at a point $\mathbf{x}_0$. The existence of the gradient, with respect to $\mathbf{x}$, of $f_k(\mathbf{x}, \alpha)$ ensures, by Theorem 25.1 of [19], that $\partial f_k(\mathbf{x}_0, \alpha) = \{\nabla_x f_k(\mathbf{x}_0, \alpha)\}$. From (ii) of Theorem 3.1 we then obtain that

$$\partial f_k(x_0) = \text{cl co}\Big[\bigcup_{\alpha \in F_k(\mathbf{x}_0)} \{\nabla_{\mathbf{x}} f_k(\mathbf{x}_0, \alpha)\}\Big],$$

where

$$F_k(\mathbf{x}_0) \equiv \{\alpha \in A \mid f_k(\mathbf{x}_0, \alpha) = f_k(\mathbf{x}_0)\}, \qquad k = 1, 2,..., s.$$

The set $\bigcup_{\alpha \in F_k(\mathbf{x}_0)}\{\nabla_x f_k(\mathbf{x}_0, \alpha)\}$ is compact because A is compact by assumption, $F_k(\mathbf{x}_0)$ is compact as a closed subset of A, and $\nabla_x f_k(\mathbf{x}_0, \alpha)$ is continuous in α. The convex hull of a compact set in a finite-dimensional space is compact, therefore closed, thus the closure operation, cl, can be omitted from the formula for $\partial f_k(\mathbf{x}_0)$. Furthermore, every element in the convex hull of a set can be represented as a convex combination of $n + 1$ elements of the set [19, Theorem 17.1], hence we get the following representation of the subdifferential set of $f_k(\mathbf{x})$ at the point $\mathbf{x}_0$:

$$\partial f_k(\mathbf{x}_0) = \Big\{ \mathbf{y} \mid \mathbf{y} = \sum_{i=1}^{n+1} \lambda_i^{(k)} \nabla_{\mathbf{x}} f_k(\mathbf{x}_0, \alpha_i^{(k)}), \\ \lambda_i^{(k)} \geqslant 0, \sum_{i=1}^{n+1} \lambda_i^{(k)} = 1, \alpha_i^{(k)} \in F_k(\mathbf{x}_0), \forall i = 1, 2,..., n + 1 \Big\}.$$

4. Necessary Conditions for Pareto Minimum in a Constrained SBA Problem

For the SBA problem described in the Introduction constrained by constraints sets as described in Theorem 2.5 we give now necessary conditions for pareto minimum. Let us agree to denote

$$d_k(\alpha) \equiv \psi_0^{(k)}(\alpha) - \sum_{j=1}^{n} x_j{}^0 \psi_j^{(k)}(\alpha)$$

for any fixed $\mathbf{x}^0 = (x_j{}^0)$.

THEOREM 4.1. *If $\mathbf{x}^0$ is a Pareto minimum for the SBA problem then*

(a) *there exist for every $k = 1, 2,..., s$*

(i) *$n + 1$ scalars $\lambda_i^{(k)} \geqslant 0$ such that $\sum_{i=1}^{n+1} \lambda_i^{(k)} = 1$,*

(ii) *$n + 1$ points $\alpha_i^{(k)} \in [a, b]$ such that $f_k(\mathbf{x}^0) = | d_k(\alpha_i^{(k)})|$,*

(iii) *$n + 1$ numbers $t_i^{(k)} = \pm 1$ such that $t_i^{(k)} = \operatorname{sign} d_k(\alpha_i^{(k)})$;*

(b) *there exist scalars $\beta_k \geqslant 0$, $k = 1, 2,..., s$, not all of them zero,*

(c) *there exist vectors $\boldsymbol{l}_j \in M_j^*$, $j = 1, 2,..., p + 1$, not all of them equal to $\mathbf{0}$, belonging to the dual cones of the cones of feasible directions and of the cone of tangent directions to the constraints sets Q_j, $j = 1, 2,..., p + 1$, at $\mathbf{x}^0$ (see Theorem 2.5 above), such that*

$$\sum_{k=1}^{s} \beta_k \left[\sum_{i=1}^{n+1} \lambda_i^{(k)} t_i^{(k)} \Psi^{(k)}(\alpha_i^{(k)}) \right] = \sum_{j=1}^{p+1} \boldsymbol{l}_j ,$$

where $\Psi^{(k)}(\alpha)$ *stands for the vector in* R^n *whose components are* $\psi_i^{(k)}(\alpha)$, $i = 1, 2, \ldots, n$.

Proof. Let us define

$$f_k(\mathbf{x}, \alpha, t) \equiv t\left(\psi_0^{(k)}(\alpha) - \sum_{j=1}^{n} x_j \psi_j^{(k)}(\alpha)\right)$$

and then define

$$f_k(\mathbf{x}) \equiv \underset{(\alpha,t)\in[a,b]\times[-1,1]}{\mathrm{Max}} f_k(\mathbf{x}, \alpha, t).$$

Then,

$$\nabla_{\mathbf{x}} f_k(\mathbf{x}, \alpha, t) = -t\Psi^{(k)}(\alpha)$$

and the theorem follows from Theorem 2.5 with the aid of the calculation of $\partial f_k(\mathbf{x}_0)$ carried out in Section 3.

5. Conclusion

We have shown how one abstract multicriterion optimization scheme can be applied to a simultaneous best approximation problem. Here, necessary conditions for Pareto optimality of the approximation problem were derived, but some additional effort seems to be required before a complete characterization result along these lines can be achieved. Another point which calls for further investigation is the question whether a *vectorial alternation theorem* can be reached in the present formulation (compare with [15]). The formulation of the SBA problem here is quite general both in allowable approximating families and in constraints sets to which it applies. In a specific case the dual cones related to the constraints sets have to be constructed but once this is done those dual cones can be used again in every problem where such constraints appear. See lecture 10 of [13].

Acknowledgments

Work for this paper was supported by NIH Grants HL 18968, HL 4664, and RR7. The material presented here is based on part of the author's doctoral thesis written under the direction of Professor Adi Ben-Israel at the Technion, Haifa, Israel. The author wishes to thank Professor Ben-Israel for his constant help and encouragement.

References

1. A. Bacopoulos, Nonlinear Chebychev approximation by vector-norms, *J. Approximation Theory* **2** (1969), 79–84.

2. A. Bacopoulos and I. Singer, On convex vectorial optimization in linear spaces, *J. Optimization Theory Appl.* **21** (1977), 175–188.
3. E. Bredendiek, Charakterisierung und Eindeutigkeit bei Simultanapproximationen, *Z. Angew. Math. Mech.* **50** (1970), 403–410.
4. M. P. Carrol, Simultaneous L_1 approximation of a compact set of real valued functions, *Numer. Math.* **19** (1972), 110–115.
5. M. P. Carroll, On simultaneous L_1 approximation, *in* "Approximation Theory" (G. G. Lorentz, Ed.), pp. 295–297, Academic Press, New York, 1973.
6. M. P. Carroll and H. W. McLaughlin, L_1 approximation of vector-valued functions, *J. Approximation Theory* **7** (1973), 122–131.
7. Y. Censor, "Contributions to Optimization Theory: Multiobjective Problems," Doctoral thesis, Technion, Haifa, 1975. [in Hebrew].
8. Y. Censor, Pareto optimality in multiobjective problems, *Appl. Math. Optimization* **4** (1977), 41–59.
9. J. B. Diaz and H. W. McLaughlin, Simultaneous approximation of a set of bounded real functions, *Math. Comp.* **23** (1969), 583–594.
10. C. B. Dunham, Simultaneous Chebyshev approximation of functions on an interval, *Proc. Amer. Math. Soc.* **18** (1967), 472–477.
11. M. Freimer and P. L. Yu, Some new results on compromise solutions for group decision problems, *Management Sci.* **22** (1976), 688–693.
12. K. R. Gehner, Characterization theorems for constrained approximation problems via optimization theory, *J. Approximation Theory* **14** (1975), 51–76.
13. I. V. Girsanov, "Lectures on Mathematical Theory of Extremum Problems," Lecture. Notes in Economics and Mathematical Systems No. 67, Springer-Verlag, Berlin/Heidelberg/New York, 1972.
14. J.-P. Laurent, "Approximation et optimisation," Hermann, Paris, 1972.
15. K.-P. Lim, Simultaneous approximation of compact sets by elements of convex sets in normed linear spaces, *J. Approximation Theory* **12** (1974), 332–351.
16. R. A. Lorentz, Nonuniqueness of simultaneous approximation by algebraic polynomials, *J. Approximation Theory* **13** (1975), 17–23.
17. D. G. Moursund, Chebyshev approximations of a function and its derivatives, *Math. Comp.* **18** (1964), 382–389.
18. B. N. Pshenichnyi, "Necessary Conditions for an Extremum," Dekker, New York, 1971.
19. R. T. Rockafellar, "Convex Analysis," Princeton Univ. Press, Princeton, N.J., 1970.
20. J. Stoer and C. Witzgall, "Convexity and Optimization in Finite Dimensions I," Springer-Verlag, Berlin/Heidelberg/ New York, 1970.
21. P. L. Yu, A class of solutions for group decision problems, *Management Sci.* **19** (1973), 936–946.