Sanjo Zlobec

Two characterizations of Pareto minima in convex multicriteria optimization

Aplikace matematiky, Vol. 29 (1984), No. 5, 342–349

Persistent URL: http://dml.cz/dmlcz/104104

TWO CHARACTERIZATIONS OF PARETO MINIMA IN CONVEX MULTICRITERIA OPTIMIZATION*)

SANJO ZLOBEC

(Received May 5, 1983)

We give two conditions, each of which is both necessary and sufficient for a point to be a global Pareto minimum. The first one is obtained by studying programs where each criterion appears as a single objective function, while the second one is given in terms of a "restricted Lagrangian". The conditions are compared with the familiar characterization of properly efficient and weakly efficient points of Karlin and Geoffrion.

Key words: Pareto minimum, properly efficient point, weakly efficient point, characterization of optimality, convex functions.

1. INTRODUCTION

For real functions in n variables $\Phi^1, \ldots, \Phi^m$ a point x^* is a Pareto minimum (also called: "efficient point") if there is no other x such that

$$\Phi^k(x) \leqq \Phi^k(x^*), \quad k \in P = \{1, \ldots, m\}$$

with at least one strict inequality. Possibly the first constructive characterizations of Pareto minima for convex criteria Φ^k, $k \in P$ seem to be given in [3], see also [4]. These characterizations are given in terms of $2^{\operatorname{card} p}$, where p is the cardinality of P, systems of inequalities and in terms of the "minimal index set of active constraints" (see also [1] and [18]). In this paper we will give two different characterizations of Pareto minima. Our results are stated in terms of one or more Lagrangian type functions and therefore they come closer to the classical results of Karlin [10] and Geoffrion [9]. Their results, in the case of convex criteria, characterize more restrictive "properly efficient points" and less restrictive "weakly efficient points".

*) Research partly supported by the Natural Sciences and Engineering Research Council of Canada and by le Ministere de l'Education du Québec (F.C.A.C.)

The results of this paper can be generalized and set up in the framework of the abstract Dubovitskii-Milyutin optimization theory as it was done for necessary conditions for Pareto minimum by Censor [7]. For related ideas on Pareto minima and applications in economics and control see e.g. [6], [8], [11] and [14].

2. THE MULTI-LAGRANGIAN CHARACTERIZATION OF A PARETO MINIMUM

In order to avoid familiar technicalities, let us assume that all convex criteria are differentiable. We recall that the *cone of directions of constancy* of Φ^k at some arbitrary point x^* is the set (see e.g. [4]).

$$D_k(x^*) = \{d: \exists \bar{\alpha} > 0 \ni \Phi^k(x^* + \alpha d) = \Phi^k(x^*) \quad \forall \alpha \in [0, \bar{\alpha}]\}.$$

For an arbitrary set M we denote its *polar* by

$$M^+ = \{u: u^\mathsf{T} x \geqq 0 \ \forall x \in M\} .$$

For each $r \in P$, let us denote

$$P^r = P \setminus \{r\} = \{k \in P: k \neq r\} .$$

Also, for a fixed point x^*, let us denote by $F^r(x^*)$ the feasible set of the program (P, r), i.e.

$$F^r(x^*) = \{x: \Phi^k(x) \leqq \Phi^k(x^*), \ k \in P^r\}$$

and by $P^r_=(x^*)$ the corresponding minimal set of active constraints (see [1] or [4]), i.e.

$$P^r_=(x^*) = \{k \in P^r: x \in F^r(x^*) \Rightarrow \Phi^k(x) = \Phi^k(x^*)\} .$$

Finally, denote

$$P_=(x^*) = \bigcup_{r \in P} P^r(x^*) = \{k \in P: x \in F^r(x^*) \Rightarrow \Phi^k(x) = \Phi^k(x^*) \text{ for some } r \in P^k\}$$

and note that

$$P \setminus P_=(x^*) = \{k \in P: \exists \hat{x} \in F^r(x^*) \in \Phi^k(\hat{x}) < \Phi^k(x^*), \text{ for every } r \in P\} .$$

It is obvious (see also [7], [11], [12]) that x^* is a Pareto minimum if, and only if, x^* is an optimal solution of the following m programs:

$$\text{(P, r)} \qquad \operatorname{Min} \Phi^r(x) \quad \text{s.t.} \quad \Phi^k(x) \leqq \Phi^k(x^*), \quad k \in P^r, \qquad r = 1, \ldots, m .$$

By characterizing the optimality of x^* in (P, r) we are characterizing a Pareto minimum. Since the constraints in (P, r) may not satisfy the Slater condition, the Kuhn-Tucker conditions do not characterize optimality for (P, r). (The constraints $f^k(x) \leqq$

$\leqq 0$, $k \in P$ satisfy Slater's condition if there is an x^* such that $f^k(x^*) < 0$, $k \in P$, e.g. [4].) However, the complete characterizations (the "BBZ conditions") from [4] are readily applicable and they give the following result.

2.1. Theorem. *Let $\Phi^1, \ldots, \Phi^m$ be convex and differentiable criteria. Then an $x^* \in R^n$ is a Pareto minimum if, and only if, for every $r \in P$ the system*

$$\text{(2.1)} \qquad \nabla\Phi^r(x^*) + \sum_{k \in P^r \setminus P^r_=(x^*)} \lambda_k \nabla\Phi^k(x^*) \in \{ \bigcap_{i \in P^r_=(x^*)} D_i(x^*)\}^+$$

$$\lambda_k \geqq 0\,, \quad k \in P^r \setminus P^r_=(x^*)$$

has a solution. ■

If the index set $P^r_=$ is empty, then the intersection on the right-hand side in (2.1) is defined as the whole space, implying that its polar is zero. After adding all systems, and using the properties of polars, we obtain the following necessary condition for optimality.

2.2. Corollary. *Let $\Phi^1, \ldots, \Phi^m$ be convex and differentiable criteria. If $x^* \in R^n$ is a Pareto minimum, then there exist positive numbers $\lambda_k > 0$, $k \in P$ such that*

$$\text{(2.2)} \qquad \sum_{k \in P} \lambda_k \nabla\Phi^k(x^*) \in \{ \bigcap_{k \in P_=(x^*)} D_k(x^*)\}^+ . \qquad ■$$

We will return to the above corollary in Section 4.

3. THE SINGLE-LAGRANGIAN CHARACTERIZATION OF A PARETO MINIMUM

Our next result is stated in terms of the derivative of the single "restricted" Lagrangian

$$L(\lambda, x) = \sum_{k \in P \setminus P_=(x^*)} \lambda_k \Phi^k(x) .$$

3.1. Theorem. *Let the criteria Φ^k, $k \in P$ be convex and differentiable. Then an $x^* \in R^n$ is Pareto minimal if, and only if, either $P = P_=(x^*)$ or there exist non-negative numbers $\lambda_k \geqq 0$, $k \in P \setminus P_=(x^*)$ not all zero, such that*

$$\text{(3.1)} \qquad \sum_{k \in P \setminus P_=(x)^*} \lambda_k \nabla\Phi^k(x^*) \in \{ \bigcap_{k \in P_=(x^*)} D_k(x^*)\}^+ .$$

Proof. An x^* is *not* Pareto minimal if, and only if, there is an x, different from x^*, such that $\Phi^k(x) \leqq \Phi^k(x^*)$, $k \in P$ with at least one strict inequality. Hence

$$\text{(3.2)} \qquad \begin{aligned} \Phi^k(x) &< \Phi^k(x^*)\,, \quad k \in P \setminus P_=(x^*) \\ \Phi^k(x) &= \Phi^k(x^*)\,, \quad k \in P_=(x^*) . \end{aligned}$$

(First, note that $x \in F^r(x^*)$ for every $r \in P$ and then use the definition of $P_=(x^*)$ and standard convexity arguments.) But the consistency of (3.2) is equivalent to the consistency of the system

$$\nabla \Phi^k(x^*)\, d < 0\,, \quad k \in P \setminus P_=(x^*)$$

(3.3)
$$d \in \bigcap_{k \in P_=(x^*)} D_k(x^*)$$

and further, by a theorem of the alternative (say, Dubovitskii-Milyutin's one, e.g. [4] or Ben-Israel's from [2]), to the nonexistence of $\lambda_k \geqq 0$, $k \in P \setminus P_=(x^*)$ not all zero, such that (3.1) holds. The complementary statement proves the theorem. (If $P_=(x^*) = \emptyset$, then the set relation (3.3) becomes $d \in R^n$.)

4. COMPARISONS WITH CLASSICAL RESULTS

First, let us recall the following *sufficiency* result from Karlin's book [10]:

4.1. Proposition. *If the criteria Φ^k, $k \in P$ are convex and differentiable, and if there exist positive numbers $\lambda_k > 0$, $k \in P$, such that at some point x^**

(4.1)
$$\sum_{k \in P} \lambda_k \nabla \Phi^k(x^*) = 0$$

then x^* is a Pareto minimum.

Remark. It can be seen, by applying standard convexity arguments (as for instance in [13, p. 213 e.f.]), that the condition in Proposition 4.1 is sufficient that a point x^* be even properly efficient. Moreover, in view of Geoffrion's result [9, p. 620], above sufficient condition is also necessary for proper efficiency. (We recall that a Pareto minimal point x^* is properly efficient for the criteria Φ^k, $k \in P$ if there exists a scalar $\beta > 0$ such that for each $k \in P$, and $x \in R^n$ satisfying $\Phi^k(x) < \Phi^k(x^*)$, there exists at least one $\bar{k} \in P \setminus \{k\}$ with $\Phi^{\bar{k}}(x) > \Phi^{\bar{k}}(x^*)$ and $(\Phi^k(x^*) - \Phi^k(x)) / (\Phi^{\bar{k}}(x) - \Phi^{\bar{k}}(x^*)) \leqq \beta$, e.g. [8].) ■

In the comparison we will use the point-to-set mapping

$$E(x) = \{ \bigcap_{k \in P_=(x)} D_k(x) \}^+$$

appearing in (2.2). If Φ^k, $k \in P$ are convex analytic functions then $D_k(x)$, $k \in P$ are subspaces independent of x (see e.g. [4], [5]). Then we use the notation D_k rather than $D_k(x)$. However, $E(x)$ remains a function of x. The function $E(x)$ can be considered as a "measure of discrepancy" between Pareto and properly efficient solutions.

Our first comparison gives the following result.

4.2. Corollary. *Let the criteria Φ^k, $k \in P$ be convex and differentiable. If $E(x^*) = 0$ at some Pareto minimum x^*, then x^* is properly efficient.*

Proof. Since x^* is a Pareto minimum, we know that (2.2) in Corollary 2.2 holds. But $E(x^*) = 0$ implies that also (4.1) in Proposition 4.1 holds. The proper efficiency of x^* now follows from the above Remark. ■

A sufficient condition for $E(x^*) = 0$ is, for example, when each of the m programs (P, r) satisfies Slater's condition. (This does not imply that all m criteria Φ^k, $k \in P$ satisfy Slater's condition!) If at least one criterion Φ^k in the set $P^r_=$, for some r, is strictly convex then the other extreme situation occurs, i.e. $E(x^*) = R^n$.

Next we turn our attention to Karlin's *necessary* condition, also from [10]:

4.3. Proposition. *If the criteria Φ^k, $k \in P$ are convex and differentiable, and if x^* is a Pareto minimum, then there exist nonnegative numbers $\lambda_k \geqq 0$, $k \in P$, and not all zero, such that*

$$\sum_{k \in P} \lambda_k \nabla \Phi^k(x^*) = 0 . \tag{4.2}$$

Remark. The above statement characterizes the so-called "weakly efficient" minima. (A point x^* is not a weakly efficient minima if, and only if, there is a point x such that $\Phi^k(x) < \Phi^k(x^*)$ for every $k \in P$.)

The results will be illustrated by examples. Their purpose is to demonstrate that our conditions can identify some weakly efficient points as Pareto nonoptimal and that they can also identify some points which are not properly efficient as Pareto optimal. The familiar optimality conditions fail to make this identification.

4.4. Example. Consider the two criteria

$$\begin{aligned} \Phi^1 &= \max\{0, x^2 \operatorname{sgn} x\} \\ \Phi^2 &= (x-1)^2 . \end{aligned} \tag{4.3}$$

The set of Pareto minima is the closed interval between $x = 0$ and $x = 1$. First, we will demonstrate that the conditions from Propositions 4.1 and 4.3 are not successful at some selected points.

(i) Is $x^* = -1$ Pareto minimal? Since (4.1) is here

$$\lambda_1 \nabla \Phi^1(x^*) + \lambda_2 \nabla \Phi^2(x^*) = \lambda_1 . 0 + \lambda_2(-4) = 0$$

we note that Karlin's necessary condition is satisfied with $\lambda_1 > 0$ and $\lambda_2 = 0$, although the point is *not* optimal. (Note that $x^* = -1$ is weakly Pareto minimal.)

(ii) Is $x^* = 1$ Pareto optimal? The equation (4.1) is now

$$\lambda_1 . 2 + \lambda_2 . 0 = 0$$

which does not have a positive solution $\lambda_1 > 0$, $\lambda_2 > 0$. Thus Karlin's sufficient condition is not satisfied although the point *is* optimal. (Of course, since all the interior points $0 < x^* < 1$ are properly efficient, Karlin's condition holds at these points.)

4.5. Example. Consider again the criteria (4.3). This time we will use our conditions of optimality.

(i) Is $x^* = -1$ Pareto optimal? The necessary condition for optimality (2.2) is here

$$\lambda_1 \nabla\Phi^1(x^*) + \lambda_2 \nabla\Phi^2(x^*) \in \{D_1(x^*)\}^+$$
$$\lambda_1 > 0\,, \quad \lambda_2 > 0\,.$$

Since $D_1(x^*) = R$ and hence $\{D_1(x^*)\}^+ = \{0\}$, the above becomes

$$\lambda_1 \,.\, 0 + \lambda_2(-4) = 0$$
$$\lambda_1 > 0\,, \quad \lambda_2 > 0$$

clearly not satisfied. Conclusion: $x^* = -1$ is *not* Pareto optimal.

(ii) Is $x^* = 1$ Pareto optimal? Since $P^1_= = \{2\}$ and $P^2_= = \emptyset$, the optimality conditions (2.1) become

$$\text{(4.4)} \qquad \nabla\Phi^1(x^*) \in \{D_2(x^*)\}^+ = \{0\}^+ = R$$

and

$$\nabla\Phi^2(x^*) + \lambda_1 \nabla\Phi^1(x^*) = 0$$
$$\lambda_1 \geqq 0$$

i.e.

$$\text{(4.5)} \qquad 0 + 2\lambda_1 = 0$$
$$\lambda_1 \geqq 0\,.$$

Both systems (4.4) and (4.5) are clearly satisfied, so $x^* = 1$ *is* indeed Pareto optimal.

Conclusions. Using Theorem 2.1 and Corollary 2.2 we have established Pareto optimality of $x^* = 1$. Since $x^* = 1$ is not properly efficient, the optimality could not have been established here by Proposion 4.1. We have also established that the weakly Pareto optimal point $x^* = -1$ is not Pareto optimal. This could not have been established here by Proposition 4.3.

Example 4.6. The "discrepancy function" for the two criteria Φ^1 and Φ^2 is

$$E(x) = \begin{cases} 0 & \text{if } \; x < 0 \\ (-\infty, 0] & \text{if } \; x = 0 \\ 0 & \text{if } \; 0 < x < 1 \\ R & \text{if } \; x = 1\,. \end{cases}$$

In view of Corollary 4.2, this shows that Karlin's sufficient condition is also necessary for $x \in (-\infty, 1)$ except $x \neq 0$. It also confirms that the Pareto minima on the open interval $(0, 1)$ are properly efficient.

Our characterizations also establish connections between the three types of minima. To this end we will use the following lemma.

4.7. Lemma. *Let* Φ^k, $k \in P$ *be convex differentiable criteria. If* $\nabla\Phi^k(x^*) = 0$

for some $k \in P$, *then* $k \in P_{=}(x^*)$. *On the other hand, if* $k \in P_{=}(x^*)$ *and* $\operatorname{int} F^r(x^*) \neq \emptyset$ *for every* $r \in P^k$ *then* $\nabla\Phi^k(x^*) = 0$.

Proof. If $\nabla\Phi^k(x^*) = 0$ for some $k \in P$, then x^* minimizes Φ^k. Therefore, there is no x such that $\Phi^k(x) < \Phi^k(x^*)$, i.e.

$$x \in F^r(x^*) \quad \text{for any} \quad r \in P^k \Rightarrow \Phi^k(x) = \Phi^k(x^*)\,.$$

This means that $k \in P_{=}(x^*)$. On the other hand, if $k \in P_{=}(x^*)$ for some $k \in P$, then we must have $\nabla\Phi^k(x^*) = 0$. If not, then x^* is not optimal and we can find an x such that $\Phi^k(x) < \Phi^k(x^*)$. But Φ^k is constant over $\operatorname{int} F^r(x^*)$ for some $r \in P^k$. These two statements contradict convexity of Φ^k. ■

4.8. Corollary. *Let the criteria* Φ^k, $k \in P$ *be convex and differentiable. If at some point* x^* *all gradients of* Φ^k, $k \in P$ *are different from zero and* $\operatorname{int} F^k(x^*) \neq \emptyset$, $k \in P$, *then* x^* *is Pareto optimal if, and only if, it is weakly Pareto optimal.*

Proof. The assumptions on gradients and interior points imply $P_{=}(x^*) = \emptyset$, by Lemma 4.7. This means that the polar set in (3.1) is zero. (The intersection over the empty set has been defined to be the whole space.) But now (4.2) and (3.1) coincide.

The above result, in particular, says, that if a weakly efficient point x^* is not Pareto optimal, then the gradient of at least one criterion must vanish at x^* or at least one $F^k(x^*)$, $k \in P$ must have an empty interior.

When combined with Corollary 4.2, the above corollary gives a condition when all three points: Pareto minimum, properly efficient point, ano weakly efficient point coincide. A stronger condition for that coincidence is that the constraints determining $F^k(x^*)$, $k = 1, \ldots, m$ satisfy Slater's condition.

References

[1] *R. Abrams, L. Kerzner:* A simplified test for optimality. Journal of Optimization Theory and Applications 25 (1978), 161–170.

[2] *A. Ben-Israel:* Linear equations and inequalities on finite dimensional, real or complex, vector spaces: A unified theory. Journal of Mathematical Analysis and Applications 27 (1969), 367–389.

[3] *A. Ben-Israel, A. Ben-Tal, A. Charnes:* Necessary and sufficient conditions for Pareto optimum in convex programming. Econometrica 45 (1977), 811–820.

[4] *A. Ben-Israel, A. Ben-Tal, S. Zlobec:* Optimality in Nonlinear Programming: A Feasible Directions Approach. Wiley-Interscience, New York, 1981.

[5] *A. Ben-Tal, A. Ben-Israel, S. Zlobec:* Characterization of optimality in convex programming without a constraint qualification. Journal of Optimization Theory and Applications 20 (1976), 417–437.

[6] *G. R. Bitran, T. L. Magnanti:* The structure of admissible points with respect to cone dominance. Journal of Optimization Theory and Applications 29 (1979), 473–514.

[7] *Y. Censor:* Pareto optimality in multiobjective problems. Applied Mathematics and Optimization 4 (1977), 41–59.

[8] *A. Charnes, W. W. Cooper:* Management Models and Industrial Applications of Linear Programming, Vol. I. Wiley, New York, 1961.

[9] *A. M. Geoffrion:* Proper efficiency and the theory of vector maximization. Journal of Mathematical Analysis and Applications 22 (1968), 618—630.

[10] *S. Karlin:* Mathematical Methods and Theory in Games. Programming and Economics, Vol. I, Addison-Wesley, Reading, Massachussetts, 1959.

[11] *V. V. Podinovskii:* Applying the procedure for maximizing the basic local criterion to solving the vector optimization problems. Systems Control 6, Novosibirsk, 1970 (In Russian).

[12] *V. V. Podinovskii, V. M. Gavrilov:* Optimization with Respect to Successive Criteria, Soviet Radio, Moscow, 1975 (In Russian).

[13] *R. T. Rockafellar:* Convex Analysis. Princeton University Press, 1970.

[14] *M. E. Salukvadze:* Vector-Valued Optimization Problems in Control Theory. Academic Press, New York, 1979.

[15] *S. Smale:* Global analysis and economics III. Journal of Mathematical Economics 1 (1974), 107—117.

[16] *S. Smale:* Global analysis and economics V. Journal of Mathematical Economics 1 (1974), 213—221.

[17] *S. Smale:* Global analysis and economics VI. Journal of Mathematical Economics 3 (1976), 1—14.

[18] *S. Zlobec:* Regions of stability for ill-posed convex programs. Aplikace Matematiky 27 (1982), 176—191.

Souhrn

DVĚ CHARAKTERIZACE PARETOVÝCH MINIM V KONVEXNÍ OPTIMIZACI S VÍCE KRITÉRII

SANJO ZLOBEC

Jsou udány dvě podmínky, každá z nich je nutná a postačující, aby daný bod byl globálním Paretovým minimem. První byla získána studiem programů, v nichž se každé kritérium objevuje jako jediná cílová funkce, zatím co druhá je dána ve tvaru „restringovaného Lagrangiánu". Podmínky jsou porovnány se známými charakterizacemi skutečných a slabých Karlinových a Geoffrionových bodů.

Author's address: Prof. *Sanjo Zlobec*, Department of Mathematics and Statistics, Burnside Hall, McGill University, Montreal, Quebec, Canada H 3A 2 K 6.