A Theory of Lexicographic Multi-Criteria Optimization*

Mark J. Rentmeesters
Dept. of Info. and Compt. Sci.
University of California Irvine
Irvine, CA 92697
mrentme@ics.uci.edu

Wei K. Tsai and Kwei-Jay Lin
Dept. of Elect. and Compt. Eng.
University of California Irvine
Irvine, CA 92697
{wtsai,klin}@ece.uci.edu

Abstract

The field of multi-criteria optimization is reviewed as it pertains to lexicographic optimization over real-valued vector spaces. How lexicographic optimization differs from multi-criteria optimization that is restricted to proper Pareto optima is explained. Through a survey of previous work, it is revealed that there are currently no generally applicable methods for solving lexicographic optimization problems, and it is explained that this is due to the lack of an adequate mathematical theory for such problems. A more adequate mathematical theory is then presented for lexicographic optimization in this paper.

1. Introduction

The engineering of large-scale real-time systems is an area characterized by increasing complexity and risks. To meet these challenges, we need to investigate the theoretical foundations of complex systems. Complex systems must be designed and built on a solid and powerful theoretical foundation so that the complexity can be adequately managed. Any system conceptually dependent upon ad hoc techniques rather than a solid theoretical foundation will clearly be subject to considerably greater complexity explosion and failures.

In this paper, we address optimization issues involved in the engineering of complex systems. Specifically, we study and develop formal, mathematical, multi-criteria optimization models. These models are applicable to a large variety of complex systems engineering problems, including cost, performance, reliability, and maintainability. They can be directly employed to develop efficient, reduced-complexity optimization tools. It is through the application of this clarifying, mathematical formalism that the problem of system complexity is addressed.

In general, multi-criteria optimization is useful whenever there are more than one objective criterion of interest. If the value of improving any one objective criteria can be meaningfully compared to the value of improving any other objective criteria, then the problem can usually be reduced to one of single objective optimization through the use of some sort of weighted averaging. Such problems then involve the identification of proper Pareto optima for the original, multi-criteria formulation. Lexicographic optimization, on the other hand, is a form of multi-criteria optimization in which the various objectives under consideration cannot be quantitatively traded off between each other, at least not in a meaningful and numerically tractable way.

2. Previous work

A very comprehensive survey of both the theory and the application of multi-criteria optimization as it is used in engineering and the sciences can be found in Stadler [8]. In particular, this work points out the distinction, from a theoretical perspective, between lexicographic optima and proper Pareto optima. Saber [6] summarizes some of the basic algorithmic approaches employed in goal programming, including nonlinear goal programming. This survey also describes several diverse applications of nonlinear goal programming. Eschenauer et al. [2], although not addressing lexicographic problems, contains several concrete approaches to proper Pareto optimization and contains many diverse applications of multi-criteria optimization in engineering.

Although theoretical studies of general Pareto optima, and various forms of proper Pareto optima are profuse, (see Sawaragi, et al. [7] for a comprehensive survey), very little theoretical work has been done concerning the mathematical description of lexicographic optima that would facilitate the automatic solution of such problems, except for the case of goal programming, in which all objectives functions are

[1] This work was supported in part by contracts from the Office of Naval Research N00014-94-1-0034 and N00014-95-1-0262, and the US Navy NSWCDD N60921-94-M-2714 and N60921-95-C-0041.

0-8186-7614-0/96 $5.00 © 1996 IEEE

linear. When the objective equations themselves are nonlinear, moreover, solution techniques suitable for automation are rare.

To be of utility in the automatic solution of lexicographic optima, a theory must provide explicit, objectively verifiable conditions discriminating lexicographic optima from nonoptimal points. Satisfying this criterion, Luptacik and Turnovec [5] have developed equations identifying that subset of solution points for ordinary lexicographic optimization problems at which all the objectives under consideration can be simultaneously satisfied. However, most applications of lexicographic optimization involve objectives that must be traded off one for another and, moreover, for which it is usually already known that there are no points at which all the objectives can be simultaneously satisfied. Thus, this theory is not as yet generally applicable.

It is not surprising that very little has been done in the way of efficient algorithm development for lexicographic optimization, for without sufficiently concrete, operational descriptions with which to identify the points of optimality it is hard to develop any sort of solution mechanism at all. That is to say, lack of an adequate theory for lexicographic optimization up to this point has prevented the development of good algorithms for such problems. An adequate theory will be presented in the next section.

3. Lexicographic optimization

3.1. Introduction.

Given a sequence $f_i, i = 1, \ldots, m$ of objective functions, and a solution $x^* \in \mathbf{R}^n$ which *lexicographically* minimizes the vector valued function $[f_1, \ldots, f_m]$, it is then by definition the case that x^* minimizes f_1, that it minimizes f_2 under the constraining condition that $f_1(x) = f_1(x^*)$, and that, in general, for $i = 3, \ldots, m$, x^* minimizes f_i under the constraining conditions that $f_1(x) = f_1(x^*), \ldots, f_{i-1}(x) = f_{i-1}(x^*)$. A straightforward but naive conclusion might then be that, as each of these m minimization conditions satisfied by x^* is nothing more than a simple, single-objective constrained (or, in the case of $i = 1$, unconstrained) optimization, it would then follow that x^* might also satisfy, in most situations, the classically described Kuhn-Tucker necessary conditions for each of these simple, single-objective optimization problems. In particular, one might presume, since x^* minimizes f_1, that $\nabla f_1(x^*) = 0$, and, for $i = 2, \ldots, m$, that there exist $\lambda_{ij} \in \mathbf{R}$, such that

$$\nabla_x \left[\sum_{j=i}^{i-1} \lambda_{ij} [f_j(x) - f_j(x^*)] + f_i(x) \right] \Bigg|_{x=x^*} = 0$$

as well.

However, that such is not the case in general is easily demonstrated. Although indeed it will be true that $\nabla f_1(x^*) = 0$, as long as f_1 is differentiable at x^*, if it were true that there was a $\lambda_{21} \in \mathbf{R}$ such that

$$\begin{aligned} 0 &= \nabla_x \left[\lambda_{21}(f_1(x) - f_1(x^*)) + f_2(x) \right] \big|_{x=x^*} \\ &= \left[\lambda_{21} \nabla f_1(x) + \nabla f_2(x) \right] \big|_{x=x^*} \\ &= \lambda_{21} \nabla f_1(x^*) + \nabla f_2(x^*) \\ &= \lambda_{21} \cdot 0 + \nabla f_2(x^*) \\ &= \nabla f_2(x^*), \end{aligned}$$

then clearly λ_{21} is arbitrary and may be taken to be zero. Having $\nabla f_2(x^*) = 0$, moreover, would imply that f_2 in fact achieves its absolute unconstrained minimum at x^*. It is also clear that this argument can be applied inductively, so that for all $i = 2, \ldots, m$, it would necessarily follow, if the classical Kuhn-Tucker conditions stated above were assumed to hold, that each $\lambda_{ij} = 0, j = 1, \ldots, i-1, i = 2, \ldots, m$, and each $\nabla f_i(x^*) = 0, i = 2, \ldots, m$, and therefore that the only lexicographic optimization solution points at which these classical Kuhn-Tucker conditions could possibly be valid are those in which each objective, $f_1, \ldots, f_m$, achieves the same minimum that would be obtained had it been optimized independently.

Although there are certainly applications of multi-objective optimization where it is known in advance that it is possible to simultaneously achieve the absolute minimum of each objective involved, such applications are not properly the concern of lexicographic optimization. For, when it is known that a number of objectives, $f_1, \ldots, f_m$, can be simultaneously satisfied, then a ranking or prioritization of them is immaterial; any ordering among them can achieve the same results as any other. It is only when the simultaneous satisfaction of all objectives cannot be achieved, when there is conflict between them, that it is necessary, or at least worth while, to consider a lexicographic optimization approach as a way to resolve the conflicts. Under these conditions, as has just been demonstrated, the classical Kuhn-Tucker necessary conditions for optimality do not apply.

Throughout this and the remaining sections, for any positive integers n and m and any $I \subset \{1, \ldots, m\}$, let $\mathcal{C}^{1,2}_{n,m}(x^*, I)$ denote the space of all functions $F = [f_1, \ldots, f_m] : \mathbf{R}^n \rightarrow \mathbf{R}^m$ such that f_i is twice continuously differentiable at x^* if $i \in I$, and once continuously differentiable at x^* otherwise.

3.2. Necessary conditions.

The classical Kuhn-Tucker necessary conditions for single-objective optimality may be stated as follows, (see [3]).

Theorem 1 (Kuhn Tucker) *If the functions f, $g_i, i = 1\ldots, m_g$, and $h_j, j = 1, \ldots, m_h$ are differentiable at x^*, and if the constraint functions $g_i, i = 1\ldots, m_g$, and $h_j, j = 1, \ldots, m_h$ satisfy a suitable constraint qualification criteria, Q, at x^*, then necessary conditions for x^* to be a local minimum of f under the constraining conditions that $g_i(x) \leq 0, i = 1, \ldots, m_g$ and $h_i(x) = 0, i = 1, \ldots, m_h$ are that there exist $\lambda_i, i = 1, \ldots, m_g$ and $\mu_j, j = 1, \ldots, m_h$ such that*

$$\begin{aligned} g_i(x^*) &\leq 0, i = 1, \ldots, m_g, \\ h_j(x^*) &= 0, j = 1, \ldots, m_h, \\ \lambda_i g_i(x^*) &= 0, i = 1, \ldots, m_g, \\ \lambda_i &\leq 0, i = 1, \ldots, m_g, \\ 0 &= \nabla f(x^*) + \sum_{i=1}^{m_g} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{m_h} \mu_j \nabla h_j(x^*). \end{aligned}$$

Although the constraint qualification that is employed as Q in theorem (1) may vary (see for example Bazaraa, et al. [1]), it will be sufficient for the purposes of this paper to consider but one example. The constraint qualification that will be taken as exemplary is that qualification presented originally by Kuhn and Tucker (see Fiacco and McCormick [3]).

Definition 1 (Constraint qualification Q [Kuhn Tucker]) *Let x^* be a point satisfying $g_i(x^*) \leq 0, i = 1, \ldots, m_g$ and $h_j(x^*) = 0, i = j, \ldots, m_h$, and assume that the functions $g_i(x^*), i = 1, \ldots, m_g$ and $h_j(x^*), i = j, \ldots, m_h$, are all once differentiable at x^*. Then the constraint qualification Q holds at x^* if for any non-zero vector z such that $z^T \nabla g_i(x^*) \leq 0$ for all $i \in B^* = \{i : g_i(x^*) = 0, 1 \leq i \leq m_g\}$, and $z^T \nabla h_j(x^*) = 0, j = 1, \ldots, m_h$, then z is tangent to a once-differentiable arc, the arc emanating from x^* and contained in the constraint region defined by $\{x : g_i(x) \leq 0, i = 1, \ldots, m_g, h_j(x) = 0, j = 1, \ldots, m_h\}$.*

It is worth noting that the reason the classical Kuhn-Tucker conditions fail to hold for all but those trivial, non-conflicting lexicographic optimality points is precisely because any other points fail to satisfy the constraint qualification that is imposed on the applicability of these conditions. That is, the more interesting conflict-resolving optimality points must of necessity fail to meet any constraint qualification that would then ensure that these classical Kuhn-Tucker conditions held. This is indeed seen to hold, for example, with the qualification criteria Q given above. At a lexicographic optimum point x^* for a simple sequence of objectives, $f_i, i = 1, 2$, consider the constraint $f_1(x) = f_1(x^*)$. Since $\nabla f_1(x^*) = 0$, any nonzero vector $z \in \mathbf{R}^n$ whatsoever will satisfy $z^T \nabla f_1(x^*) = 0$. Thus in order for the constraint qualification Q to hold at x^* for the constraint $f_1(x) - f_1(x^*) = 0$, it would be necessary for any once-differentiable arc $s : \mathbf{R} \rightarrow \mathbf{R}^n$ whatsoever passing through x^* to satisfy $f_1(s(r)) = f_1(x^*), r \in [0, t)$ for some $t \geq 0$. This then would again imply that, within a local neighborhood of x^*, the objective f_1 would be constant and thus neither constraining nor conflicting with f_2 or any other objective.

Thus it can be seen that whatever may be the features of a set of more general necessary conditions for characterizing lexicographic optima, these conditions will need to rely on a less restrictive class of constraint qualification criteria than those normally imposed in order that the classical Kuhn-Tucker conditions may be applied.

Given these insights concerning the extent of applicability of the classical Kuhn-Tucker necessary conditions for optimality, as typically applied, it will be possible to show, in what follows, that the more general necessary conditions to be presented, those that will be applicable even to conflict-resolving lexicographic optima, will follow from, in fact, little more than the very same theorem from which the classical conditions themselves follow. But before stating the theorem itself, it is worth pointing out, as Luenberger [4] has done, that constraint qualifications such as Q above apply to a given *representation* of a constraint region, specifically, to one expressed in terms of the constraint functions, $g_i, i = 1, \ldots, m_g$ and $h_j, j = 1, \ldots, m_h$ given and not to the constraint region itself. It is possible for the constraint qualification Q to fail at an optimality point x^* for one representation $g_i, i = 1, \ldots, m_g$ and $h_j, j = 1, \ldots, m_h$, but hold (and consequently the Kuhn-Tucker necessary conditions of theorem (1) as well) for a different representation, $\hat{g}_i, i = 1, \ldots, \hat{m}_g$ and $\hat{h}_j, j = 1, \ldots, \hat{m}_h$ of same constraint region. In particular, for lexicographic optima, it will be observed below, that a constraining condition $f_i(x) \leq f_i(x^*)$ can be augmented with the condition that $\nabla f_i(x) = \nabla f_i(x^*)$ without nullifying optimality.

It is in fact just such a simple method of re-representation that is the basis for the more general form of necessary conditions to follow. The generalized version of constraint qualification criteria Q that will be employed in this generalized condition considers the addition of augmenting constraint functions drawn from the first derivatives of the original constraint functions, as exemplified in the preceding paragraph for lexicographic optimization.

Definition 2 (Constraint qualification Q') *Given an $x^* \in \mathbf{R}^n$, $I \subseteq \{1, \ldots, m_g\}$, $G = [g_1, \ldots, g_{m_g}]^T \in \mathcal{C}^{1,2}_{n,m_g}(x^*, I)$, $J \subseteq \{1, \ldots, m_h\}$, and $H = [h_1, \ldots, h_{m_h}]^T \in \mathcal{C}^{1,2}_{n,m_h}(x^*, J)$, let x^* be a point satisfying $g_i(x^*) \leq 0, i = 1, \ldots, m_g$ and $h_j(x^*) = 0, i = j, \ldots, m_h$. Then the constraint qualification $Q'(I, J)$ holds at x^* if for any non-zero vector z such that $z^T \nabla g_i(x^*) \leq 0$ for all $i \in B^* = \{i : g_i(x^*) = 0\}$, $\nabla^2 g_i(x^*) z = 0$ for all $i \in B^* \cap I$, $z^T \nabla h_j(x^*) = 0$ for all $j = 1, \ldots, m_h$, and $\nabla^2 h_j(x^*) z = 0$*

for all $j \in J$, then z is tangent to a once-differentiable arc, the arc emanating from x^* and contained in the constraint region defined by $\{x : g_i(x) \leq 0, i = 1, \ldots, m_g, h_j(x) = 0, i = j, \ldots, m_h, \}$.

Note that $Q'(\emptyset, \emptyset) = Q$, and that if $I_1 \subseteq I_2 \subseteq \{1, \ldots, m_g\}$ and $J_1 \subseteq J_2 \subseteq \{1, \ldots, m_h\}$ then any point x^* at which constraint qualification $Q'(I_1, J_1)$ holds, constraint qualification $Q'(I_2, J_2)$ holds as well. Thus a constraint qualification $Q'(I, J)$ for any $I \neq \emptyset$, or $J \neq \emptyset$ is a weaker, more permissive constraint qualification than Q, possibly admitting more points than Q.

The application of the more complex theory embodied in the extension of theorem (1) to employ qualification Q' to the problem of lexicographic optimization, where, as has been shown, the simpler theory fails to apply, will now be described.

Definition 3 *For a given $x^* \in \mathbf{R}^n$, $I \subset \{1, \ldots, m\}$, and $F = [f_1, \ldots, f_m]^T \in \mathcal{C}^{1,2}_{n,m}(x^*, I)$ the point x^* is said to be a lexicographic regular point of the sequence $f_i, i = 1, \ldots, m$ with respect to I if for for each $i = 2, \ldots, m$, the following holds. For any non-zero vector $z \in \mathbf{R}^n$ such that $z^T \nabla f_j(x^*) = 0$, $j = 1, \ldots, i-1$, and $\nabla^2 f_j(x^*) z = 0$, $j \in I \cap \{1, \ldots, i-1\}$, z is tangent to a once-differentiable arc, the arc emanating from x^* and contained within the constraint region defined by $\{x : f_j(x) = f_j(x^*), j = 1, \ldots, i-1\}$.*

It is readily identified that if x^* is a lexicographic regular point for the sequence $f_i, i = 1, \ldots, m$ with respect to a set I, then constraint qualification $Q'(\emptyset, I \cap \{1, \ldots, k\})$ holds at x^* for each of the constraint sequences $\{f_1, \ldots, f_k\}, k = 1, \ldots, m-1$.

With this definition, it is now possible to apply the classical Kuhn-Tucker conditions to lexicographic optimization as follows.

Theorem 2 (Lexicographic Kuhn-Tucker) *Given $x^* \in \mathbf{R}^n$, $I \subset \{1, \ldots, m\}$, and $F = [f_1, \ldots, f_m]^T \in \mathcal{C}^{1,2}_{n,m}(x^*, I)$, if x^* is a lexicographic regular point for $f_1, \ldots, f_m$ with respect to I, then necessary conditions for x^* to be a local lexicographic minimum of the vector valued function F are that there exist numbers $\lambda_{ij} \in \mathbf{R}$, $\lambda_{ij} \geq 0$, $j = 1, \ldots, i-1, i = 2, \ldots, m$ and vectors $\mu_{ij} \in \mathbf{R}^n$, $j \in I_i = I \cap \{1, \ldots, i-1\}, i = 2, \ldots, m$ such that*

$$\nabla f_1(x^*) = 0, \tag{1}$$

$$\nabla f_i(x^*) + \sum_{j=1}^{i-1} \lambda_{ij} \nabla f_j(x^*) + \sum_{j \in I_i} \nabla^2 f_j(x^*) \mu_{ij} = 0, i = 2, \ldots, m. \tag{2}$$

It is readily verified that the conditions (1) and (2) of theorem (2) will hold at lexicographic minimum points even when these points demonstrate inter-objective conflicts, so long as the optimum point x^* is now a lexicographic regular point with respect to an index set I which includes the index of at least the higher ranking of each pair of conflicting objectives. Thus theorem (2) presents a characterization of lexicographic optima that is applicable to genuinely lexicographic problem formulations, that is, those in which inter-objective conflict exists and is to be resolved by prioritizing these objectives.

4. Conclusions and Further Research.

A mathematical theory applicable to lexicographic optimization has been presented. With this theory, it is now possible to describe lexicographic optima in a concrete way, namely through the use of conditions, equations of the Kuhn-Tucker type, that are satisfied at these points.

These equations are similar to the long standing and well known Kuhn-Tucker conditions for simple, single objective optimization. Hence, it should be possible to develop algorithms for lexicographic optimization that proceed by solving the new Kuhn-Tucker conditions for lexicographic optimization by modifying or extending algorithms that have already been developed. In particular, the very efficient gradient-based methods and Newton-based methods should be extendible to solution of the new Kuhn-Tucker conditions.

References

[1] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming: theory and algorithms*. John Wiley & Sons, Inc., New York, 1993.

[2] H. Eschenauer, J. Loki, and A. Osyczka, editors. *Multicriteria design optimization : procedures and applications*. Springer-Verlag, Berlin; New York, 1990.

[3] A. V. Fiacco and G. P. McCormick. *Nonlinear programming: sequential unconstrained minimization techniques*. Society for Industrial and Applied Mathematics, Philadelphia, 1990.

[4] D. G. Luenberger. *Linear and Nonlinear Programming, 2nd. Ed.* Addison Wesley, Reading, Mass., 1984.

[5] M. Luptacik and F. Turnovec. Lexicographic geometric programming. *European Journal of Operational Research*, 51(2):259–269, 1991.

[6] H. M. Saber and A. Ravindran. Nonlinear goal programming theory and practice: a survey. *Computers and Operations Research and their Applications*, 20(3):275–291, 1993.

[7] Y. Sawaragi, H. Nakayama, and T. Tanino. *Theory of Multiobjective Optimization*. Academic Press, Orlando, Florida, 1985.

[8] W. Stadler, editor. *Multicriteria optimization in engineering and in the sciences*. Plenum Press, New York, 1988.