Vilfredo Pareto and Multi-objective Optimization

Matthias Ehrgott

2010 Mathematics Subject Classification: 90C29
Keywords and Phrases: Multi-objective optimization, Pareto optimality

A multi-objective optimization problem consists in the simultaneous optimization of p objective functions $f_1, \dots, f_p$ subject to some constraints, which I will just write as $x \in \mathcal{X}$, where $\mathcal{X}$ is a subset of $\mathbb{R}^n$. It is usually assumed that there does not exist any $x \in \mathcal{X}$ such that all functions f_k attain their minimima at x. Hence, due to the absence of a total order on $\mathbb{R}^p$, it is necessary to define the minimization with respect to partial orders. So let $\mathcal{Y} := \{f(x) : x \in \mathcal{X}\}$ be the set of outcome vectors. To compare elements of $\mathcal{Y}$, I will follow the definition of Koopmans (1951). Let $y^1, y^2 \in \mathcal{Y}$. Then $y^1 \leqq y^2$ if and only if $y^1_k \leqq y^2_k$ for all $k = 1, \dots p$; $y^1 \leq y^2$ if and only if $y^1 \leqq y^2$, but $y^1 \neq y^2$ and $y^1 < y^2$ if and only if $y^1_k < y^2_k$ for all $k = 1, \dots p$.

It is here that Pareto makes his appearance. In countless books and articles on multi-objective optimization, one can find a definition like this:

Definition 1. Let $\mathcal{X} \subset \mathbb{R}^n$ be a non-empty set of feasible solutions and $f = (f_1, \dots f_p) : \mathbb{R}^n \to \mathbb{R}^p$ be a function. Feasible solution $\hat{x} \in \mathcal{X}$ is called a *Pareto optimal* solution of the multi-objective optimization problem

$$\min\{f(x) : x \in \mathcal{X}\} \tag{1}$$

if and only if there does not exist any $x \in \mathcal{X}$ such that $f(x) \leq f(\hat{x})$.

Sometimes Pareto optimality is defined with respect to outcome vectors.

Definition 2. Let $\mathcal{Y} \in \mathbb{R}^p$ be a non-empty set of outcome vectors. Outcome vector $\hat{y} \in \mathcal{Y}$ is called *Pareto optimal* if and only if there does not exist any $y \in \mathcal{Y}$ such that $y \leq \hat{y}$.

Where does the name Pareto optimal come from? Vilfredo Pareto and Francis Ysidro Edgeworth are often called as the fathers of multi-objective optimization. Sentences like the "introduction of the Pareto optimal solution in 1896" (Chen et al., 2005, p. VII); "The concept of noninferior solution was introduced at the turn of the century [1896] by Pareto, a prominent economist" (Chankong and Haimes, 1983, p. 113); "Edgeworth and Pareto were probably the first who introduced an optimality concept for such problems" (Jahn, 2004, p. 113); "wurden besonders von F.Y. Edgeworth (1845–1926) and V. Pareto (1848–1929 [sic!]) hinreichende Bedingungen für Paretomaximalität bzw. Gleichgewichtsbedingungen angegeben." (Göpfert and Nehse, 1990, p. 9) or "The foundations are connected with the names of Vilfredo Pareto (1848–1923) and Francis Ysidro Edgeworth (1845–1926)" (Löhne, 2011, p. 1) abound in textbooks. The International Society on Multiple Criteria Decision Making bestows the Edgeworth–Pareto award "upon a researcher who, over his/her career, has established a record of creativity to the extent that the field of MCDM would not exist in its current form without the far-reaching contributions from this distinguished scholar", see `http://www.mcdmsociety.org/intro.html#Awards`.

Edgeworth was an influential Professor of Economics at King's College London and from 1891 Professor of Political Economy at Oxford University. In his best known book *Mathematical Psychics* (Edgeworth, 1881) he applied formal mathematics to decision making in economics. He developed utility theory, introducing the concept of indifference curve and is best known for the *Edgeworth box*. But because multi-objective optimization is concerned with Pareto optimality rather than Edgeworth optimality, this story focuses on his contemporary.

Fritz Wilfried Pareto

According to Yu (1985, p. 49) Pareto "was a famous Italian engineer" but he is certainly much better known as an economist. The following information is taken from Stadler (1979) and the wikipedia entry (`http://en.wikipedia.org/wiki/Vilfredo_Pareto`) on Pareto.

Vilfredo Federico Damaso Pareto was born on 15 July 1848 in Paris as Fritz Wilfried Pareto, son of a French woman and an Italian civil engineer, who was a supporter of the German revolution of 1848. His name was changed to the Italian version when his family moved back to Italy in 1855 (or 1858). In 1870 he graduated from Polytechnic Institute of Turin with a dissertation entitled "The Fundamental Principles of Equilibrium in Solid Bodies". He then worked as an engineer and manager for an Italian railway company. He was very politically active, an ardent supporter of free market economy. He obtained a lecturer position in economics and management at the University of Florence in 1886 (according to wikipedia). Eventually he resigned from his positions in 1889. During the 1880s he became acquainted with leading economists of the time and he published many articles by 1893 (not all academic, though). In 1893 he moved to Lausanne where he lectured at the University of Lausanne and became the successor of Léon Walras as Professor of Political Economy. In his later years he mainly worked in Sociology. Vilfredo Pareto died at Célégny, Switzerland, on 19 August 1923. The University of Lausanne still has a Centre d'études interdisciplinaires Walras Pareto (`http://www.unil.ch/cwp`). Apart from Pareto optimality, Pareto's name is attached to the Pareto principle (or 80–20 rule), observing in 1906 that 80% of the property in Italy was owned by

Figure 1: Vilfredo Pareto 1848–1923 (Picture scanned from the second French edition of Pareto (1906) published in 1927.)

20 % of the population and the Pareto distribution, a power law probability distribution.

Pareto Optimality

The origin of the term Pareto optimality goes back to the following text from Pareto (1906, Chapter VI, Section 33).

> Principeremo col definire un termine di cui è comodo fare uso per scansare lungaggini. Diremo che i componenti di una collettività godono, in una certa posizione, del massimo di ofelimità, quando è impossibile allontanarsi pochissimo da quella posizione giovando, o nuocendo, a tutti i componenti la collettività; ogni piccotissimo spostamento da quella posizione avendo necessariamente per effetto di giovare a parte dei componenti ta collettività e di nuocere ad altri.

Or in the English translation (Pareto, 1971, p. 261):

> We will begin by defining a term which is desirable to use in order to avoid prolixity. We will say that the members of a collectivity enjoy *maximum ophelimity* in a certain position when it is impossible to find a way of moving from that position very slightly in such a manner that the ophelimity enjoyed by each of the individuals of that collectivity increases or decreases. That is to say, any small displacement in departing from that position necessarily has the effect of increasing the ophelimity which certain individuals enjoy, and decreasing that which others enjoy, of being agreeable to some and disagreeable to others.

Of course, Pareto here refers to the distribution of utility (ophelimity) among individuals in an economy rather than solutions of an optimization problem. Multi-objective optimization or mathematical optimization in general as we know it today, did not exist during Pareto's lifetime, it only developed in the 1940s. And it is some of the founding works of Operations Research and optimization that need to be cited here. Nobel Laureate in Economics T.C. Koopmans (1951) formally studied production as a resource allocation problem and the combination of activities to represent the output of commodities as a function of various factors. In this work he introduced the following definition of efficient vector (p. 60). "A point y in the commodity space is called *efficient* if it is *possible* [i.e., if $y \in (A)$], and if there exists no possible point $\bar{y} \in (A)$ such that $\bar{y} - y \geq 0$." Note that (A) is what I called $\mathcal{Y}$ in Definition 2, i.e., *possible* means that there is some x such that $y = Ax$. Koopmans does hence only talk about efficient vectors in terms of the outcome set. He proves necessary and sufficient conditions for efficiency, but he does not refer to Pareto, nor does he talk about Pareto optimal solutions as in Definition 1 – instead he refers to "an activity vector x (that) shall lead to an efficient point $y = Ax$".

Another classic reference in optimization is the seminal paper by Kuhn and Tucker (1951). They refer to the "vector maximum of Koopmans' efficient point type for several concave functions $g_1(x), \ldots, g_p(x)$". This seems to be the earliest reference to the optimization of several functions in mathematics. Kuhn and Tucker cite Koopmans (1951) when they talk about vector maximum. They also apply the term *efficient* to the solutions of vector optimization problems (i.e., in decision space) and introduce the notion of proper efficiency. But, again, no mention of Pareto. Kuhn and Tucker (1951) cite another Nobel Laureate in Economics who contributed to the foundations of multi-objective optimization, Kenneth J. Arrow.

Arrow discusses Pareto extensively in his economical work and statements of the impossibility theorem today usually refer to Pareto optimality as one of the axioms that cannot be jointly satisfied by a social choice function, but this term does not appear in Arrow's original formulation (Arrow, 1951). Arrow's important contribution to multi-objective optimization (Arrow et al., 1953) starts as follows "A point s of a closed convex subset S of k-space is *admissible* if there is no $t \in S$ with $t_i \leq s_i$ for all $i = 1, \ldots, k$, $t \neq s$." This is, of course, the same as

Koopmans' definition of efficient point (whose paper Arrow et al. (1953) cite), and again is relevant in the outcome set of a multi-objective problem rather than the set of feasible solutions – no trace of Pareto here, either.

There are a number of other definitions of Pareto optimal, efficient, or admissible points. Zadeh (1963) defines "A system $S_0 \in \mathcal{C}$ is *noninferior* in $\mathcal{C}$ if the intersection of $\mathcal{C}$ and $\Sigma_>(S_0)$ is empty." $\Sigma_>(S_0)$ is the set of all systems which are better than S_0 with respect to a partial order $\geq$. Chankong and Haimes (1983) later use the same definition. While Zadeh cites Koopmans and Kuhn and Tucker, Pareto remains notably absent. The final term that is common today is that of a *nondominated* point.

Multiobjective Optimization and Economics

When did the term *Pareto optimal* first appear in the literature? As we have seen, it was not used in early mathematical works on multi-objective optimization. The answer is once again in economics. Little (1950, p. 87) in a discussion of the distribution of income (i.e., in the same context as Pareto himself) uses the term Pareto 'optimum' (with the quotation marks). The origin of the term is, therefore, clearly found in economics. It has then apparently mostly been used in economics, appearing in journals such as *Public Choice* and *Journal of Economic Theory*. As shown above, it was not used by the economists who are credited with having contributed to the origins of the mathematical theory of multi-objective optimization, but migrated to mathematics later on. The first journal articles that I could find are Basile and Vincent (1970) and Vincent and Leitmann (1970). These articles also used the term *undominated* as an alternative. This then turned to nondominated in Yu and Leitmann (1974).

Economics had a strong influence on the early history of multi-objective optimization, especially Pareto's original definition of the term *maximum ophelimity* and the origin of the term Pareto optimum in Little (1950). The move into mathematics and optimization coincides with the mathematization of economics by scholars such as Koopmans and Arrow and finally the introduction of the topic into mathematical optimization by Kuhn and Tucker. It seems to have taken quite a while for Pareto's name to appear in the mathematical optimization literature.

The consequence of the history of Pareto optimality outlined above, is that at present there are quite a few terms (efficient, noninferior, nondominated, admissible, Pareto optimal) that express the same idea. Since multi-objective optimization often distinguishes between decision vectors $x \in \mathcal{X}$ and outcome vectors $y \in \mathcal{Y}$, one can find a large number of combinations of these terms in the literature used in parallel today, such as Pareto optimal decisions and efficient outcomes.

It turns out that the history of multi-objective optimization (vector optimization) is quite an interesting read, and I would like to refer interested readers to Stadler (1979) as a starting point. The history of multiple criteria decision making in general is the topic of the book Köksalan et al. (2011). These works also consider roots of multi-objective optimization in game theory and the theory of ordered spaces and vector norms.

References

K. J. Arrow. *Social Choice and Individual Values*. Cowles Commission for Research in Economics Monograph No. 12. John Wiley & Sons, New York, 1951.

K. J. Arrow, E. W. Barankin, and D. Blackwell. Admissible points of convex sets. In H.W. Kuhn and A.W. Tucker, editors, *Contributions to the Theory of Games*, volume 2, pages 87–91. Princeton University Press, Princeton, 1953.

G. Basile and T. L. Vincent. Absolutely cooperative solution for a linear, multiplayer differential game. *Journal of Optimization Theory and Applications*, 6:41–46, 1970.

V. Chankong and Y. Y. Haimes. *Multiobjective Decision Making – Theory and Methodology*. Elsevier Science, New York, 1983.

G. Chen, X. Huang, and X. Yang. *Vector Optimization – Set-Valued and Variational Analysis*, volume 541 of *Lecture Notes in Economics and Mathematical Systems*. Springer Verlag, Berlin, 2005.

F. Y. Edgeworth. *Mathematical Psychics*. C. Kegan Paul & Co., London, 1881.

A. Göpfert and R. Nehse. *Vektoroptimierung*, volume 74 of *Mathematisch-Naturwissenschaftliche Bibliothek*. BSB B.G. Teubner Verlagsgesellschaft, Leipzig, 1990.

J. Jahn. *Vector Optimization – Theory, Applications, and Extensions*. Springer Verlag, Berlin, 2004.

M. Köksalan, J. Wallenius, and S. Zionts. *Multiple Criteria Decision Making – From Early History to the 21st Century*. World Scientific Publishing, Singapore, 2011.

T. C. Koopmans. Analysis of production as an efficient combination of activities. In T.C. Koopmans, editor, *Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics Monograph No. 13, pages 33–97. John Wiley & Sons, New York, 1951.

H. W. Kuhn and A. W. Tucker. Nonlinear programming. In J. Neyman, editor, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, Berkeley, 1951.

I. M. D. Little. *A Critique of Welfare Economics.* The Clarendon Press, Oxford, 1950.

A. Löhne. *Vector Optimization with Infimum and Supremum.* Springer Verlag, Berlin, 2011.

V. Pareto. *Manuale di Economia Politica.* Società Editrice Libraria, Milan, 1906.

V. Pareto. *Manual of Political Economy.* Augustus M. Kelley Publishers, New York, 1971.

W. Stadler. A survey of multicriteria optimization or the vector maximum problem, Part I: 1776-1960. *Journal of Optimization Theory and Applications*, 29:1–52, 1979.

T. L. Vincent and G. Leitmann. Control-space properties of cooperative games. *Journal of Optimization Theory and Applications*, 6:91–113, 1970.

P. L. Yu. *Multiple Criteria Decision Making: Concepts, Techniques and Extensions.* Plenum Press, New York, 1985.

P. L. Yu and G. Leitmann. Compromise solutions, domination structures, and Salukvadze's solution. *Journal of Optimization Theory and Applications*, 13: 362–378, 1974.

L. A. Zadeh. Optimality and non-scalar-valued performance criteria. *IEEE Transactions on Automatic Control*, 8:59–60, 1963.

Matthias Ehrgott
Department of Engineering Science
The University of Auckland
New Zealand
m.ehrgott@auckland.ac.nz