On regularization and vector optimization of machine design variables

L B Matusov
Institute of Mechanical Engineering of the Russian Academy of Sciences, Moscow, Russia
matusoff.l@yandex.ru

Abstract. The correct determination of the feasible solutions set and the set of Pareto-optimal solutions are the most important in engineering optimization problems. To create a machine with optimal parameters, you need to know exactly the boundaries of the feasible set on which to search for these parameters. And to claim that the resulting machine design is optimal, we need to approximate the Pareto set with the necessary accuracy. To create a feasible set the method of obtain his constrains was proposed. Earlier, the results of our research on determining the rate of convergence of this method and approximation of feasible set of solutions were solved. In this paper, the problems of approximation and regularization Pareto optimal set are solved. Due to the fact that the set of Pareto-optimal solutions is not stable, even small errors in calculating the values of the system performance criteria can significantly change this set. It follows from this that, approximating a feasible solutions set with a given accuracy, we cannot guarantee approximations of the Pareto set. Such problems are called ill-posed according to Tikhonov. In these case regularization of the Pareto-optimal set is a solution of these ill-posed problem. To get a complete solution to this problem, acceptable for most practical tasks, is quite difficult. In this paper, this problem is solved for criteria that satisfy Lipschitz condition. The results obtained here are not only theoretical in nature, but are already used in the design and identification of parameters of mathematical models of machines.

1. Introduction. Basic definitions of multicriteria optimization

Suppose that a mathematical model of the system under study is given, and this model depends on the design variables \( \alpha_1, \ldots, \alpha_n \). Let \( \alpha = (\alpha_1, \ldots, \alpha_n) \).

As a rule, specialists set the limits of change for each of the design variables, that we will call design variable constraints. These constraints distinguish the parallelepiped \( \Pi \).

In addition to parametric constraints, usually in the conditions of the problem a functional and criterion constraints are included criteria [1].

Let be \( D \) the set of points that satisfy all the constraints. It is natural to call \( D \) as the set of feasible points, or feasible solutions set. We formulate one of the main problems of vector optimization.

It is necessary to find a set \( P \subset D \), for which

\[
\Phi(P) = \min_{\alpha \in D} \Phi(\alpha),
\]

(1)
where $\Phi(\alpha) = (\Phi_1(\alpha), \Phi_2(\alpha), ..., \Phi_k(\alpha))$ is the vector of criteria [1-3].

$P$ is called the Pareto set, as well as its image in the space of criteria - $\Phi(P)$.

Suppose that the criteria are continuous functions that satisfy the Lipschitz condition, i.e. for all vectors $\alpha$ and $\beta$ from the domain of definition of the criterion $\Phi$, there exists a number $L_\nu$ such that

$$|\Phi_\nu(\alpha) - \Phi_\nu(\beta)| \leq L_\nu \max_j |\alpha_j - \beta_j|.$$ 

Or equivalently: there is $L'_\nu$ such that

$$|\Phi_\nu(\alpha) - \Phi_\nu(\beta)| \leq L'_\nu \sum_{j=1}^{r} |\alpha_j - \beta_j|.$$ 

2. Pareto set approximation

The problem of constructing a Pareto set is quite complicated. This is explained by the fact that approximating a feasible set with a given accuracy [2], it is impossible to guarantee the approximation of the Pareto set. Such tasks are called ill-posed according to Tikhonov. A number of works propose a solution to this problem with sufficiently strong constrains imposed on the criteria or system of preferences of a specialist. The results proposed below were obtained only under the assumption of continuity of criteria and the fulfillment of the Lipschitz condition [2].

Let $P$ be the Pareto set in the parameter space; $\Phi(P)$ - his image; $\epsilon$ - is the set of permissible errors for each criterion. It is necessary to construct a finite Pareto set $\Phi(P_\epsilon)$, approximating $\Phi(P)$ with accuracy up to $\epsilon$. Let $\Phi(D_\epsilon)$ be an $\epsilon$-approximation of $\Phi(D)$ and $P_\epsilon$ - the Pareto optimal subset of $D_\epsilon$. As already mentioned, the approximation problem $\Phi(P)$ is ill-posed according to Tikhonov. We give a definition of this concept.

Let $P$ be functional (operator) from the space $X, P: X \sqsubseteq Y$. Suppose that there exists $y^* = \inf P(x)$, and $V_{\epsilon}(y^*)$ is a neighborhood of the required solution $y^*$. Let us single out an element $x^*$ (or a set of elements) in the space $X$ and its $\delta$-neighborhood $V_\delta(x^*)$. We call $X_\delta^{\epsilon}$ - the solution to the problem of finding the extremal value of functional $P$ if the solution simultaneously satisfies the conditions $x_\delta^{\epsilon} \in V_\delta(x^*)$ and $P(x_\delta^{\epsilon}) \in V_{\epsilon}(y^*)$. If at least one of these conditions is not satisfied for arbitrary $\epsilon$ and $\delta$ then such a problem is called Tikhonov's ill-posed.

A similar definition can be given when the $P$ - operator from the space $X$ into a space $Y$. Let $X_\epsilon = \{\Phi(D_\epsilon), \Phi(D)\}$; $Y = \{\Phi(P_\epsilon), \Phi(P)\}$, where $\epsilon \to 0$, and let $P: X \to Y$ - be the operator relating each element of $X$ to its Pareto optimal subset. Then, in accordance with the foregoing, the problem of constructing the sets $\Phi(D_\epsilon)$ and $\Phi(P_\epsilon)$, belonging simultaneously to the $\epsilon$-neighborhoods of $\Phi(D)$, and $\Phi(P)$, respectively, is ill-posed according to Tikhonov. In this case, a metric or topology must be defined in the spaces $X$ and $Y$ corresponding to the system of preferences of a specialist on $\Phi(D)$.

We define a topology on $X$ using a neighborhood of an arbitrary point $x$. Let $W_\epsilon(x)$ remains from such $\Phi(D_\epsilon)$, such that for any $\Phi(\alpha) \in x$, there exists $\Phi(\beta) \in \Phi(D_\epsilon)$, for which $|\Phi(\beta) - \Phi(\alpha)| \leq \epsilon$, for all $v$. Either $W_\epsilon(x)$ contains such $\Phi(D_\epsilon)$, that for any $\Phi(\alpha) \in \Phi(D_\epsilon)$ there is $\Phi(\beta) \in x$ such that $|\Phi(\beta) - \Phi(\alpha)| \leq \epsilon$, for all $v$. Similarly we define the neighborhood $W_{\epsilon}(y)$ for arbitrary $y \in Y$, and hence the topology on $Y$. We define the $\epsilon$-neighborhood of the point $\Phi(x^0) \in \Phi(P)$

$$V_{\epsilon} = \{\Phi(\alpha) \in \Phi(P) : \left| \Phi_v(\alpha^0) - \Phi_v(\alpha) \right| \leq \epsilon_v, \quad v = 1, ..., k \}.$$ 

In Theorem 1 below, we construct the Pareto optimal set $\Phi(P_\epsilon)$, in which for any point $\Phi(x^0) \in \Phi(P)$ and any of its $\epsilon$-neighborhoods $V_{\epsilon}$, there is a point $\Phi(\beta) \in \Phi(P_\epsilon)$, belonging to $V_{\epsilon}$. Conversely, in the $\epsilon$-neighborhood of any point $\Phi(\beta) \in \Phi(P_\epsilon)$ there exist point $\Phi(x^0) \in \Phi(P)$. $\Phi(P_\epsilon)$ is called approximation having the property M. Let there is $\Phi(D_\epsilon)$ - approximation of the set $\Phi(D)$ [2].
We say that the approximation \( \Phi(P_\epsilon) \) has the property \( M_1 \) if for any point \( \Phi(\alpha') \in \Phi(P) \) and any of its \( \epsilon \)-neighborhoods \( V \), there exists a point \( \Phi(\beta) \in \Phi(P_\epsilon) \), belonging to \( V \).

**Lemma 1.** If the criteria are continuous and satisfy the Lipschitz condition there is an approximation with the property \( M_1 \).

**Proof.** The proof of the lemma is carried out by analyzing the neighborhoods of the so-called suspicious points of \( \Phi(D_\alpha) \), i.e. such points in the vicinity of which there may be truly Pareto points. If in the neighborhoods of suspicious points there are new Pareto optimal solutions, then they are added to \( \Phi(P_\epsilon) \). Together with \( \Phi(P_\epsilon) \) they form \( \epsilon \)-approximation of the Pareto set. Let us define the set of suspicious points. Consider some point 
\[
\Phi(\beta) \in \Phi(P_\epsilon), \quad M_\epsilon^+ = \{ \Phi(\beta) \in \Phi(D_\alpha): \Phi(\beta) \geq \Phi(\alpha) \} \quad (\nu = 1, 2, \ldots, k), \quad M_\epsilon^- = \{ \Phi(\beta) \in \Phi(D_\alpha): \Phi(\beta) \leq \Phi(\alpha) \} \quad (\nu = 1, 2, \ldots, k)
\]

And let 
\[
M_\epsilon^+ = \Phi(D_\alpha) \setminus (M_\epsilon^- \cup \Phi(P_\epsilon)), \quad M_\epsilon^- = M_\epsilon^+ \cup M_\epsilon^\nu, \quad M = U M_\alpha.
\]

Let \( \Phi(\alpha') \) from \( \Phi(D_\alpha) \) be a Pareto point, then \( \Phi(\beta) \) improves the value of at least one criterion of an arbitrary point \( \Phi(\alpha') \in \Phi(P_\epsilon) \). If such \( \Phi(\beta) \) exists, then add it to \( \Phi(P_\epsilon) \). Otherwise \( \Phi(\alpha') \) does not contain a point \( \Phi(\alpha') \in \Phi(D_\alpha) \) with an accuracy of \( \epsilon \). This operation is repeated for all points, belonging to SM. Let \( U \Phi(\mu^1) \cup \Phi(P_\epsilon) = \Phi(P_\epsilon) \) and \( U \Phi(\mu^1) \cup D_\epsilon = D_\epsilon \) for all \( i \), where \( \mu^1 \) is a point obtained after performing the above procedure. So \( \Phi(P_\epsilon) \) can contain points, that are not Pareto optimal and must be discarded.

As a result, we get the set \( \Phi(P_\epsilon') \) which is Pareto optimal subset of \( \Phi(U \mu^1 \cup D_\epsilon) \) and \( \epsilon \)-approximation of \( \Phi(P_\epsilon) \). This completes the proof of the lemma.

The approximation \( \Phi(P_\epsilon') \) obtained in this way has the property \( M_1 \). However, the inverse statement, in the general case, is not true, because \( \Phi(P_\epsilon') \) may contain points, analysis which is useless.

We say that the \( \epsilon \)-approximation of the Pareto optimal set \( \Phi(P_\epsilon) \), constructed in Lemma 1 has the property \( M_2 \) if there exists a point \( \Phi(\beta) \in \Phi(P) \) within the \( \epsilon \)-neighborhood of any point of \( \Phi(\alpha) \in \Phi(P_\epsilon) \).

**Lemma 2.** There exists a subset \( \Phi(P_\epsilon^*) \) of the set \( \Phi(P_\epsilon') \), which possesses the property \( M_2 \).

**Proof.** Let \( \Phi(\alpha) \in \Phi(P_\epsilon') \), \( B \) be an arbitrary subset of \( \{1, 2, \ldots, k\} \) and \( N_\Phi(\alpha) = \{ \Phi(\beta) \in \Phi(P_\epsilon'): \Phi(\beta) \leq \Phi(\alpha) + \epsilon \} \) for any \( \nu \) of \( B \).

\[
N_\Phi(\alpha) = \{ \Phi(\beta) \in \Phi(P_\epsilon'): \Phi(\beta) \leq \Phi(\alpha) - \epsilon \}, \quad \text{for any } \nu \in \{1, 2, \ldots, k\}: \text{B. Let also that}
\]

\[
N^*_\Phi(\alpha) = N^*_\Phi(\alpha) \cup N^*_\Phi(\alpha). \quad \text{As before, we will investigate neighborhood of points } \alpha \text{ and } \beta, \text{ for which}
\]

\[
\Phi(\beta) \in \Phi(\alpha) \in \Phi(P_\epsilon'). \quad \text{If, for all points } P \text{-net [1] belonging to } \Pi \alpha \text{ and } \Pi \beta \text{ and satisfying previous requirements for approximating feasible set the condition is fulfilled: for any } \alpha' \text{ belonging to the intersection}
\]

3
П \alpha \text{ and D } \text{ there exists } \mu, \text{ belonging to the intersection } \Pi_\beta \text{ and D, such than } 
\Phi(\mu) \leq \Phi(\alpha') + \varepsilon', \text{ for any } \nu \text{ of B, where } \varepsilon' \text{ - is a residually small quantity, then the point } \Phi(\alpha') \text{ can be excluded from } \Phi(P_\varepsilon'), \text{ since its } \varepsilon-\text{neighborhood does not contain any point from } \Phi(P). \text{By carrying out similar procedure for all } \Phi(\alpha') \in \Phi(P_\varepsilon'), \text{ we get a subset } \Phi(P_\varepsilon') \text{ of the set } \Phi(P_\varepsilon'), \text{ having the property } M_2.

3. Pareto set regularization

It is easy to show that, using the procedures described in Lemmas 1 and 2, and computing \Phi(P_\varepsilon'), we actually proved the following:

**Theorem 1.** If the criteria are continuous and satisfy the Lipschitz condition, then \Phi(P_\varepsilon') is an approximation of the Pareto set \Phi(P) possessing the property \text{M}.

Let be \varepsilon_j \text{- convergent to 0, decreasing sequence of positive numbers. It well known that the solution of an ill-posed problem, which is the problem of approximating a Pareto set comes down to constructing a regularizing sequence. In this case, it is a sequence of sets } \Phi(P_\varepsilon_j), j=1,2,\ldots \text{ such that for any corresponding sequences } \Phi(D_\varepsilon_j) \text{ and any } \varepsilon_j \text{ neighborhoods of the sets } \Phi(P) \text{ and } \Phi(D), \text{ the sets } \Phi(P_\varepsilon_j) \text{ and } \Phi(D_\varepsilon_j), \text{ starting with some } j_0, \text{ belong to the corresponding neighborhoods.}

Suppose that, in accordance with Lemmas 1 and 2, the sequences \Phi(P_\varepsilon_j), \Phi(D_\varepsilon_j), P_\varepsilon_j, D_\varepsilon_j \text{ are constructed for the sequence } \varepsilon_j. \text{Then the following holds.}

**Theorem 2.** The sequence \Phi(P_\varepsilon_j) - is regularizing.

**Proof.** Since that the property M, valid for any \Phi(P_\varepsilon_j), \text{ and the definition of neighborhoods } W_\varepsilon(\Phi(P)) \text{ and } W_\varepsilon(\Phi(D)) \text{ it immediately follows that the conditions defining regularizing sequence are performed.}

4. Remarks

1. In a number of papers, the problem of regularizing the Pareto set using the Hausdorff metric. It is known, however, that to construct on an feasible domain of solutions, a metric corresponding to the specialist’s preference system is not possible or very difficult. The class of problems described by the Hausdorff metric is very narrow, because when applied to a general case, as a rule a distortion of the preferences structure of the specialist will occur. Also arises question: why should this metric be generated by any predefined distance? Indeed, with a change in this distance, convergence may change. Therefore, in the general case, it is necessary to introduce a topology, similar done above. This topology is a generalization of the Hausdorff metric, but this does not distort the system of preferences of the specialist.

2. Since in practical problems of multicriteria design, the number vectors belonging to the set of Pareto optimal solutions is small due to the presence of strong functional and criteria constrains, then the procedure for studying the neighborhood of suspicious points is constructive, i.e. implemented on a computer at an acceptable time.

Similarly, we can obtain an approximation and regularization of the set Pareto in the parameter space. The latter is especially important for solving problems of multicriteria identification) [4,5].

3. The results obtained here are not only theoretical in nature, but are already used in the design and identification of parameters of mathematical models of machines.

5. Conclusion

1. A method for approximating the set of Pareto optimal solutions with given accuracy is developed.

2. The approximation algorithm is presented.

3. The problem of regularizing the Pareto set is solved.
References

[1] Sobol IM and Statnikov RB 2006 *The choice of optimal parameters in problems with many criteria*. Moscow Drofa.

[2] Matusov LB 2019 Vector Optimization of a Complicated Intellectual and Mechanical System *Europ. J. of Nat. Hist*. N. 5 p. 70-74.

[3] Statnikov R B, Bordsky A, Matusov J, Sobol’ I and Statnikov A 2012 Definition of the Feasible Solution Set in Multicriteria Optimization Problems with Continuous, Discrete, and Mixed Design Variables. *Journal of Nonlinear Analysis* 71, 12 pp.109-117.

[4] Statnikov R B 1999 *Multicriteria Design. Optimization and Identification*. Dordrecht/Boston/London Kluwer Academic Publishers

[5] Statnikov R B and Statnikov A 2011 *The Parameter Space Investigation Method Toolkit*, Boston/London Artech House.