Solution of Special Classes of Multi-extremal Problems

Igor A. Bykadorov

Sobolev Institute of Mathematics SB RAS Acad. Koptyug avenue 4, 630090 Novosibirsk, Russia Novosibirsk State University Pirogova street 2, 630090 Novosibirsk, Russia Novosibirsk State University of Economics and Management Kamenskaja street 56, 630099 Novosibirsk, Russia bykadorov.igor@mail.ru

Abstract

We suggest an approach to solve special classes of multi-extremal problems to optimize the monotone combination (e.g., sum, product) of several functions, under the assumption that the effective algorithms to optimize each of this item are known (e.g., each of these functions has some properties of generalized concavity: linear fractional, etc.) The algorithm proposed is iterative. It realizes one of the idea of the branchand-bound method and consists in successive correcting of the low and the upper bounds of optimal value of objective functions. Moreover, we use the methodology of multi-objective optimization, studying the image of Pareto boundary in the image space. In each iteration, the total area of the region, guaranteed to contain the image optimal point, decreases at least twice.

1 Introduction

At the moment, many types of multi-extremal problems are studied in optimization theory (see, for example, the famous monographs [Horst & Tuy, 1993], [Horst & Pardalos, 1995]). In particular, we study the problem of searching for a constrained extrema of superpositions of several functions under the assumption that the effective algorithms are known to optimize each of the functions involved. One of the most popular classes of such problems is, for example, the minimization of the sum of quasi-convex functions.

In the paper, the main attention is paid to the constrained minimization of the sum of two functions under the assumption that an algorithm for minimizing each term is known. However, the general idea of the algorithm allows us to transfer it to the case of a larger number of terms. Moreover, this approach is suitable for optimization, for example, the product of several functions.

Let us consider the problem

$$F(x) = F_1(x) + F_2(x) \to \min_{x \in X},\tag{1}$$

where $X \subset \mathbb{R}^n$ is a set while F_i are real-valued functions defined on X. We assume that effective algorithms are known for solving each of the following problems:

$$F_i(x) \to \min_{x \in X}, \quad i = 1, 2.$$
 (2)

Copyright © by the paper's authors. Copying permitted for private and academic purposes.

In: Yu. G. Evtushenko, M. Yu. Khachay, O. V. Khamisov, Yu. A. Kochetov, V.U. Malkova, M.A. Posypkin (eds.): Proceedings of the OPTIMA-2017 Conference, Petrovac, Montenegro, 02-Oct-2017, published at http://ceur-ws.org

Note that to solve problem (1) the effective algorithms (taking into account the special structure of the problem) are known only for the case when the set X is polyhedral and the functions F_i have a special form (for example, they are linear fractional). Let us mention briefly some works: [Choo et al., 1982], [Warburton, 1985], [Benson, 2002], [Kuno, 2002], [Gruzdeva & Strekalovsky, 2016]. A feature of the described algorithms in mentioned works is that they take into account a special kind of problem, therefore they cannot be transferred directly to the case when the functions F_i have a more general form.

The main goal of our paper is to construct an effective algorithm for solving a problem of the form (1). The proposed algorithm realizes the idea of the branch-and-bound method and consists in successively refinement of the boundaries of the optimal value of the objective function. The very preliminary variant of this approach can be found in [Bykadorov, 2016].

2 Statement of the Problem and Preliminary Discussions

Consider the problem

$$(P): f(x) = f_1(x) + f_2(x) \to \min_{x \in X}$$

where $X \subset \mathbb{R}^n$ is a set while f_i are real-valued functions defined on X. We assume that effective algorithms are known for solving each of the following problems:

$$f_i(x) \to \min_{x \in X}, \quad i = 1, 2.$$

 $\nu_i^0 = \min_{x \in X} f_i(x), \quad i = 1, 2.$ (3)

Let us denote

Let us associate with each
$$\nu_i \in \mathbb{R}$$
 the sets $X_i(\nu_i) = \{x \in X : f_i(x) \le \nu_i\}, i = 1, 2, \text{ and consider the problems}$

$$(P_1(\nu_2)): f_1(x) \to \min_{x \in X_2(\nu_2)}, \qquad (P_2(\nu_1)): f_2(x) \to \min_{x \in X_1(\nu_1)}.$$

Let $x_{P_1(\nu_2)}$ and $x_{P_2(\nu_1)}$ be the solutions of the problems $(P_1(\nu_2))$ and $(P_2(\nu_1))$, respectively. For arbitrary choice of ν_1 and ν_2 , it is possible that problems $(P_1(\nu_2))$ and $(P_2(\nu_1))$ have no solutions. But due to (3), problems $(P_1(\nu_2^0))$ and $(P_2(\nu_1^0))$ and $(P_2(\nu_1^0))$ are solvable and, moreover, $f_1(x_{P_2(\nu_1^0)}) = \nu_1^0$ and $f_2(x_{P_1(\nu_2^0)}) = \nu_2^0$.

Let us denote

$$\nu_1^{00} = \min_{x \in X_2\left(\nu_2^0\right)} f_1(x) = f_1\left(x_{P_1\left(\nu_2^0\right)}\right), \qquad \nu_2^{00} = \min_{x \in X_1\left(\nu_1^0\right)} f_2(x) = f_2\left(x_{P_2\left(\nu_1^0\right)}\right). \tag{4}$$

Definition. A pair (ν_1, ν_2) is said to be **attainable** if a point $x \in X$ exists such that $f_1(x) = \nu_1, f_2(x) = \nu_2$. **Remark.** By construction, pairs (ν_1^0, ν_2^{00}) and (ν_1^{00}, ν_2^0) are attainable.

Let x^* be the solution of Problem (P) and $f_i(x^*) = \nu_i^*$, i = 1, 2. Due to (3) and (4), $\nu_i^* \in [\nu_i^0, \nu_i^{00}]$, i = 1, 2. Consider right isosceles triangle *ABC* with vertices

$$A = \left(\nu_1^0 + \nu_{1,2}^0, \nu_2^0\right), \qquad B = \left(\nu_1^0, \nu_2^0\right), \qquad C = \left(\nu_1^0, \nu_2^0 + \nu_{1,2}^0\right),$$

where

$$\nu_{1,2}^0 = \min\left\{\nu_1^{00} - \nu_1^0, \nu_2^{00} - \nu_2^0\right\}$$
(5)

(see figure 1). Then problem (P) is equivalent to the following:

in triangle ABC, find a point such that its coordinates (ν_1, ν_2) form an attainable pair and moreover

$$\nu_1 + \nu_2 = \nu_1^* + \nu_2^*$$

Let us denote $T_i = \left[\nu_i^0, \nu_i^{00}\right]$, i = 1, 2. In what follows, we assume that the following condition is fulfilled. **Condition** (A). For any pair $(\nu_1, \nu_2) \in T_1 \times T_2$, the points $x' \in X$ and $x'' \in X$ exist such that

$$f_1(x') = \nu_1, \quad f_2(x') = f_2(x_{P_2(\nu_1)}), \quad f_1(x'') = f_1(x_{P_1(\nu_2)}), \quad f_2(x'') = \nu_2.$$

Remark. Let $(\nu_1, \nu_2) \in T_1 \times T_2$. Then the pairs $(\nu_1, f_2(x_{P_2(\nu_1)}))$ and $(f_1(x_{P_1(\nu_2)}), \nu_2)$ are attainable. Lemma 1. The following statements are true.

- Let $\nu'_1 \in T_1, \nu''_1 \in T_1$. If $\nu'_1 \le \nu''_1$ then $f_2\left(x_{P_2(\nu''_1)}\right) \ge f_2\left(x_{P_2(\nu'_1)}\right)$.
- Let $\nu'_2 \in T_2, \nu''_2 \in T_2$. If $\nu'_2 \le \nu''_2$ then $f_1\left(x_{P_1\left(\nu''_2\right)}\right) \ge f_1\left(x_{P_1\left(\nu'_2\right)}\right)$.

Note. The proofs of this and subsequent statements are rather technical. We plan to bring these formal proofs in the extended version of this paper.

Consider the function $G: T_1 \to \mathbb{R}$ defined as follows:

$$G(\nu_1) = \min \{ f_2(x) : x \in X_1(\nu_1), \nu_1 \in T_1 \}.$$

Due to Lemma 1, we have

Corollary 1.1. The function G is decreasing.

Consider the set

$$Y = \{(\nu_1, G(\nu_1)) : \nu_1 \in T_1\}$$
(6)

(see figure 2). We associate with each pair $(\nu_1, \nu_2) \in Y$ the line $H(\nu_1, \nu_2)$ passing through it and parallel to the hypotenuse of triangle *ABC*. Then the pair (ν_1^*, ν_2^*) of interest to us (the values of the functions f_1 and f_2 in the optimum) is characterized as follows (see figure 2): for each pair $(\nu_1, \nu_2) \in Y$, the line $H(\nu_1, \nu_2)$ lies "above" the straight line $H(\nu_1^*, \nu_2^*)$.

The remainder of this section we devote to describing a situation when Condition (A) holds. Lemma 2. The following statements are true.

- Let $\nu_1 \in T_1$. If $f_1(x_{P_2(\nu_1)}) < \nu_1$ then $x_{P_2(\nu_1)} \notin \arg\min_{x \in X} f_2(x)$.
- Let $\nu_2 \in T_2$. If $f_2(x_{P_1(\nu_2)}) < \nu_2$ then $x_{P_1(\nu_2)} \notin \arg\min_{x \in X} f_1(x)$.

Lemma 3. The following statements are true.

- Let $\nu_1 \in T_1$ and function f_2 be quasi-convex on set X. If $f_1(x_{P_2(\nu_1)}) < \nu_1$ then $x' \in X$ exists such that $f_1(x') = \nu_1, f_2(x') = f_2(x_{P_2(\nu_1)}).$
- Let $\nu_2 \in T_2$ and function f_1 be quasi-convex on set X. If $f_2(x_{P_1(\nu_2)}) < \nu_2$ then $x'' \in X$ exists such that $f_1(x'') = f_1(x_{P_1(\nu_2)})$, $f_2(x'') = \nu_2$.

Corollary 3.1. The following statements are true.

- Let function f_2 be quasi-convex on set X. For each $\nu_1 \in T_1$, point x' exists such that $f_1(x') = \nu_1$, $f_2(x') = f_2(x_{P_2(\nu_1)})$.
- Let function f_1 be quasi-convex on set X. For each $\nu_2 \in T_2$, point $x'' \in X$ exists such that $f_1(x'') = f_1(x_{P_1(\nu_2)})$, $f_2(x'') = \nu_2$.

Corollary 3.2. Let functions f_1 and f_2 be quasi-convex on X. Then Condition (A) holds.

3 The Main Idea of the Algorithm

The algorithm is iterative, realizes one of the ideas of the branch-and-bound method, and consists in the sequential refinement of the estimates of the values ν_1^*, ν_2^* and $\nu_1^* + \nu_2^*$, as well as the reduction the total area of the region containing the point (ν_1^*, ν_2^*) .

Let $\nu_1 \in [\nu_1^0, \nu_1^0 + \nu_{1,2}^0]$. (For the definition of ν_1^0 see (3), while the definition of $\nu_{1,2}^0$ see in (5)). We set $\nu_2 = G(\nu_1)$. Note that $(\nu_1, \nu_2) \in Y$ by the definition of set Y, see (6).

The following cases are possible:

- $\nu_1 + \nu_2 < \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$, see figure 3;
- $\nu_1 + \nu_2 = \nu^0$, see figure 4;
- $\nu_1 + \nu_2 > \nu^0$, $\nu_2 < \nu_2^0 + \nu_{1,2}^0$, see figure 5;
- $\nu_1 + \nu_2 > \nu^0$, $\nu_2 \ge \nu_2^0 + \nu_{1,2}^0$, see figure 6.

In each of these cases we can exclude from further consideration the regions that (due to Corollary 2) does not contain the point of our interest (ν_1^*, ν_2^*) . These areas correspond to the shaded parts of the triangle *ABC*. As a result, the estimates for the value $\nu_1^* + \nu_2^*$ are refined, and the area of the region, which is guaranteed not containing the point (ν_1^*, ν_2^*) , is also increased.

In the next steps, the described procedure is applied to each of the obtained unshaded triangles, or to one of them (for example, the largest, "the most perspective").

4 Some Remarks

1. As initial lower bounds for ν_1^*, ν_2^* and $\nu_1^* + \nu_2^*$, we can take, for example, the following:

$$\underline{\nu}_i = \nu_i^0, \quad i = 1, 2, \qquad \underline{\nu} = \underline{\nu}_1 + \underline{\nu}_2,$$

and as the upper bounds, the values

$$\overline{\nu}_i = f_i(x^i), i = 1, 2, \qquad \overline{\nu} = \min\left\{\underline{\nu}_1 + \overline{\nu}_2, \overline{\nu}_1 + \underline{\nu}_2\right\}$$

where

$$x^{i} \in \{x \in X : f_{i}(x) = \underline{\nu}_{i}\}, \quad i = 1, 2.$$

2. If we set

$$\nu_1 = \frac{\nu_1 + \overline{\nu}_1}{2}, \quad \nu_2 = f_2 \left(x_{P_2(\nu_1)} \right)$$

then due to Corollary 1.1, we can delete from the triangle ABC the region whose area is not less than half the area of the triangle ABC, see all the cases shown in figures 3 - 6. Therefore, in each iteration, the total area of the region, guaranteed to contain the image optimal point, decreases at least twice. This allows us to tell about the effectiveness of the algorithm.

3. The disadvantage of the proposed approach is to recognize the possible increase in the number of resulting triangles, this leads to an increase in the volume of stored information. However, in the case of an excessive increase in the number of these triangles, one can be chosen (for example, the largest one, i.e., "promising") and temporarily "forget" about the others, see figure 7, thus obtaining new estimates of the quantities ν_1^*, ν_2^* and $\nu_1^* + \nu_2^*$ for this selected triangle. These new estimates may allow us to exclude some of the "forgotten" triangles from further consideration, since we remove all parts of the triangles lying "above" the corresponding hypotenuse (this part may coincide with the whole triangle, see figure 8). Then we can consider all the "updated" triangles, choose one of them as the most "promising" for the next step. Thus, the number of considered triangles does not necessarily increase, and, moreover, may even decrease.

5 Formal Description of the Algorithm

Step 0. Define the values $\underline{\nu_1}, \underline{\nu_2}, \underline{\nu}, \overline{\nu_1}, \overline{\nu_2}, \overline{\nu}$ such that $\underline{\nu_1} \leq \nu_1^* \leq \overline{\nu_1}, \ \underline{\nu_2} \leq \nu_2^* \leq \overline{\nu_2}, \ \underline{\nu} \leq \nu_1^* + \nu_2^* \leq \overline{\nu}$. Choose ε , the required accuracy of calculations.

Set l = 1 as the total number of triangles under consideration, s = 1 as the number of the considered triangle.

Step 1. Choose $\nu_1 \in [\underline{\nu_1}, \overline{\nu_1}]$. For example, set $\nu_1 = \frac{\underline{\nu_1} + \overline{\nu_1}}{2}$. Solve the problem $P_2(\nu_1)$. Set $\nu_2 = f_2(x_{P_2(\nu_1)})$. Step 2.

(2a) Let the case shown in figure 3 is realized, i.e., $\nu_1 + \nu_2 < \nu \equiv \nu_1 + \nu_2 + \nu_{1,2}$, where $\nu_{1,2} = \min \{ \underline{\nu_1} - \overline{\nu_1}, \underline{\nu_2} - \overline{\nu_2} \}$ (cf. (5)). Calculate for the considered triangle the values x, y, z (see figure 9):

$$x[s] = \nu_1, \qquad y[s] = \nu_1 + \nu_2 - \nu_2, \qquad z[s] = \nu_2.$$

Calculate x, y, z for the new triangle: $x[l+1] = \underline{\nu_1}, y[l+1] = \nu_1, z[l+1] = \nu_2$. Set l = l+1. Go to Step 3.

(2b) Let one of the cases shown in figure 4 and 5 is realized, i.e., $\nu_1 + \nu_2 = \nu$, or $\nu_1 + \nu_2 > \nu$ but $\nu_2 < \underline{\nu_2} + \nu_{1,2}$. Set $x[s] = \nu_1$, $y[s] = \overline{\nu_2}$, $z[s] = \underline{\nu_2}$, $x[l+1] = \underline{\nu_1}$, $y[l+1] = \underline{\nu_1} + \overline{\nu_2} - \nu_2$, $z[l+1] = \nu_2$, l = l+1. Go to Step 4. (2c) Let the case shown in figure 6 is realized, i.e., $\nu_1 + \nu_2 > \nu$ but $\nu_2 \geq \underline{\nu_2} + \nu_{1,2}$. Set $x[s] = \nu_1$, $y[s] = \overline{\nu_2}$, $z[s] = \nu_2$. Go to Step 4.

Step 3. Compare the coordinates of the vertices of all the triangles obtained. If the situation shown in figure 8, are realized, then delete the triangles (or parts thereof) located "between" the two hypotenuses. Renumber

the remaining triangles. Recalculate for each of the triangles the value y (since in the triangles the trapezium part was removed, with bases parallel to the hypotenuse).

Step 4. Among all triangles, choose the largest (i.e., such that the value y - x is maximal). We assign the number s to this triangle.

Step 5. For triangle with the number s, calculate the values x[s], y[s], z[s], see figure 9. Set

 $\underline{\nu_1} = x[s], \qquad \overline{\nu_1} = y[s], \qquad \underline{\nu_2} = z[s], \qquad \overline{\nu_2} = z[s] + y[s] - x[s], \qquad \underline{\nu} = x[s] + z[s], \qquad \overline{\nu} = y[s] + z[s]$

If $\overline{\nu} - \underline{\nu} \equiv y[s] - x[s] \leq \varepsilon$ then STOP. Otherwise, go to Step 1.

6 Conclusion

The application of the proposed approach to the general form of the problem (1) requires that effective algorithms for solving each of the problems (2) be known. In addition, Condition (A) is required. In particular, we can assume that the functions F_i have some properties of generalized convexity (quasi-convexity, pseudo-convexity, see, for example, [Avriel et al., 1988]).

Possible modifications of the proposed approach deal with different ways of constructing the curve Y (see (6)) using already known achievable points (for example, by interpolations, approximations).

Besides, the described approach is also applicable to the optimization of "monotonic" combinations (for example, products) of functions. In this case, for the product of two functions, in image space, the level lines of the function $f_1 \cdot f_2$ are hyperbolas. Hence, instead of the right isosceles triangles, we deal with the "curvilinear" right triangles by replacing the hypotenuse with a piece of the corresponding hyperbola. It seems that the generality of the proposed approach is its advantage over other methods that take into account the structure of the problem.

Of course, the practical value of the proposed approach is questionable since no experimental results are presented. It is necessary to add practical experiments demonstrating the superiority of the proposed approach over other global optimization methods. We plan to do it in the future, in the extended versions of the paper.

Finally, the approach can be applied to the problems that arise in marketing: optimization of communication expenditure [Bykadorov et al., 2002] and the effectiveness of advertising [Bykadorov et al., 2009a], pricing [Bykadorov et al., 2009b]; to monopolistic competition models: retailing [Bykadorov et al., 2014], investments in R&D [Antoshchenkova & Bykadorov, 2017], market distortion [Bykadorov et al., 2016], and international trade [Bykadorov et al., 2015].

The author considers it his pleasant duty to express deep gratitude to the anonymous reviewers for very valuable comments. It is hoped that their excellent comments have allowed to improve the content of the paper and the presentation of the material.

References

- [Antoshchenkova & Bykadorov, 2017] Antoshchenkova, I.V., & Bykadorov, I.A. (2017). Monopolistic competition model: The impact of technological innovation on equilibrium and social optimality. Automation and Remote Control, 78(3), 537-556. doi:10.1134/S0005117917030134
- [Avriel et al., 1988] Avriel, M., Dewert, W.E., Schaible, S., & Zang, I. (1988). Generalized concavity. New York a.o.: Plenum press.
- [Benson, 2002] Benson, H.P. (2002). Global optimization algorithm for the nonlinear sum of ratios problem. Journal of Optimization Theory and Applications, 112(1), 1-29. doi:10.1023/A:1013072027218
- [Bykadorov, 2016] Bykadorov, I.A. (2016). On an Approach to Solving Special Classes of Multi-extremal Problems. In: Kabanikhin, S. et al. (Eds.). Proceedings of the XIII International Asian School-seminar "Problems of complex systems optimization", Novosibirsk, 12-16 December 2016 [in Russian] (pp. 91-99) http://conf.ict.nsc.ru/opcs2016/ru/proceedings
- [Bykadorov et al., 2016] Bykadorov, I., Ellero, A., Funari, S., Kokovin, S., & Pudova, M. (2016). Chain Store Against Manufacturers: Regulation Can Mitigate Market Distortion. In: Kochetov, Yu. et al. (Eds.). Proceedings of the 9th International Conference "Discrete Optimization and Operations Research" (Lecture Notes in Computer Sciences, 9869, pp. 480-493). Heidelberg, Germany: Springer. doi:10.1007/978-3-319-44914-2_38

- [Bykadorov et al., 2009a] Bykadorov, I., Ellero, A., Funari S., & Moretti, E. (2009). Dinkelbach Approach to Solving a Class of Fractional Optimal Control Problems. Journal of Optimization Theory and Applications, 142(1), 55-66. doi:10.1007/s10957-009-9540-5
- [Bykadorov et al., 2002] Bykadorov, I., Ellero, A., & Moretti, E. (2002). Minimization of communication expenditure for seasonal products. *RAIRO Operations Research*, 36(2), 109-127. doi:10.1051/roi2002012
- [Bykadorov et al., 2009b] Bykadorov, I., Ellero, A., Moretti, E., & Vianello, S. (2009). The role of retailer's performance in optimal wholesale price discount policies. *European Journal of Operational Research*, 194(2), 538-550. doi:10.1016/j.ejor.2007.12.008
- [Bykadorov et al., 2015] Bykadorov, I., Gorn, A., Kokovin, S., & Zhelobodko, E. (2015). Why are losses from trade unlikely? *Economics Letters*, 129, 35-38. doi:10.1016/j.econlet.2015.02.003
- [Bykadorov et al., 2014] Bykadorov, I.A., Kokovin, S.G., & Zhelobodko, E.V. (2014). Product Diversity in a Vertical Distribution Channel under Monopolistic Competition. Automation and Remote Control, 75(8), 1503-1524. doi:10.1134/S0005117914080141
- [Choo et al., 1982] Choo, E.U. & Atkins, D.R. (1982). Bicriteria Linear Fractional Programming. Journal of Optimization Theory and Applications, 36(2), 203220. doi:10.1007/BF00933830
- [Gruzdeva & Strekalovsky, 2016] Gruzdeva, T. & Strekalovsky, A. (2016). An Approach to Fractional Programming via D.C. Constraints Problem: Local Search. In: Kochetov, Yu. et al. (Eds.). Proceedings of the 9th International Conference "Discrete Optimization and Operations Research" (Lecture Notes in Computer Sciences, 9869, pp. 404-417). Heidelberg, Germany: Springer. doi:10.1007/978-3-319-44914-2_32
- [Horst & Pardalos, 1995] Horst, R. & Pardalos, P.M. (Eds.) (1995). Handbook of Global Optimization Springer. doi:10.1007/978-1-4615-2025-2
- [Horst & Tuy, 1993] Horst, R. & Tuy H. (1993). Global Optimization. Deterministic approach. Berlin: Springer Verlag.
- [Kuno, 2002] Kuno, T. (2002). A branch-and-bound algorithm for maximizing the sum of several linear ratios. Journal of Global Optimization, 22(1-4), 155-174. doi:10.1023/A:1013807129844
- [Warburton, 1985] Warburton, A.R. (1985). Parametric Solution of Bicriterion Linear Fractional Programming. Operation Research, 33(1), 7484. doi:10.1287/opre.33.1.74



Figure 1: The illustration: triangle ABC



Figure 2: The illustration: the curve Y



Figure 3: The case $\nu_1 + \nu_2 < \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$



Figure 4: The case $\nu_1 + \nu_2 = \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$



Figure 5: The case $\nu_1 + \nu_2 < \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$



Figure 6: The case $\nu_1 + \nu_2 < \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$



Figure 7: Selection of the most "promising" triangle



Figure 8: The situation when the number of triangles decreases



Figure 9: Illustration for a formal description of the algorithm