The method of conditional minimization based on the selection of the optimal value of the objective function

I Ya Zabotin, K E Kazaeva and O N Shulgina
Kazan (Volga Region) Federal University, 18, Kremlyovskaya st., Kazan, 420008

E-mail: iyazabotin@mail.ru

Abstract. One of the convex programming methods that allows to reduce the solution of the original problem to the successive minimization of auxiliary functions on the whole space or on simple sets is the method of parametrization of the objective function. A method close to the one mentioned above is proposed in this paper. As in the known method, in order to find an iterative point, we solve the minimization problem of some auxiliary function constructed with cuts of the parametrized objective function and the constraint functions. The proposed method has some advantages over the parametrization method of the objective function due to a different principle of setting parameter values. The method convergence is substantiated. The properties associated with its implementation are discussed.

1. Introduction

In practice, in solving problems of conditional minimization, methods that are united by the following general scheme are often used. At each step of these methods, to find an approximation, a certain auxiliary function is constructed on the basis of objective functions and constraints. As an iterative point, the minimum point of this auxiliary function is taken either on the entire space or on a set that is simpler than the admissible set.

Methods of this type include the method of external penalty functions (e.g., [1 - 3]), the method of barrier functions (e.g., [1, 2]), variants of methods of centres and external centres (e.g., [4]), the method of parametrization of the objective function (e.g., [2, 3]), etc.

The method proposed here is ideologically close to the last of the known methods mentioned. Note that in the parametrization method of the objective function, the lower estimate of the optimal value is used at the initial step, which is not always easy to obtain. In addition, the parameter values at all iterations of the method should not exceed the optimal value of the objective function. However, if the minimization problem of auxiliary function is solved approximately at some step, then the next value of the parameter might exceed the optimal value, and further application of the method becomes impossible. In the proposed method, there are no such problems with the choice of parameter values. We also note that iterative points of the basic sequence of approximations belong to the admissible domain of the general problem.

2. Problem statement

The problem

\[ \min \{ f(x) : x \in D \}, \] (1)
is solved, where \( D = \{ x \in R_n : F(x) \leq 0 \} \) is a bounded set, \( F(x) \) and \( f(x) \) are convex functions defined in \( R_n \).

Let \( f^* = \min \{ f(x) : x \in D \} \), moreover,

\[
f^* < 0,
\]

(2)

\( X^* = \{ x \in D : f(x) = f^* \}, x^* \in X^*, F^*(x) = \max \{ F(x), 0 \}, K = \{ 0, 1, \ldots \}. \)

Just note that the condition (2) is easily satisfied by changing the objective function of the original problem. For example, this could be done according to the following lemma.

Lemma 1. Let \( z \in D, f(x) = f(x) - (f(z) + \beta), \) where \( \beta > 0, f^* = \min \{ f(x) : x \in D \}. \) Then \( f^* < 0. \)

Proof. By the hypothesis, \( f^* \leq f(x) \) for all \( x \in D \), including \( x = z. \) Then \( f^* \leq f(z) = f(z) - \beta = -\beta < 0. \) The lemma is proved.

Note that the set of solutions of the problem \( \min \{ f(x) : x \in D \} \) coincides with the set \( X^* \) of solutions of problem (1).

The proposed method for solving the problem (1) produces a sequence of approximations \( x_k, k \in K, \) and consists in the following.

We choose a convex closed set \( D_0 \subseteq R_n \) such that \( x^* \in D_0. \) Choose \( \delta_0 > 0 \) and set \( a_0 = 0, f_0(x) = f(x), \)

\[
f^*_0(x) = \max \{ f_0(x), 0 \}, \quad \Phi_0(x) = f^*_0(x) + F^*(x), \quad i = 0, k = 0.
\]

1. The solution \( y_i \) of the problem

\[
\min \{ \Phi_i(x) : x \in D_0 \}. \quad (3)
\]

is found. If the inequalities

\[
\Phi_i(y_i) > 0, \quad F(y_i) \leq 0,
\]

(4)

are fulfilled simultaneously, then \( y_i \in X^* \), and the process terminates.

2. If \( \Phi_i(y_i) = 0 \), then set \( i_k = i, x_k = y_{i_k}, r = i_k, \)

\[
\alpha_{i_k+1} = \alpha_i + \delta_{i_k+1},
\]

where \( \delta_{i_k+1} = \delta_i \), the value of \( k \) is increased by one and the transition to step 3 follows.

Otherwise, i.e. for \( \Phi_i(y_i) > 0 \), it is assumed that

\[
\alpha_{i+1} = \alpha_i + \delta_{i+1},
\]

where \( \delta_{i+1} = \delta_i/p, \quad p > 1, \) and subsection 3 holds.

3. Set \( f_i(x) = f(x) + \alpha_{i+1}, \)

\[
f_i(x)^* = \max \{ f_i(x), 0 \}, \quad \Phi_i(x) = f_i(x)^* + F^*(x).
\]

The value of \( i \) is increased by one and the transition to step 1 follows.

3. The method’s discussion

Let us make some remarks to the method. First of all, note that for all \( i \in K \) and \( x \in R_n \) the inequality \( \Phi_i(x) \geq 0 \) holds. We set

\[
Y_i^* = \text{Argmin} \{ \Phi_i(x) : x \in D_0 \}.
\]

Theorem 1. If \( a_i = -f^i \) for some \( i \in K, \) then \( y_i \in X^* \) and \( \Phi_i(y_i) = 0. \)

Proof. Since \( y_i \in Y_i^* \), then \( \Phi_i(y_i) \leq \Phi_i(x) \) for all \( x \in D_0. \) Then, taking into account that \( x^* \in D_0, \) we have

\[
\max \{ f(y_i) - f^*, 0 \} + F^*(y_i) \leq \max \{ f(x^*) - f^* - f^0 , 0 \} + F^*(x^*).
\]

But \( F^*(x^*) = 0, \) therefore

\[
\max \{ f(y_i) - f^*, 0 \} + F^*(y_i) \leq 0. \quad (5)
\]
Assuming that $F'(y_i) > 0$, then inequality (5) becomes contradictory, since $\max\{f(y_i) - f^*, 0\} \geq 0$. Therefore, $F'(y_i) = 0$, i.e. $y_i \in D$, and by (5) $f(y_i) - f^* \leq 0$. On the other hand, $f(y_i) \geq f^*$, since, as just shown, $y_i \in D$. Thus, $f(y_i) = f^*$, and the first assertion is proved. The second assertion of the theorem follows from (5) and the inequality $\Phi(y_i) \geq 0$.

Lemma 2. Let the point $\bar{x}$ be such that $\Phi(\bar{x}) = 0$, $i \in K$. Then $\bar{x} \in D$.

Proof. By the hypothesis,

$$f_+^*(\bar{x}) + F^*(\bar{x}) = 0.$$  \hfill (6)

But $f_+^*(x) \geq 0$ and $F^*(x) \geq 0$ for all $x \in R_n$, therefore (6) could occur only in the case when $f_+^*(\bar{x}) = 0$ and $F^*(\bar{x}) = 0$. The last equality means that $F(\bar{x}) \leq 0$, i.e. $\bar{x} \in D$. The lemma is proved.

From Lemma 2, the following corollary immediately follows.

Corollary. If for the solution $y_i$ of the problem (3) the equality $\Phi(y_i) = 0$ holds, then $y_i \in D$.

Let us justify the criterion of optimality contained in step 1 of the method.

Theorem 2. Assume that for some $i \in K$ the inequalities (4) hold for the point $y_i$. Then $y_i \in X^*$.

Proof. Since $y_i \in Y_i^*$, with the hypothesis of Theorem 0 $< \Phi(y_i) \leq \Phi(x), x \in D_0$, or

$$0 < f_+^*(y_i) + F^*(y_i) \leq f_+^*(x) + F^*(x) \forall x \in D_0.$$

Here we set $x = x^*$. Then, in view of the equalities $F^*(x^*) = 0$ and $F^*(y_i) = 0$ we have $0 < f_+^*(y_i) \leq f_+^*(x^*)$ or

$$0 < \max\{f(y_i), 0\} = f(y_i) = f(x^*) + a_i \leq \max\{f(x^*), 0\} = f(x^*) = f(x^*) + a_i.$$  \hfill (7)

Hence $f(y_i) \leq f^*$. On the other hand, $f(y_i) \geq f^*$, since by the condition $y_i \in D$. Consequently, $f(y_i) = f^*$, and the theorem is proved.

In solving the problem (3), the following statements might be useful.

Lemma 3. If for some $i \in K$ inclusion $z \in D_0$ and inequalities

$$F(z) \leq 0, f(z) \leq 0$$

hold simultaneously for a point $z \in R_n$, then $z \in Y_i^*$.

Proof. The inequalities (7) mean that $F^*(z) = 0$ and $f_+^*(z) = \max\{f(z), 0\} = 0$, that is, $\Phi(z) = 0$. But $\Phi(x) \geq 0$ for all $x \in R_n$, including for all $x \in D_0$. Therefore, $\Phi(x) \geq \Phi(z)$ for all $x \in D_0$, so $z \in Y_i^*$. The lemma is proved.

Lemma 4. If the set $D_0$ is such that $D \subset D_0$, and the inequalities (7) hold for the point $z \in R_n$, then $z \in Y_i^*$.

The proof of the lemma is similar to the proof of Lemma 3.

Now we show how Lemmas 3 and 4 could be used in step 1 of the method for finding the points $y_i$.

Suppose that the equality $\Phi(y_i) = 0$ is satisfied for the approximation $y_i$ found at the $i$-th iteration, since $x_k = y_i$. Then, according to the corollary to Lemma 2, in addition to the inclusion $y_i \in D_0$ the inequality $F(y_i) \leq 0$ holds. In this case, compute the value of $f_{i+1}(y_i)$ at the next $(i + 1)$-th iteration. If it turns out that $f_{i+1}(y_i) \leq 0$, then by Lemma 3 the inclusion

$$y_i \in Y_{i+1}$$

holds. Therefore, on the $(i + 1)$-th iteration, the problem of minimizing the function $\Phi_{i+1}(x)$ on $D_0$ might not be solved, assuming

$$y_{i+1} = y_i, x_{k+1} = y_{i+1}.$$  \hfill (i)

In this case $f_{i+1}^*(y_{i+1}) = 0, F^*(y_{i+1}) = 0$, i.e.

$$\Phi_{i+1}(y_{i+1}) = 0,$$

and on the next $(i + 2)$-th iteration, it is also useful to calculate the value of $f_{i+2}(y_{i+1})$ and compare it with zero before starting the solution of the problem (3).

Let
According to the method for specifying numbers $\delta_i$, $i \in K$, the set $K^*$ consists of an infinite number.

Theorem 3. Any limit point of the sequence $\{x_k\}, k \in K^*$, belongs to the set $X^*$.

Proof. Let $x_k, k \in K' \subset K^*$, be a convergent subsequence of the sequence $\{x_k\}, k \in K^*$, and let $\bar{x}$ be the limit point of this subsequence. We prove that $\bar{x} \in X^*$.

According to step 2 of the method and the corollary to Lemma 2, $x_k = y_{ik}$,

$$y_{ik} \in D \forall k \in K.$$  \hfill (8)

Then, by virtue of the closedness of the set $D$, the inclusion $\bar{x} \in D$ holds, so

$$f(\bar{x}) \geq f^*.$$  \hfill (9)

Further, for all $k \in K$ we have the equalities $\Phi_{ik}(x_k) = \Phi_{ik}(y_{ik}) = f^*_k(y_{ik}) + F^*(y_{ik}) = 0$. Hence, taking into account (8), $f^*_k(y_{ik}) = 0$ or $\max\{f(y_{ik}) + \alpha_{ik}, 0\} = 0, k \in K$. So

$$f(y_{ik}) + \alpha_{ik} \leq 0 \forall k \in K.$$  \hfill (10)

Since $y_{ik+1} \in Y^*_{ik+1}, k \in K'$, it follows that $\Phi_{ik+1}(y_{ik+1}) \leq \Phi_{ik+1}(x) \forall k \in K', x \in D_0$. Let $x = x^*$ in the last inequality. Then for all $k \in K'$ the following relations hold:

$$0 < \Phi_{ik+1}(y_{ik+1}) = f^*_{k+1}(y_{ik+1}) + F^*(y_{ik+1}) \leq f^*_k(x^*) + F^*(x^*).$$

And since $F^*(x^*) = 0$, then $\max\{f(x^*) + \alpha_{ik+1}, 0\} > 0$ or $f(x^*) + \alpha_{ik+1} > 0$. From this and (10) it follows that

$$f(y_{ik}) + \alpha_{ik} < f^*(x^*) + \alpha_{ik+1} \forall k \in K'.$$

But $\alpha_{ik+1} = \alpha_{ik} + \delta_{ik+1}, k \in K'$, where $\delta_{ik+1} = \delta_{ik} / p$, thus

$$f(y_{ik}) < f^* + \delta_{ik} / p \forall k \in K'.$$  \hfill (11)

Since $\delta_{ik} / p \to 0$ as $k \to \infty, k \in K'$, it follows from (11) that $f(\bar{x}) \leq f^*$. This and (9) imply the assertion of the Theorem.

Note that the point $y_i$, as a solution of the problem (3), is necessary only in the case when

$$\Phi^*_i = \min\{\Phi_i(x): x \in D_0\} = 0.$$ 

If in the process of solving the problem (3) it turns out that the value of $\Phi^*_i$ is guaranteed to be positive, then according to step 2 of the method the point $y_i$ is not needed, and its search could be stopped.

In connection with this remark, during the testing of the method for solving the problem (3), a cutting algorithm was applied [5] with an approximation of the epigraph of the function $\Phi_i(x)$, since at each step of the algorithm there was a lower bound on the value of $\Phi^*_i$.

References
[1] Vasil’ev F P 2011 Optimization methods [in Russian] MCCME 620
[2]Evtushenko Yu 1981 Methods of Solving Extremal Problems and Their Application in Optimization Systems (Moscow: Nauka) 432
[3] Sukharev A G, Timokhov A V and Fedorov V V 1986 Course of optimization methods [in Russian] (Moscow: Nauka) 328
[4] Grossman K and Kaplan A A 1981 Nonlinear programming based on unconditional minimization [in Russian] (Novosibirsk: Nauka) 183
[5] Zabotin I Ya and Yarullin R S 2014 A Cutting Method for Finding Discrete Minimax with Dropping of Cutting Planes Lobachevskii Journal of Mathematics 35 (2) 157–163