Exact penalty method for minimizing the local Lipschitz functions on a convex set

Lyudmila Polyakova

Saint-Peterburg State University,

Department of Applied Mathematics and Control Processes,
Saint-Petersburg, Russia

13.02.2018

Introduction

Penalty function methods are simple and widely known methods for solving nonlinear programming problems.
The basic idea of the method is to approximate the reduction of the problem of minimizing the constraints to the problem of minimizing a function without any restrictions, with the auxiliary function chosen in such a way that it coincides with the given function to be minimized on the set and grows quickly outside.

Consider the constrained optimization problem: to find

$$\inf_{x\in X} f(x), \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a locally Lipschitz function on $\mathbb{R}^n$.
Suppose that the set $X \subset \mathbb{R}^n$ is non-empty, compact and given in the form

$$X = \{\, x \in \mathbb{R}^n \mid \varphi(x) = 0\},$$

where φ is a non-negative locally Lipschitz function on $\mathbb{R}^n$

$$\varphi(x) \geq 0 \quad \forall x \in \mathbb{R}^n.$$

Then the set X is a closed set of minimizers points of the function φ on $\mathbb{R}^n$. As usual, for solving of (1) by the method of penalty functions we introduce a function

$$F(c,x) = f(x) + c\varphi(x),$$

where c is a non-negative number, called *a penalty parameter*. Consider the unconstrained minimization problem

$$\inf_{x\in\mathbb{R}^n} F(c,x). \tag{2}$$

For the application of penalties it is essential to assume the existence of solutions of the auxiliary problems. Assume that the infimum in (2) is achieved for any non-negative number c.

Denote by

$$f^* = \min_{x \in X} f(x), \quad F^*(c) = \min_{x \in \mathbb{R}^n} F(c, x), \quad x(c) = \arg\min_{x \in \mathbb{R}^n} F(x, c).$$

$$x^{**} = \arg\min_{x \in \mathbb{R}^n} f(x), \quad f^{**} = f(x^{**}).$$

Suppose that $f^{**} > -\infty$. Note that $f^* \geq F^*(c)$ for any positive c.

Choose a monotonically increasing sequence of non-negative numbers $\{c_k\}(k = 0, 1, 2, \dots)$ that tends to $+\infty$,

$$0 = c_0 < c_1 < c_2 < \cdots < c_k < \dots, \quad c_k \to +\infty.$$

Then

$$x^{**} = x(c_0) = f^{**}.$$

Of course, at first we check whether x^{**} belongs to the set X (x^{**} is a global minimizer of f on $\mathbb{R}^n$).

If $x(c_k)$ is a sequence of solutions of auxiliary problems (2), then the following inequalities:

1) $$F^*(c_k) = F(c_k, x(c_k)) \leq F(c_{k+1}, x(c_{k+1})) = F^*(c_{k+1}) \quad \forall k > 0$$

2) $$f(x(c_k)) \leq f(x(c_{k+1}));$$

3) $$f(x(c_k)) \leq F(c_k, x(c_k)) \leq f^* \tag{3}$$

hold.

Lemma

Under the above assumptions we have the inequality

$$\varphi(x(c_{k+1})) \leq \varphi(x(c_k)) \quad \forall k > 0. \tag{4}$$

Denote

$$\mathcal{L}(\varphi, x^{**}) = \{x \in \mathbb{R}^n \mid \varphi(x) \leq \varphi(x^{**})\}.$$

The set $\mathcal{L}(\varphi, x^{**})$ is compact and bounded from below by the function φ. From (4) it follows that all the points of the sequence $\{x(c_k)\}$ lie in the set $\mathcal{L}(\varphi, x^{**})$.

Lemma

For any sequence of positive numbers $\{c_k\}$, $c_k \to +\infty$, *we have*

$$\varphi(x(c_k)) \underset{k\to+\infty}{\longrightarrow} 0. \qquad (5)$$

Thus, the sequence of points $\{x(c_k)\}$ is a minimizing sequence for the function φ.

Exact penalty function. Statement of the problem.

Exact penalty functions are the penalty functions for which there exists a parameter $c^* > 0$ such that for any $c \geq c^*$ the set of minimum points of $F(c, x)$ on $\mathbb{R}^n$ coincides with the set of solutions of (1).

A parameter c^* is called an exact penalty parameter for the family of functions $F(c, x)$.

Therefore, any number greater then an exact penalty parameter is also an exact penalty parameter.

Obviously, the implementation of the method of exact penalty functions is primarily depend on the properties of the function φ.

In practice it would be useful to find conditions which guarantee that there exists an exact penalty parameter $c^* \geq 0$ such that the set

$$\{x \in \mathbb{R}^n \mid x = \arg \min_{x \in \mathbb{R}^n} F(c, x)\}$$

coincides with the set

$$\{x \in \mathbb{R}^n \mid x = \arg \min_{x \in X} f(x)\}$$

Consider the optimization problem: find

$$\inf_{x \in X} f(x), \tag{6}$$

where

$$X = \{x \in \mathbb{R}^n \mid f_1(x) \leq 0\},$$

the function $f : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz and $f_1 : \mathbb{R}^n \to \mathbb{R}$ is convex.

Assume that X is not an isolated point and Slater's condition is satisfied, i.e., there exists a point $\hat{x} \in X$ for which the inequality $f_1(\hat{x}) < 0$ holds.

Consider the case where the minimum point of the function f does not belong to X.

For (6) we construct the problem equivalent to the original and solve it by using exact penalty functions.

Consider the optimization problem: find

$$\inf_{x \in X} f(x), \tag{7}$$

where

$$X = \{x \in \mathbb{R}^n \mid \varphi(x) = 0\}, \quad \varphi(x) = \max\{0, f_1(x)\}. \tag{8}$$

Under our assumptions, problems (6) and (7) are equivalent.

Introduce the function

$$F(c,x) = f(x) + c\varphi(x), \quad c > 0,$$

and consider an unconstrained minimization problem

$$\inf_{x \in \mathbb{R}^n} F(c,x). \tag{9}$$

Assume that the infimum in (9) is achieved for any positive number c. and the Lebesgue set

$$\mathcal{L}(f,x_0) = \{x \in \mathbb{R}^n \mid f(x) \le f(x_0)\},$$

is bounded, where x_0 is a point in $\mathbb{R}^n$. Note that for the problem (6) this condition is satisfied.

Suppose that in problem (1) the set X is given in the form (8), f_1 is a finite convex function on $\mathbb{R}^n$, and there is a point $\hat{x} \in \mathbb{R}^n$, for which $f_1(\hat{x}) < 0$, i.e., the set X has interior points.

In this case, the functions f_1 and φ are convex and satisfy expansions

$$f_1(x+\alpha g) = f_1(x) + \alpha f_1'(x,g) + o_1(\alpha,x,g), \ \ x,g \in \mathbb{R}^n, \ \ \alpha \in \mathbb{R},$$

$$\varphi(x+\alpha g) = \varphi(x) + \alpha \varphi'(x,g) + o(\alpha,x,g), \ \ x,g \in \mathbb{R}^n, \ \ \alpha \in \mathbb{R},$$

where $f_1'(x,g), \varphi'(x,g)$ are the direction derivatives of the functions f_1 and φ at x and

$$\frac{o_1(\alpha,x,g)}{\alpha} \underset{\alpha\downarrow 0}{\longrightarrow} 0, \quad \frac{o(\alpha,x,g)}{\alpha} \underset{\alpha\downarrow 0}{\longrightarrow} 0. \tag{10}$$

In this case,

$$f_1'(x,g) = \max_{v \in \partial f_1(x)} \langle v, g \rangle, \quad \varphi'(x,g) = \max_{v \in \partial \varphi(x)} \langle v, g \rangle.$$

where $\partial f_1(x),\ \partial \varphi(x)$ is the subdifferential of the functions f_1 and φ at x and

$$o(\alpha, x, g) \geq 0,\ \ o_1(\alpha, x, g) \geq 0 \quad \forall x, g \in \mathbb{R}^n, \alpha > 0.$$

If $x \in \text{int } X$, then $\partial\varphi(x) = 0_n$. If $x \in \text{bd } X$, then

$$\partial\varphi(x) = \text{co } \{0_n, \partial f_1(x)\}.$$

Thus, for each boundary point $x \in \text{bd } X$ the zero point belongs the subdifferential $\partial\varphi(x)$ and it is also a boundary point, since minimizers of the function f_1 are not contained in the set X.
Denote by $\Gamma(X, x)$ the cone of feasible directions at a point $x \in X$

$$\Gamma(X, x) = \text{cl } \{g \in \mathbb{R}^n \mid \exists \alpha_0 > 0, \quad x + \alpha g \in X \quad \forall \alpha \in [0, \alpha_0)\}.$$

In this case, the cone $\Gamma(X, x)$ is convex and closed.

The set

$$N(X,x) = \{g \in \mathbb{R}^n \mid \langle g, z - x\rangle \leq 0 \ \forall z \in X\}$$

called the normal cone to a set X at $x \in X$. The cone $N(X,x)$ is also closed and convex. Under the above assumptions at the point $x \in \text{bd } X$, $(f_1(x) = 0)$ the equality

$$N(X,x) = -[\text{cone } (X - x)]^* = -\Gamma^*(X,x)$$

holds, where $\Gamma^*(X,x)$ is the cone dual to the cone $\Gamma(X,x)$. The set

$$\gamma_1(X,x) = \left\{g \in \mathbb{R}^n \mid f_1'(x,g) \leq 0\right\}$$

called a cone of nonincreasing directions of the function f_1 at $x \in X$. Let

$$\gamma_\varphi(X,x) = \left\{g \in \mathbb{R}^n \mid \varphi'(x,g) \leq 0\right\}.$$

Lemma

Under our assumptions regarding the set X at any boundary point $x \in \text{bd } X$ the equalities

$$\Gamma(X, x) = \gamma_1(X, x) = \gamma_\varphi(X, x) = -\left[\text{cone } \{\partial(f_1(x))\}\right]^*$$

hold.

Example

. Suppose that

$$f_1(x) = 2|x_1| + |x_2| - 1, \ x = (x_1, x_2) \in \mathbb{R}^2.$$

The set X is given in the form (8)

$$\varphi(x) = \max\{0, f_1(x)\}.$$

Consider the point $\tilde{x} = (0,1) \in \text{bd}\, X$. We have

$$\partial f_1(\tilde{x}) = \text{co}\{(2,1), (-2,1)\}, \quad \partial\varphi(\tilde{x}) = \text{co}\{(0,0), (2,1), (-2,1)\}.$$

$$\text{cone}\ \{\partial f_1(\tilde{x})\} = \text{cone}\ \{\partial\varphi(\tilde{x})\},$$

Regularity condition

The implementation of exact penalty functions methods first of all depends on the properties of the function φ. Therefore various conditions are imposed on φ to make it possible to solve our problem.
Choose a point $z \in \mathbb{R}^n$ which does not belong to X, and project it onto this set. Suppose that a point $x \in X$ is the projection of z $(x = \text{pr}(z))$ onto the set X. Since X is closed, then the projection of z exists, although it may not be unique.

Consider the set

$$\mathcal{A}(X) = \{x \in \mathrm{bd}(X) \mid \exists z \notin X,\ x = \mathrm{pr}\ (z)\}.$$

At each point $x \in \mathcal{A}(X)$ the normal cone $N(X,x)$ is not only the zero vector.

Lemma

If

$$z \notin X \text{ and } \ x = \mathrm{pr}(z),$$

then the inclusion $(z - x) \in N(X,x)$.

Under considering optimization problems with constraints properties of the functions φ play an enormous role. In follows we assume that the function φ satisfies the regularity condition.

Regularity condition

We say, that a regularity condition is satisfied for the function φ if for each boundary point $x^* \in \text{bd } X$ there exist positive real numbers $\varepsilon(x^*)$ and $\beta(x^*)$, such that

$$\frac{o(\alpha, \bar{x}, g)}{\alpha} > -\varphi'(\bar{x}, g) + \beta(x^*), \tag{11}$$

$\forall \alpha \in (0, \varepsilon(x^*)]$, $\forall x \in \mathcal{A}(X) \cap S(x^*, \varepsilon(x^*))$, $\forall g \in N(X, x)$, $||g|| = 1$.

where

Here $S(x, r)$ denotes the closed ball of radius r centered at x_0

$$S(x, r) = \{x \in \mathbb{R}^n \mid \ ||x - x_0|| \le r\}.$$

The regularity condition is the condition on behavior of the function φ only at the boundary points of the set X.
We can prove that in order to satisfy the regularity condition it is necessary that the function φ is nondifferentiable at boundary points of X

Theorem

If the set X is compact and for the function φ regularity condition (11) is satisfied, then for the family of penalty functions

$$F(c_k, x) = f(x) + c_k \varphi(x)$$

there exists an exact penalty parameter c^*, *i.e.*

$$x(c_k) \in X \quad \forall c_k \geq c^*. \tag{12}$$

Consider a special case of the set X.
Let X be a non-empty compact set in $\mathbb{R}^n$.
As the function φ we take the function of the Euclidean distance from a point $x \in \mathbb{R}^n$ to the set X ($\varphi(x) = \rho(X, x)$).
Recall that $\rho(X, x)$ is the Lipschitz function with constant 1.
Then this function satisfies our requirements imposed on φ.

Theorem

For a family of penalty functions

$$F(c_k, x) = f(x) + c_k \rho(X, x)$$

we can take the Lipschitz constant of the function f on the set $\mathcal{L}(\varphi, x^{**})$ as exact penalty parameter .

Lemma

If the function φ, defining the set X of the form (8), at a point $x \in \mathrm{bd}\,(X)$ satisfies the regularity condition (11), then

$$\varphi'(x^*, g) \geq \beta(x) \quad \forall g \in N(X, x), \quad ||g|| = 1.$$

Corollary.

If the function φ at the point $x \in \text{bd}\,(X)$ satisfies the regularity condition (11), then the inequality

$$\min_{\substack{||g|| = 1 \\ g \in N(X,x)}} \varphi'(x,g) \geq \beta(x)$$

holds

Calculation of the exact penalty parameter for the set defined by a strongly convex function f_1

Let f_1 be a strongly convex function on $\mathbb{R}^n$ and $m > 0$ be its strong convexity constant then

$$f_1(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f_1(x_1) + \lambda_2 f_1(x_2) - m\lambda_1\lambda_2 ||x_1 - x_2||^2$$

$$\forall\ x_1, x_2 \in \mathbb{R}^n,\ x_1 \neq x_2,\ \lambda_1, \lambda_2 \in (0,1),\ \lambda_1 + \lambda_2 = 1.$$

Let x_1^* be a minimizer of the function f_1 on $\mathbb{R}^n$. Then $x_1^* \in \text{int } X$ because the Slater condition for the set X is satisfied.
Let $S(x_1^*, r)$ be a ball of the maximum radius centered at x_1^*, contained in a set X. Then we have

$$||v(x)|| \geq 2mr \ \forall v(x) \in \partial f_1(x), \ \forall x \notin S(x_1^*, r). \qquad (13)$$

Thus, for any boundary point of the set X the next inequality

$$||v(x)|| \geq 2mr \quad \forall x \in \ \text{bd} X$$

is true.

Theorem.

If the function f_1 is strongly convex, the set X is given in form (8), then the function φ, defining the set satisfies regularity condition (11) at any boundary point of the set.
In this case as the exact penalty parameter we can take the value of $\dfrac{L}{2mr}$, where L is the Lipschitz constant of f on the set $\mathcal{L}(\varphi, x^{**})$.

Example.

Suppose that

$$f_1(x) = \langle x, x\rangle - 1, \quad \varphi(x) = \max\{0, f_1(x)\}, \quad X = \{x \in \mathbb{R}^n \mid \varphi(x) = 0\}.$$

It is obvious that X is the unit ball in $\mathbb{R}^n$. It is easy to verify that

$$\min_{x \in \mathrm{bd}\, X} \quad \min_{\substack{\|g\| = 1, \\ g \in N(X,x)}} \quad \varphi'(x, g) \geq 2.$$

On minimizing of the maximum function of strongly convex functions

Consider the problem of an unconstrained minimization of the maximum function of strongly convex functions with a constant step.
Construct an optimization algorithm in which for finding the descent direction it is necessary to solve a quadratic programming problem.
It is proved that in this algorithm the generated sequence converges linearly to the minimizer of the objective function.

Introduction

Consider the problem of an unconstrained minimization of the maximum function of strongly convex functions with a constant step.
Let

$$f(x) = \max_{i \in I} f_i(x).$$

where $f_i(x)$, $i \in I = \{1, \ldots, s\}$, are twice continuously differentiable strongly convex functions. Denote by $m_i > 0, i \in I$, their strong convexity parameters. That is,

$$f_i(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f_i(x_1) + (1-\lambda)f_i(x_2) - m_i\lambda(1-\lambda)||x_1 - x_2||^2,$$

$$\forall \lambda \in [0,1], \quad \forall x_1, x_2 \in \mathbb{R}^n, \quad i \in I.$$

Then f is a strongly convex function where $m = \min_{i \in I} m_i > 0$ is its strong convexity parameter. Note that the function f can be assumed in the form

for each point x there exist numbers $\lambda_i(x),\ i \in I$

$$\lambda_i(x) \geq 0 \quad \forall i \in I, \quad \sum_{i=1}^{s} \lambda_i(x) = 1,$$

that

$$f(x) = \sum_{i=1}^{s} \lambda_i(x) f_i(x),$$

$$\lambda_i(x)(f_i(x) - f(x)) = 0 \quad \forall\, i \in I.$$

Consider the optimization problem: find

$$\min_{x \in \mathbb{R}^n} f(x). \tag{14}$$

The solution of problem (14) exists and a minimizer of f on $\mathbb{R}^n$ is unique.
Write the necessary minimum conditions of problem (14):
in order for a point x^* is a minimizer of the function f on $\mathbb{R}^n$ it is the necessary and sufficient that

$$0_n \in \partial f(x^*) = \text{co} \left\{ \bigcup_{i \in R(x^*)} f'(x^*) \right\}, \tag{15}$$

where

$$R(x^*) = \{ i \in I \mid f_i(x^*) = f(x^*) \}.$$

Note that condition (15) can be reformulated in this way:
in other for a point x^* is a minimizer of the function f on $\mathbb{R}^n$, it is necessary and sufficient that there exist such numbers $\lambda_i(x^*)$,

$$\lambda_i(x^*) \geq 0 \quad \forall\, i \in I, \quad \sum_{i=1}^{s} \lambda_i(x^*) = 1,$$

that

$$0_n = \sum_{i=1}^{s} \lambda_i(x^*) f_i'(x^*)$$

and

$$\lambda_i(x^*)(f_i(x^*) - f(x^*)) = 0 \quad \forall\, i \in I.$$

Assume that there is such constant $M \geq 1$, that the inequality

$$\langle f_i''(x)w, w\rangle \leq M||w||^2 \quad \forall\, x \in \mathbb{R}^n, \quad \forall\, w \in \mathbb{R}^n, \quad \forall\, i \in I, \qquad (16)$$

holds where $f_i''(x)$, $i \in I$, are matrices of second derivatives of the functions f_i at the point x.

With each point $x \in \mathbb{R}^n$ we connect the optimization problem: find

$$\min_{w \in \mathbb{R}^n} \max_{i \in I} \left\{ (f_i(x) - f(x))M + \langle f_i'(x), w \rangle + \frac{1}{2}||w||^2 \right\}. \qquad (17)$$

The solution of this problem $w(x) \in \mathbb{R}^n$ is also unique.

$$w(x) = \arg \min_{w \in \mathbb{R}^n} \max_{i \in I} \left\{ (f_i(x) - f(x))M + \langle f_i'(x), w(x) \rangle + \frac{1}{2}||w(x)||^2 \right\}.$$

Denote by

$$\nu(x) = \max_{i \in I} \left\{ (f_i(x) - f(x))M + \langle f_i'(x), w(x) \rangle + \frac{1}{2}||w(x)||^2 \right\}.$$

From the necessary and sufficient conditions of problem (17) follow the existence of such coefficients

$$\lambda_i(x) \geq 0, \quad i \in I, \quad \sum_{i=1}^{s} \lambda_i(x) = 1,$$

and

$$w(x) = -\sum_{i=1}^{s} \lambda_i(x) f_i'(x),$$

$$\nu(x) = \sum_{i=1}^{s} \lambda_i(x)(f_i(x) - f(x))M - \frac{1}{2}||w(x)||^2 \leq 0.$$

Lemma

1) The function $\nu(x)$ is equal to zero if and only if the point x is a minimizer of the function f on $\mathbb{R}^n$.
2) If the vector $w(x) = 0_n$, then $\nu(x) = 0$.
3) The direction $w(x)$ is a descent direction of the function f at the point x.

Describe a method for minimizing a function f on $\mathbb{R}^n$.

Take an arbitrary point $x_0 \in \mathbb{R}^n$. If x_0 is a minimizer of the function f on $\mathbb{R}^n$ then the process stops.
Suppose that we have found a point $x_k \in \mathbb{R}^n$. If x_k is a minimizer of the function f on $\mathbb{R}^n$ then the process stops.
Suppose that x_k is not a minimizer of f on $\mathbb{R}^n$.
Define the descent direction $w(x_k)$, that is, solve the optimization problem (17) if $x = x_k$.

Then there are coefficients

$$\lambda_i(x_k) \geq 0, \quad i \in I, \quad \sum_{i=1}^{s} \lambda_i(x_k) = 1,$$

that

$$w(x_k) = -\sum_{i=1}^{s} \lambda_i(x_k) f_i'(x_k) \neq 0_n,$$

$$\nu(x_k) = \sum_{i=1}^{s} \lambda_i(x_k)(f_i(x_k) - f(x_k))M - \frac{1}{2}||w(x_k)||^2 < 0.$$

Put

$$\bar{\alpha} = \frac{1}{M} \quad \text{and} \quad x_{k+1} = x_k + \bar{\alpha} w(x_k).$$

Lemma.

At each point x_k the next inequality

$$f(x_{k+1}) - f(x_k) \leq \frac{1}{M}\nu(x_k) \tag{18}$$

holds.
If the sequence $\{x_k\}$ is infinite, then

$$f(x_{k+1}) - f(x_k) \to 0.$$

Inequality (18) shows that the process is relaxing.

Lemma

In this method at each point x_k the inequality

$$\nu(x_k) \leq m\left(f(x^*) - f(x_k)\right), \tag{19}$$

is true, where x^* is a minimizer of f on $\mathbb{R}^n$.
From (18) and (19) the inequality

$$f(x_{k+1}) - f(x_k) \leq \frac{m}{M}(f(x^*) - f(x_k)) \tag{20}$$

holds.

Theorem.

With an arbitrary starting point $x_0 \in \mathbb{R}^n$ the sequence $\{x_k\}$ converges to the minimizer x^* of f with the rate of geometric progression:

$$f(x_{k+1}) - f(x^*) \leq q(f(x_k) - f(x^*)),$$

where $q = 1 - \dfrac{m}{M}$, and there is a positive number $Q > 0$, for which the next inequality

$$||x_k - x^*|| \leq Q(\sqrt{q})^k$$

holds.

The main computational difficulty is the determination of a descent direction $w(x_k)$. Problem (17) can be reduced to the quadratic programming problem.
Denote by

$$z = \max_{i \in I} \left\{ (f_i(x) - f(x))M + \langle f_i'(x), w(x) \rangle \right\}.$$

Then problem (17) is equivalent to the following quadratic programming problem in the space $\mathbb{R}^n \times \mathbb{R}$:
to minimize the function

$$z + \frac{1}{2}\langle w, w\rangle$$

under constraints

$$T = \{[w, z] \in \mathbb{R}^n \times \mathbb{R} \mid \langle f_i'(x), w(x)\rangle - z \leq (f(x) - f_i(x))M,\ i \in I\}.$$

Example.

As the test functions we consider a maximum function

$$f(x) = \max_{i \in I} f_i(x), \quad x \in \mathbb{R}^{10}, \quad I = 1, \dots, 5.$$

where

$$f_i(x) = \langle A_i x, x \rangle + \langle b_i, x \rangle, \quad b_i \in \mathbb{R}^{10}, \quad i \in I,$$

$$a_i(j,k) = e^{j/k} \cos(j \cdot k) \sin i, \quad j < k, \quad a_i(j,k) = a_i(k,j), \quad j > k,$$

$$a_i(i,i) = 2\frac{|\sin i| \cdot i}{j} + \sum_{k \neq j} |a_i(j,k)|,$$

$$b_i(j) = e^{j/i} \sin(i \cdot j), \quad j,k \in 1:10.$$

Matrices $A_i, i \in I$, are positive definite and Table 1 shows, respectively, the maximum (λ_{max}) and minimum (λ_{min}) eigenvalues of matrices $2A_i$, the value of each of the functions $f_i, i \in I$, at the minimizer x^* and the norms of the gradient of these functions at x^*. For our function $M = 36.327, \ m = 0.69$.

The point $x^* =$ $(-0.0546, -0.0241, -0.0057, 0.0231, 0.0558, -0.2434,$ $0.0685, 0.1321, 0.0772, 0.0336)^T \in \mathbb{R}^{10}$ is a minimizer of $f(x)$ on $\mathbb{R}^n$, $f(x^*) = -0.72576$ is a solution of the problem, $||w(x^*)||$ $= 0.000092, \;\; R(x^*) = \{2, 3, 4, 5\}$, where

$$R(x^*) = \{i \in I \mid f_i(x^*) = f(x^*)\}.$$

i	$\lambda_{\max}$	$\lambda_{\min}$	$f_i(x^*)$	$\|\|f_i'(x^*)\|\|$
1	26.973	7.154	-311.19750	12805.888
2	29.793	8.306	-0.725756	155.679
3	4.735	1.374	-0.725756	37.765
4	27.507	9.641	-0.725756	14.479
5	36.327	12.909	-0.725756	5.934

Table 1.

Table 1 shows that the function f_1 is not active at the optimum point, but because of its gradient at x^* is large enough at the point $x^* + \lambda f'(x^*)$ for small λ, it becomes active.
It is obvious that in most cases the number M is not known beforehand. Therefore, you can use this algorithm for various M.

For this example the minimizer x^* was obtained for 85 steps if $M = 145.28$, when $M = 72.64$ we have 38 steps when $M = 36.32$ we have 19 steps when $M = 18.16$ we have 11 steps.
If $M = 9$ and less algorithm didn't make sense.
The problem was considered solved if the condition $||w(x_k)|| < 10^{-4}$ is performed.