Journal of Global Optimization (2005) 31: 437–453
DOI 10.1007/s10898-004-7018-4

On the Solution of Min-sum-min Problems

ALEXEI DEMYANOV
Faculty of Mathematics and Mechanics, Saint-Petersburg State University, Staryi Peterhof, 198504, St. Petersburg, Russia (e-mail: alex@ad9503.spb.edu)

(Received: 11 March 2003; revised: 10 December 2003; accepted: 4 January 2004)

Abstract. One class of min-sum-min problems is discussed in the paper. Min-sum-min problems appear in a natural way in many applications (e.g., in cluster analysis, pattern recognition, classification theory etc.). Like min-max-min problems, min-sum-min problems represent a very important family of nonsmooth problems. Problems of this type can be treated by means of the existing tools of Nonsmooth Analysis. However, most of algorithms available provide a local minimizer only, since they are based on necessary conditions which are of local nature. In the paper it is proved that the original problem can be reduced to the problem of minimizing a finite number of sum-functions. A necessary condition for a global minimum and a sufficient condition for a local minimum are stated. The necessary condition is of nonlocal nature. An algorithm (so-called Exchange algorithm) for finding points, satisfying necessary conditions, is described. An ε-Exchange algorithm is formulated, allowing, in principle, to escape from a 'shallow' local minimizer. An example is presented to illustrate the results and algorithms. An application of the proposed algorithms to solving one clustering problem is also given. Numerical results are provided.

AMS Subject Classifications: 90C30, 49J40.

Key words: cluster analysis, exchange algorithm, ε-exchange algorithm, necessary conditions for a minimum, sufficient conditions, sum-min function.

1. Introduction and Statement of the Problem

In this paper we are concerned with the following problem:

Let sets $I=\{1,\dots,N\}$ and $J_i=\{1,\dots,m_i\}, i\in I$, be given, where N and m_i are positive integers. Define the function

$$F(x)=\sum_{i\in I}\min_{k\in J_i}\varphi_{ik}(x),$$

where $x\in\mathbb{R}^n, \varphi_{ik}:\mathbb{R}^n\to\mathbb{R} \quad (n\in\mathbb{N})$.

PROBLEM P. Find a point $x^*\in\Omega$, such that

$$F(x^*)=\min_{x\in\Omega}F(x),$$

where $\Omega\subset\mathbb{R}^n$. That is, it is required to find a minimizer of the functional F on the set Ω.

Min-sum-min problems appear in a natural way in many applications (e.g., in cluster analysis, pattern recognition, classification theory etc., see (Bagirov et al., 1999; Bagirov, Rubinov and Yearwood, 2002; Bagirov, Rubinov and Yearwood, 2001). Like min-max-min problems, min-sum-min problems represent a very important family of nonsmooth problems.

Problems of this type can be treated by means of the existing tools of Nonsmooth Analysis (see, e.g. (Pardalos and Resende, 2002)). However, most of algorithms available provide a local minimizer only, since they are based on necessary conditions which are of local nature. In particular, it is very difficult to overcome some 'specific' obstacles, which are the main features of this class of problems:

– Essential nonsmoothness (even if φ_{ik}'s are smooth)
– Large number of 'shallow' local minimizers.

As it will be shown below by examples, even if the functions φ_{ik}'s are 'very good' (say, real-analytic), it doesn't help in any way to solve this problem.

An important property of sum-min functions to be exploited here is that the problem of minimizing such a function can be reduced to solving a finite (though maybe very large) number of relatively simple min-sum-type problems. Similar property holds for min-max-min problems (Vershik, Malozemov and Pevnyi, 1975) and was crucial for developing numerical algorithms in (Demyanov, 2003; Demyanov, 2002; Demyanov, Demyanov and Malozemov, 2002).

In this paper we will describe a new approach to solving such problems. Obviously, we will need to assume some properties of the functions under consideration – in particular, even though the optimality condition does not require anything from the functions φ_{ik}'s, for constructing numerical methods for finding minimizers, we will suppose that we are able to solve a 'more simple' kind of problem, involving φ_{ik}'s.

One problem of this type was discussed in (Demyanov, 2003). Namely, the case where $|J_i| = 2$ and the functions φ_{ik} are of the form $\varphi_{ik}(x_k)$, with $x = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}, n_1 + n_2 = n$.

The Paper is organized as follows: In Section 2 an equivalent formulation of the problem is given. It is proved that the original problem is reduced to the problem of minimizing a finite number of sum-functions. A necessary condition for a global minimum and a sufficient condition for a local minimum are stated in Section 3. The necessary condition is of nonlocal nature. An algorithm (so-called Exchange algorithm) for finding points, satisfying necessary conditions, is described in Section 4. In Section 5 an ε-Exchange algorithm is formulated. Using this algorithm it is possible, in principle, to escape from a 'shallow' local minimizer. An example is presented in Section 6 to illustrate the results and algorithms. An application of the proposed algorithms to solving one clustering problem is shown in Section 7. Numerical results are also provided.

2. An Equivalent Problem Formulation

So, our aim is to find a minimizer of F on the set $\Omega \subset \mathbb{R}^n$. Put

$$J = J_1 \times \cdots \times J_N$$

and for every $j = (j_1, \dots, j_N) \in J$ let us introduce the function

$$F_j(x) = \sum_{i \in I} \varphi_{i j_i}(x). \tag{1}$$

Evidently,

$$F(x) \leq F_j(x), \tag{2}$$

for every $j \in J$.

Let us formulate another

PROBLEM P1. Find $j^* \in J$ such that

$$\inf_{x \in \Omega} F_{j^*}(x) = \min_{j \in J} \inf_{x \in \Omega} F_j.$$

Now we will prove, that the Problem P is equivalent to the Problem P1.

THEOREM 1. *The following equality is valid*:

$$\inf_{x \in \Omega} F(x) = \min_{j \in J} \inf_{x \in \Omega} F_j(x). \tag{3}$$

Proof. From (2) for every $x \in \Omega$ we have

$$\inf_{x \in \Omega} F(x) \leq \inf_{x \in \Omega} F_j(x) \tag{4}$$

for each $j = (j_1, \dots, j_N) \in J$. Since it is true for an arbitrary $j \in J$, (4) implies

$$\inf_{x \in \Omega} F(x) \leq \min_{j \in J} \inf_{x \in \Omega} F_j(x). \tag{5}$$

Now let us take an arbitrary $\overline{x} \in \Omega$. Put $j(\bar{x}) = (j_1(\bar{x}), \dots, j_N(\bar{x}))$, where $j_i(x) \in J_i$ are such that

$$\varphi_{ij_i(\bar{x})}(\bar{x}) = \min_{k \in J_i} \varphi_{ik}(\bar{x}).$$

Then

$$F(\bar{x}) = \sum_{i \in I} \varphi_{ij_i(\bar{x})}(\bar{x}) = F_{j(\bar{x})}(\bar{x}) \geq \inf_{x \in \Omega} F_{j(\bar{x})}(x) \geq \min_{j \in J} \inf_{x \in \Omega} F_j(x).$$

This inequality is valid for an arbitrary $\bar{x} \in \Omega$ and we have

$$\inf_{x \in \Omega} F(x) \geq \min_{j \in J} \inf_{x \in \Omega} F_j(x). \tag{6}$$

The inequalities (5) and (6) imply (3). □

This theorem has two remarkable consequences:

1. It shows, that the global minimizer of F on the set Ω exists, by modulo of existence of a minimizer of the function F_j on Ω, for all $j \in J$.
2. Due to (3) the problem of minimizing the functional F is reduced to solving a finite number of problems of minimizing functions of the form (1). So, we have proved that if we are able to minimize functions of the type (1), then in principle it is possible to solve the Problem P by solving the problem of minimizing F_j for each $j \in J$. However, in real applications the number

$$|J| = \prod_{i \in I} |J_i| = \prod_{i \in I} m_i \tag{7}$$

is quite large and it is practically impossible to use this equivalence directly. However, making use of this result, we will develop some methods for finding local minimizers of the functional F.

3. A Necessary Condition for a Global Minimum and a Sufficient Condition for a Local Minimum

Let $x \in \mathbb{R}^n$. Consider *the set of active indices* at the point x, defined in the following way:

$$J(x, 0) = \{j \in J \mid \varphi_{ij_i}(x) = \min_{k \in J_i} \varphi_{ik}(x) \forall i \in I\} \subset J. \tag{8}$$

It is easy to see, that

$$F_j(x) = \sum_{i \in I} \varphi_{ij_i}(x) = F(x) \;\; \forall j \in J(x, 0). \tag{9}$$

THEOREM 2. *For a point $x^* \in \Omega$ to be a global minimizer of F on the set Ω it is necessary that*

$$\inf_{x\in\Omega} F_j(x) = F(x^*)\forall j \in J(x^*, 0). \tag{10}$$

If at some point $x^* \in \Omega$ condition (10) *holds and $\varphi_{ik} \in C(\Omega)$, then x^* is a local minimizer of F on Ω.*

Proof. Necessity. We have

$$\inf_{x\in\Omega} F_j(x) \leq F_j(x^*) = F(x^*)\forall x^* \in \Omega, \quad \forall j \in J(x^*, 0). \tag{11}$$

Since x^* is a global minimizer of F on the set Ω, condition (2) implies that

$$\inf_{x\in\Omega} F_j(x) \geq \inf_{x\in\Omega} F(x) = F(x^*). \tag{12}$$

Thus, equality (10) follows from (11) and (12).

Sufficiency. Suppose, that condition (10) holds, and the point x^* is not a local minimizer of F on the set Ω. Then for any $\delta > 0$ there exists $x_\delta \in B(x^*, \delta) \cap \Omega$, such that $F(x_\delta) < F(x^*)$. We have

$$F(x_\delta) = F_{j(\delta)}(x_\delta) = \sum_{i\in I} \varphi_{ij_i(\delta)}(x_\delta),$$

where $j(\delta) = (j_1(\delta), \ldots, j_n(\delta)) \in J(x_\delta, 0)$. Since φ_{ik} are continuous on Ω, then the inclusion $J(x_\delta, 0) \subset J(x^*, 0)$ holds for all $x_\delta \in B(x^*, \delta) \cap \Omega$ if δ is small enough. Thus, $j(\delta) \in J(x^*, 0)$. Hence, we get

$$\inf_{x\in\Omega} F_{j(\delta)}(x) \leq F_{j(\delta)}(x_\delta) < F(x^*),$$

which contradicts (10). The theorem is proved. □

Condition (10) is a necessary condition for a point to be a global minimizer and a sufficient condition for a local minimum of the functional F on the set Ω. A point $x^* \in \Omega$ satisfying condition (10) is called a stationary point of F on Ω. Note, that this condition is of a nonlocal nature. Though this theorem implies that each stationary point is a local minimizer, the converse is not true: not every local minimizer is a stationary point (see Section 6, Remark 4).

4. An Exchange Algorithm

In this section we will construct a method for finding stationary points, based on the necessary and sufficient condition for a minimum.

Suppose, that $\varphi_{ik} \in C(\Omega)$ and for every $j \in J$ there exists (*and we are able to find it*) a point $x_j \in \Omega$ such that

$$F_j(x_j) = \min_{x \in \Omega} F_j(x).$$

Condition (10) leads us to the following algorithm:

1. Choose an arbitrary $u_0 \in \Omega$.
2. Let a point $u_s \in \Omega$ be given.
3. Construct the set $J(u_s, 0)$ and for every $j \in J(u_s, 0)$ check whether condition (10) is satisfied, that is, check the equality

$$F_j(x_j) = F(u_s). \tag{13}$$

– If it holds for all $j \in J(u_s, 0)$ then the point u_s is a local minimizer of F on Ω and the algorithm terminates.
– Otherwise, there exists $j^* \in J(u_s, 0) \subset J$, such that

$$\min_{x \in \Omega} F_{j^*}(x) = F_{j^*}(x_{j^*}) < F(u_s). \tag{14}$$

Put $u_{s+1} = x_{j^*}$. The inequalities (2) and (14) yield

$$F(u_{s+1}) \leq F_{j^*}(u_{s+1}) < F(u_s). \tag{15}$$

Go to step 2.

Since the number of points in the set J is finite, inequality (15) implies that the algorithm terminates in a finite number of steps, resulting in a stationary point $u^* \in \Omega$ (which is a local minimizer of F).

Remark 1. The algorithm described above can be used, if at every step the number $|J(u_s, 0)|$ is substantially less then $|J|$. That is, to check the minimality condition we will not need to solve a large number of problems of minimizing the functions F_j's.

5. An ε-Exchange Algorithm

An Exchange algorithm allows us to find a stationary point (which is a local minimizer) of F in a finite number of steps. However, the functional F may possess a huge number of local minimizers on the set Ω, and even though the definition of a stationary point removes some of them from our consideration (since not every local minimizer is a stationary point), we need some more efficient methods to 'improve' the minimizer. The following method allows us to 'leave' a local minimizer and to find a 'better' one (in the sense of minimizing F).

Again, we will assume that $\varphi_{ik} \in C(\Omega)$ and we are able to solve the problem of minimizing the function F_j (as in (1)), for every $j \in J$.

1. Fix $\varepsilon > 0$. Let $v_0 \in \Omega$ be a stationary point. Construct *the set of ε-active indices* $J(v_0, \varepsilon)$ at the point v_0 as follows:

$$J(v_0, \varepsilon) = \{j \in J \mid \varphi_{ij_i}(v_0) \leq \min_{k \in J_i} \varphi_{ik}(v_0) + \varepsilon \ \ \forall i \in I\}.$$

For each $j \in J(v_0, \varepsilon)$ let us find a point x_j, such that

$$\min_{x \in \Omega} F_j(x) = F_j(x_j). \tag{16}$$

2. Let a stationary point $v_t \in \Omega$ be given.
   - If

$$\min_{j \in J(v_t, \varepsilon)} F_j(x_j) = F(v_t), \tag{17}$$

   then the algorithm terminates. A point v_t, satisfying (17), is called an *ε-local minimizer* of F on the set Ω.
   - If

$$\min_{j \in J(v_0, \varepsilon)} F_j(x_j) = F_{j(\varepsilon)}(x_{j(\varepsilon)}) < F(v_t),$$

   then applying the Exchange Algorithm with $x_{j(\varepsilon)}$ as the initial point, in a finite number of steps we obtain a stationary point $v_{t+1} \in \Omega$, and

$$F(v_{t+1}) \leq F(x_{j(\varepsilon)}) \leq F_{j(\varepsilon)}(x_{j(\varepsilon)}) < F(v_t).$$

Again, since $|J|$ is finite, the algorithm terminates in a finite number of steps.

Remark 2. Note, that taking ε quite large we obtain $J(v, \varepsilon) = J$. In this case the checking of ε-minimality is equivalent to the solution of the Problem P1 (see Theorem 1).

Remark 3. While executing the Exchange or the ε-Exchange Algorithm, at every step we need to check conditions (13) and (17) respectively. However, note, that to prove that a point is not a stationary point (or ε-local minimizer), we don't need the precise solution of these problems. If, for example, the problem of minimizing F_j's is complicated from the numerical point of view, it is sufficient to find only one index j and one point $x(j)$, such that the value of $F_j(x(j))$ is strictly less then the value of F at the given point. Therefore, the following strategy can be adopted: at first steps of the algorithms we use some rough method, just to find a point, which delivers strictly smaller value to F, not spending much efforts for this, and later we start using more sophisticated methods for finding precise values of $\min_{x\in\Omega} F_j(x)$.

As it was mentioned before, this problem is essentially nonsmooth and extremely multi-extremal. We have made several steps to become able to solve this type of problems:

1. The definition of a stationary point excludes some local minimizers from our considerations (since not every local minimizer is a stationary point). It allows one to avoid considering some of 'shallow' local minimizers.
2. The Exchange Algorithm can be used for finding stationary points.
3. The ε-Exchange Algorithm can be used to find a better local minimizer by extending the set of active indices. By varying ε we can balance between the 'depth' of the search and the 'quality' of the minimizer.

6. An Example

To illustrate the described method let us consider the following example.

Let $\Omega = \mathbb{R}$, $I = \{1\}$, $J_1 = \{1, 2\}$, that is the functional F has the form

$$F(x) = \min_{k\in J_1} \varphi_{1k}(x),$$

where

$$\varphi_{11}(x) = \frac{1}{2}x(x+2)(x-2)^2, \quad \varphi_{12}(x) = 10(x-1)(x+1).$$

Let $u_0 = 2$, it is easy to see (Figure 1) that u_0 is a local minimizer of F on Ω. Let us check the condition (10): $J(u_0, 0) = \{1\}$, since $\varphi_{11}(u_0) = 0$, $\varphi_{12}(u_0) = 30$. Thus, $F_1(x) = \varphi_{11}(x)$. Since

$$\inf_{x\in\mathbb{R}} F_1(x) = F_1(x_1) < -4 < 0 = F(u_0), \tag{18}$$

where $x_1 = \frac{1}{4}(-1 - \sqrt{17})$, we conclude that the condition (10) fails to hold, and therefore u_0 is not a stationary point.

Remark 4. The above example reveals the fact that though a stationary point is a local minimizer (for continuous functions), the converse is not true: not every local minimizer is a stationary point.

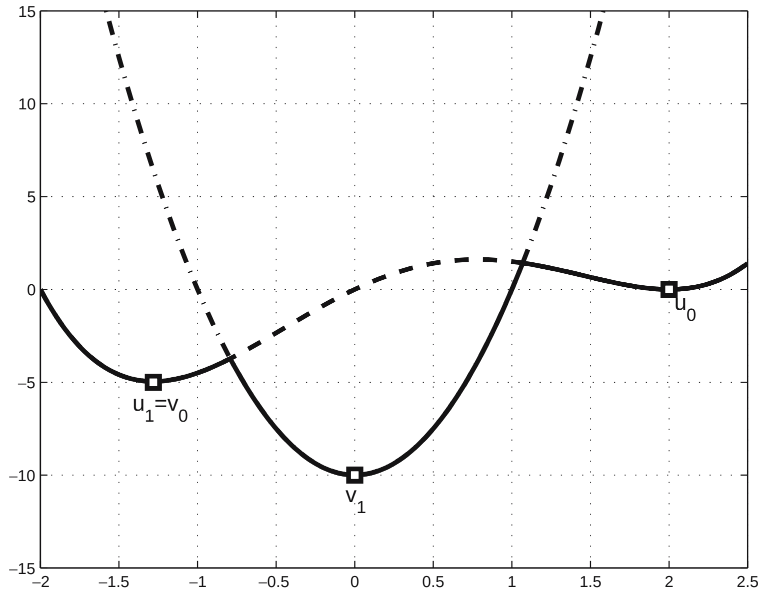

Figure 1. F, φ_{11} and φ_{12}.

Now, according to the Exchange Algorithm, we will check whether the point $u_1 = x_1$ satisfies the optimality condition. We have $J(u_1, 0) = \{1\}$ and hence the condition (10) and the equality (18) imply that u_1 is a stationary point for the functional F on the set Ω (see Figure 1).

In order to escape from this local minimizer, let us apply the ε-Exchange Algorithm.

Put $\varepsilon = 13$, $v_0 = u_1$. We have $J(v_0, \varepsilon) = \{1, 2\}$. So, $F_1(x) = \varphi_{11}(x)$, $F_2(x) = \varphi_{12}(x)$. However, since

$$\inf_{x \in \mathbb{R}} F_2(x) = F_2(v_1) = -10 < -5 < F(v_0), \tag{19}$$

where $v_1 = 0$, we conclude that v_0 is not an ε-local minimizer for $\varepsilon = 13$.

It is easy to see that since $J(v_1, \varepsilon) = \{1, 2\}$, the equality (19) implies that v_1 is an ε-local minimizer for F on the set Ω for $\varepsilon = 13$.

It is easy to see that v_1 is an ε-local minimizer of F for every $\varepsilon > 0$, so one concludes that v_1 is a global minimizer for F on R.

Remark 5. This example illustrates, that even if the functions φ_{ij}'s are analytic, it doesn't make the problem of minimizing F much easier. Nice properties that φ_{ij}'s enjoy can help minimizing F_j's, but the condition (10) still needs to be checked.

7. An Application to Clustering Problems

Let a set of points $\{t_1, \dots, t_N\} \subset \mathbb{R}^n$ be given. Let $I = \{1, \dots, N\}$, $J_i = \{1, 2, 3\}$ for all $i \in I$, and $x = (x^1, x^2, x^3) \in \Omega = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$. Introduce the functions

$$\varphi_{i1}(x) = \|t_i - x^1\|^2, \quad \varphi_{i2}(x) = \|t_i - x^2\|^2, \quad \varphi_{i3}(x) = \|t_i - x^3\|^2$$

where $\|x\|^2 = \langle x, x\rangle$.

Note that φ_{ik}'s are continuously differentiable and convex. Consider the following clustering problem.

PROBLEM CP. Find $x^* = (x^{*1}, x^{*2}, x^{*3}) \in \Omega = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ such that

$$F(x^*) = \min_{x \in \Omega} F(x),$$

where

$$F(x) = \sum_{i \in I} \min_{k \in J_i} \varphi_{ik}(x).$$

Take any $j \in J = J_1 \times \cdots \times J_N$, then the functions $F_j(x)$, defined in (1) take the form

$$F_j(x) = \sum_{i \in I} \varphi_{ij_i}(x) = \sum_{i \in \sigma_1(j)} \varphi_{i1}(x) + \sum_{i \in \sigma_2(j)} \varphi_{i2}(x) + \sum_{i \in \sigma_3(j)} \varphi_{i3}(x),$$

where

$$\sigma_k(j) = \{i \in I \mid j_i = k\}, \quad k = 1, 2, 3.$$

Clearly,

$$\min_{x \in \Omega} F_j(x) = F_j(x_j), \tag{20}$$

where $x_j = (x^1(j), x^2(j), x^3(j))$, and

$$x^k(j) = \frac{1}{|\sigma_k(j)|} \sum_{i \in \sigma_k(j)} t_i. \tag{21}$$

Remark 6. Often in clustering problems the following functional is considered:

$$F(x) = \sum_{i \in I} \min_{k \in J_i} \varphi_{ik}(x),$$

where

$$\varphi_{ik}(x) = \|t_i - x^k\|.$$

We use $\varphi_{ik} = \|t_i - x^k\|^2$, since in this case we have the explicit formula (21) for finding minimizers of F_j's. That is, we don't need to spend much efforts for minimizing F_j's, while solving the problem of minimizing F.

7.1. EXAMPLE

Let the set of points $\{t_1, \ldots, t_{41}\} \subset \mathbb{R}^2$ be as shown in Table I. We have to find 3 clusters $x^1, x^2, x^3 \in \mathbb{R}^2$ which minimize the functional

$$F(x) = \sum_{i \in I} \min\{\|t_i - x^1\|^2,\ \ \|t_i - x^2\|^2,\ \ \|t_i - x^3\|^2\},$$

where $I = \{1, \ldots, 41\}$. Here $\varphi_{ik} = \|t_i - x^k\|^2$ and $J_i = \{1, 2, 3\}$ for all $i \in I$.

First, let us solve 1-cluster problem

$$\min_{x \in \mathbb{R}^2} F_1(x) = F_1(y^*),$$

where

$$F_1(x) = \sum_{i \in I} \|t_i - x\|^2.$$

From (21) it follows that $y^* = (3.8902, 5.6098)$ and $F_1(y^*) = 200.432$.

Table I. The set of points $\{t_1, \ldots, t_{41}\} \subset \mathbb{R}^2$

i	t_i	i	t_i	i	t_i	i	t_i	i	t_i
1	(2, 6)	9	(7, 2)	17	(4, 7)	25	(2, 1)	33	(1.5, 6)
2	(1, 6)	10	(6, 3)	18	(6, 7)	26	(2, 0)	34	(2.5, 6)
3	(3, 5)	11	(6, 1)	19	(6, 6)	27	(1, 5)	35	(5.5, 8)
4	(2, 7)	12	(5, 3)	20	(4, 5)	28	(2, 8)	36	(4.5, 8)
5	(2, 5)	13	(7, 3)	21	(4, 6)	29	(4, 9)	37	(4.5, 7)
6	(3, 5)	14	(5, 7)	22	(1, 7)	30	(5, 9)	38	(5.5, 7)
7	(6, 2)	15	(5, 8)	23	(1, 8)	31	(1.5, 5)	39	(6, 9)
8	(5, 2)	16	(5, 6)	24	(2, 2)	32	(2.5, 5)	40	(5.5, 9)
								41	(6, 8)

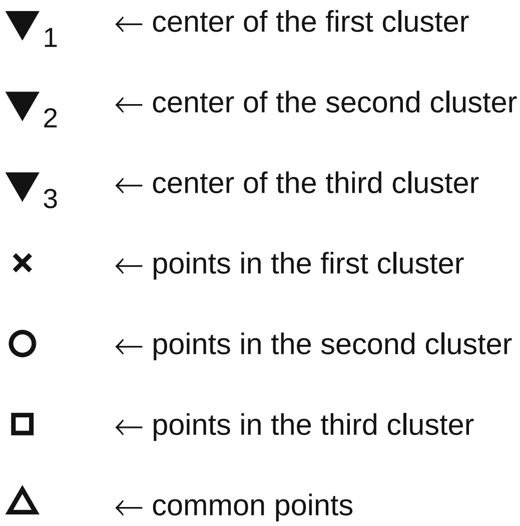

Figure 2. The legend.

At the first sight, it seems interesting to put $u_0 = (y^*, y^*, y^*) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$ as the initial point for applying the Exchange Algorithm. However, in this case $J(u_0, 0) = J = J_1 \times \cdots \times J_{41}$, so just to check the optimality condition (10) it is necessary to solve $|J| = 3^{41}$ problems of minimizing functions F_j's, which is impossible.

That is why we will slightly move the centers: say, let's take $u_0 = (u_0^1, u_0^2, u_0^3)$, with $u_0^1 = y^*$, $u_0^2 = y^* + \Delta_1$ and $u_0^3 = y^* + \Delta_2$, where $\Delta_1 = (0.1, 0)$ and $\Delta_2 = (-0.1, 0.1)$.

Starting the Exchange Algorithm, in 5 steps we obtain a stationary point $u_1 = (u_1^1, u_1^2, u_1^3)$, with $u_1^1 = (2, 1)$, $u_1^2 = (6, 2.2857)$, $u_1^3 = (3.5968, 6.8065)$. The function value is $F(u_1) = 155.477$. The results of application of the Exchange Algorithm are depicted in Figure 3.

The notations used in the Figures are explained in Figure 2. Common points denote the points, whose indices generate the set of ε-active indices $J(v, \varepsilon)$.

Now put $\varepsilon = 8$ and let us take $v_0 = u_1$ as the initial point for the ε-Exchange Algorithm. t_{27} and t_{19} are the ε-common points, that is the indices $\{19, 27\}$ generate the set of ε-active indices $J(v_0, 8)$.

At the first iteration of ε-Exchange algorithm we obtained the point $v_1 = (v_1^1, v_1^2, v_1^3)$, where $v_1^1 = (2, 6)$, $v_1^2 = (4.8, 1.9)$, $v_1^3 = (5.0938, 7.5625)$ with $F(v_1) = 99.547$. The situation is illustrated in Figure 4.

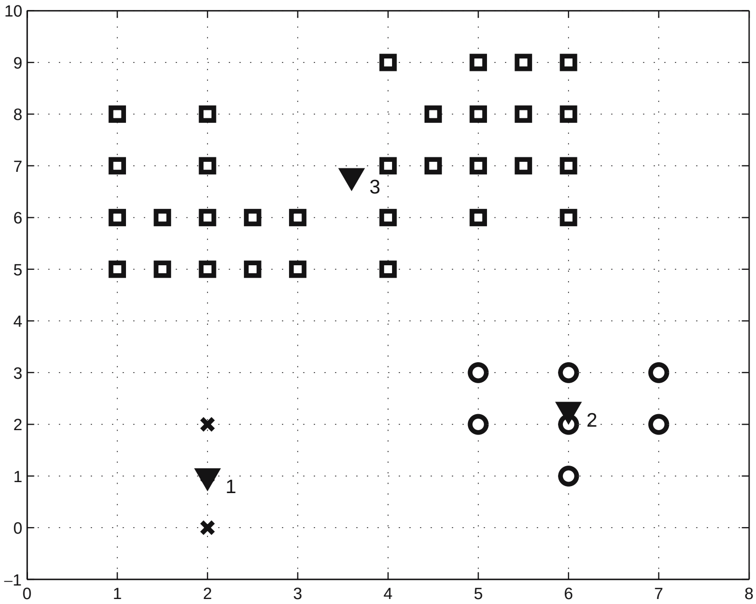

Figure 3. The results of the Exchange algorithm, $F(u_1) = 155.47$

.

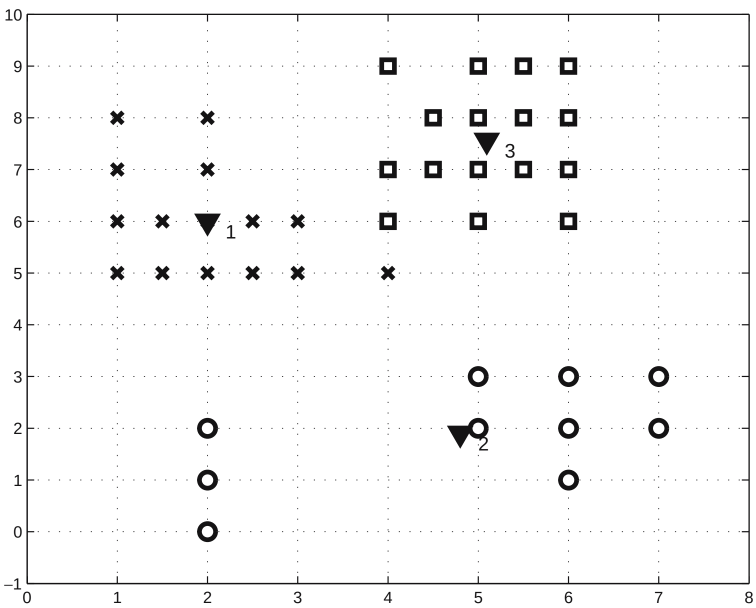

Figure 4. The first iteration of the ε-Exchange algorithm, $F(v_1) = 99.547$

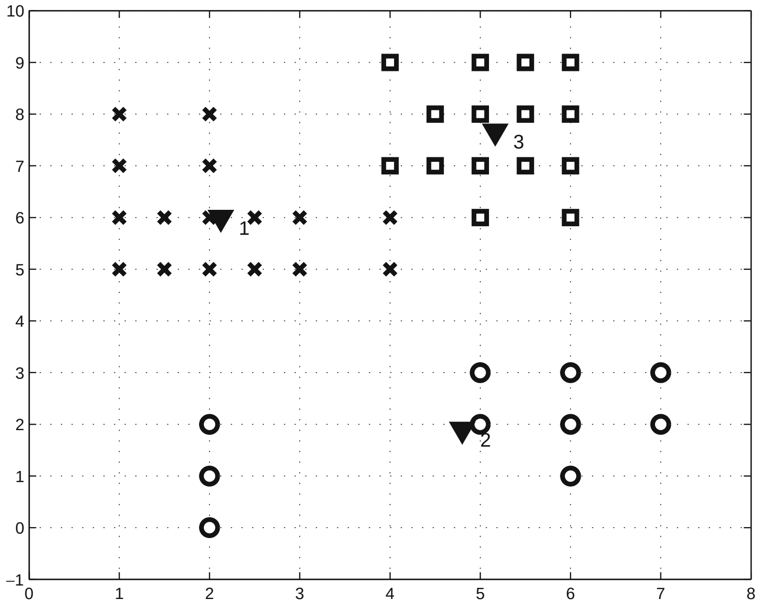

Figure 5. An ε-local minimizer v_2 for $\varepsilon = 8$ with $F(v_2) = 99.417$

At the second iteration of the algorithm we get the point $v_2 = (v_2^1, v_2^2, v_2^3)$ where $v_2^1 = (2.125, 6)$, $v_2^2 = (4.8, 1.9)$, $v_2^3 = (5.1667, 7.6667)$ with $F(v_2) = 99.417$. This point is an ε-local minimizer of F for $\varepsilon = 8$. The resulting clusters are shown in Figure 5.

7.2. NUMERICAL RESULTS

We have applied the described algorithm to several datasets: Cleveland Heart Disease (Heart) and Diabetes. The results of 10-fold validation are depicted in Table II.

1. Cleveland Heart Decease Database. We have chosen 6 features, and solved 3-cluster problem. We obtained 3 clusters, and then we say that the cluster is 'of the 1st type', if the points from the 1st set dominate. In the Table III you can see the detailed description of what we obtained when applied the Exchange algorithm to the whole set. As you can see, in the 1st cluster the points from the second set dominate, but 17 of 74 points are misidentified. It means that, if we have a new point (that is,

Table II. Numerical results

	Heart	Diabetes
Dataset size	297	768
Dataset dimension	13	8
Number of clusters	3	6
10-fold training set correctness, %	82.98	76.45
10-fold testing set correctness, %	82.95	76.33

Table III. Heart results

Cluster number	1	2	3
Type of the cluster	2	2	1
Size of the cluster	74	65	158
Number of misidentified points	17	8	23
Percentage of misidentified points	23%	12%	15%

we have to make a diagnosis), and if it proves to be in the 1st cluster, then it is very probable (since the error is 23 percent) that this point is also from the 2nd set.

The total number of misidentified points is 48 (that is, the overall accuracy is 83.84 percents), and after 10-fold validation we obtained what is described in the Table II.

2. Diabetes Database.
In this case we were looking for 6 clusters. As it is shown it Table IV, the 2nd cluster contains mainly the points of the 1st type (7% of mis-

Table IV. Diabetes results

Cluster number	1	2	3	4	5	6
Type of the cluster	1	1	2	1	1	2
Size of the cluster	172	149	95	187	97	68
Number of misidentified points	70	11	18	54	8	20
Percentage of misidentified points	41%	7%	19%	29%	8%	29%

identified points), while the 1st cluster contains 41% of misidentified points, though the points of the 1st type dominate. In principle, this additional information can be used for stating the diagnosis: that is, if a new point belongs to the 2nd cluster, then it is very probable, that it is of the 1st type. At the same time, if a new point is in the 1st cluster, we are not so sure about it.

The results obtained are comparable with those reported in (Astorino and Gaudioso, 2002; Bazirov, Rubinov and Yearwood, 2002; Bagirov, Rubinov and Yearwood, 2001).

8. Concluding Remarks

Thus, we have derived necessary and sufficient conditions for a point to be a minimizer of the sum-min function. Based on these conditions we became able to construct some numerical methods for solving min-sum-min problems.

To check practical efficiency of the proposed algorithms, we applied these methods to the problems of cluster analysis.

On several datasets, containing up to 1.000 points (Diabetes and Heart disease datasets) they demonstrated quite a competitive performance, and we believe that these ideas can be employed for solving different large-scale problems. That is, the problems where the function under consideration involves summation over several hundreds or thousands min-functions.

Acknowledgements

The author is thankful to the anonymous Referees for their careful reading the manuscript and useful suggestions made.

References

1. Astorino, A. and Gaudioso, M. (2002), Polyhedral separability through successive LP, *Journal of Optimization Theory and Applications*, 112(2), 265–293.
2. Bagirov, A.M., Rubinov, A.M., Stranieri, A. and Yearwood, J. (1999), The global optimization approach to the clustering analysis. Research Report 45/99, University of Ballarat, Australia.
3. Bagirov, A.M., Rubinov, A.M. and Yearwood, J. (2002), A global optimization approach to the classification, *Optimization and Engineering*, 3, 129–155.
4. Bagirov, A.M., Rubinov, A.M. and Yearwood, J. (2001), Using global optimization to improve classification for medical diagnosis, *Topics in Health Information Management*, 22, 65–74.

5. Demyanov, A.V. (2003), An exchange algorithm for minimizing sum-min functions. In: Qi, L., Teo, K.L. and Yang, X.Q. (eds.), *Optimization and Control with Applications*, Kluwer Academic Publisher, Nowell, MA. USA, 2003, pp. 213–234.
6. Demyanov, A.V. (2002), On the solution of minmaxmin problems, *Vestnik of Saint-Petersburg University*, Ser. 1, Issue 3 (No. 17), 3–9.
7. Demyanov, A.V., Demyanov, V.F. and Malozemov, V.N. (2002), Minmaxmin problems revisited, *Optimization Methods and Software*, 2002, 17(5), 783–784.
8. Mangasarian, O.L. (1997), Mathematical programming in data mining, *Data Mining and Knowledge Discovery*, 1, 183–201.
9. Pardalos, P. and Resende, M. (eds.), (2002), *Handbook of Applied Optimization*, Oxford University Press.
10. Vershik, A.M., Malozemov, V.N. and Pevnyi, A.B. (1975), The best piecewise polynomial approximation, *Siberian Mathematical Journal*, XVI(5), 925–938.