# Combinatorial Optimization Problems Related to Machine Learning Techniques

**Conference Paper** · December 2014



2 authors:

Michael Khachay
Russian Academy of Sciences

Maria Poberiy
Russian Academy of Sciences

# Combinatorial Optimization Problems Related to Machine Learning Techniques

Michael Khachay
Krasovsky Institute of
Mathematics and Mechanics
16 S. Kovalevskoy str., 620990 Ekaterinburg
Omsk State Technical University
11 Mira ave., 644050 Omsk, Russia
Email: mkhachay@imm.uran.ru

Maria Poberii
Krasovsky Institute of
Mathematics and Mechanics
16 S. Kovalevskoy str., 620990 Ekaterinburg
Omsk State Technical University
11 Mira ave., 644050 Omsk, Russia
Email: maschas_briefen@mail.ru

***Abstract*—A brief survey of computational complexity and approximability results concerning efficient cluster analysis techniques and learning procedures in the class of piece-wise linear majority classifiers is provided. Also, new results confirming the connection between the structural minimization risk principle and theoretic combinatorics (along with rigorous proofs) are presented.**

## I. Introduction

Optimization and machine learning appear to be extremely close fields of the modern computer science. Various areas in machine learning: SVM-learning and kernel machines (see, e.g. [1]), PAC-learning [2] and boosting [3], [4], cluster analysis, etc. are continuously presenting new challenges for designers of optimization methods due to the steadily increasing demands on accuracy, efficiency, space and time complexity and so on.

Sometimes a learning problem can be successfully reduced to some kind of combinatorial optimization problems: max-cut, p-median, TSP, etc. To this end, all the known results for the latter problem (approximation algorithms, polynomial-time approximation schemas, approximation thresholds) can find their application in design precise and efficient learning algorithms for the former.

In this paper we try to observe some new results proving such a mutual cooperation between CO and ML.

This is chiefly a survey paper. In Section II we overview several recent results obtained for two special cases of well-known Weighted Clique Problem and their application in cluster analysis. Further, in Section III, some aspects of ensemble learning are covered. In particular, in Subsection III-A we give a brief survey on combinatorial results concerning the Minimum Affine Separating Committee (MASC) problem. Finally, in Subsection III-B we present new results bridging well-known Structural risk minimization principle and the Integer Partition Problem (IPP), which is an object of interest in combinatorics and number theory.

## II. Cluster Analysis and the Weighted Clique Problem

We start with giving an informal description of the special class of cluster analysis problems. Suppose, we deals with a collection of uniform objects. Any object can be in two states, we call them *on-line* and *off-line*. During the experiment, for any object in question, its current state can be revealed up to some additive noise. The input dataset contains such results. The problem is to distinguish the given number of active objects.

It is convenient, to formulate this problem mathematically in terms of graph theory. Let a complete, weighted, undirected graph $G = (V, E, w)$ be given. Here, $w$ is a weighting (cost) non-negative function $w : E \to \mathbb{R}_+$ defined on the edge-set of the graph $G$. For any subgraph $G' = (V', E') \subset G$, the weight $w(G')$ is defined by the equation $w(G') = \sum_{e\in E'} w(e)$. As usual, we call any complete subgraph $G'$ of the graph $G$ *a clique*.

*Problem 1 (min-EWCP):* Input: a graph $G$ of order $n$ and a natural number $m < n$. It is required to find a clique of order $m$ having the minimum possible weight.

As it is proved in [5], the min-EWCP is intractable and inapproximable in the general case.

*Theorem 1:* (i) The Min-EWCP is strongly NP-hard. (ii) For any ratio $r = O(2^n)$, the Min-EWCP has no polynomial time r-approximation algorithms unless $P = NP$.

Following [5], we describe two special cases of the Min-EWCP problem: we call them *metric* (Min-EWCP-M) and *square Euclidean* (Min-EWCP-SE):

(i) in the metric subclass, the weighting function $w$ is defined by symmetric zero-main-diagonal $n\times n$-matrix $W = \|w_{ij}\|$ satisfying the triangle inequality $w_{ij} + w_{jk} \geq w_{ik}$;

(ii) in the square Euclidean case, vertices of the graph $G$ are points in some finite-dimensional Euclidean space and entries of the weight matrix $W$ is obtained as squared Euclidean among them.

As the general min-EWCP, the defined above its special cases are intractable as well.

*Theorem 2:* The min-EWCP-M and min-EWCP-SE problems are strongly NP-hard.

Fortunately, the both problems belong to Apx class. In [5], for each of these problems, a polynomial-time 2-approximation algorithm is developed. For brevity, we skip their formal description. Theorem 3 summarizes their main properties.

*Theorem 3:* (i) for min-EWCP-M problem there exist a polynomial time 2-approximation algorithm with running time of $O(n^2)$ and asymptotically tight approximation ratio; (ii) for min-EWCP-SE problem there exist a polynomial time 2-approximation algorithm with running time of $O(n^2)$ and tight approximation ratio;

## III. Committee Separability Related Problems

### A. Minimum Affine Separating Committee (MASC) Problem

We start with the simplest two-class setting of the classic pattern recognition problem. For a finite sample

$$\xi = ((x_1, y_1), \ldots, (x_m, y_m)), \tag{1}$$

where $x_i \in \mathbb{R}^d$ and $y_i \in \{-1, 1\}$, for a given class

$$\mathcal{H} \subset [\mathbb{R}^d \to \{-1, 1\}]$$

of two-valued functions, which are called *classifiers*, and a given utility functional $I : H \to \mathbb{R}_+$, it is required to find an optimal (or suboptimal) classifier $h = h_\xi$ w.r.t. the functional $I$. Usually, the functional $I$ has a form of expected misclassification loss

$$I[h] = \int\limits_{X \times Y} L(y - h(x))dP(x, y)$$

depending on unknown probability measure $P$. Due to this uncertainty, the initial problem

$$\min_{h \in H} I[h] \tag{2}$$

could not be solved as is and should be refined. According to the famous Empirical Risk Minimization (ERM) principle, the unformalized functional $I$ is replaced with the empirical mean functional (of the misclassification loss)

$$I_{emp}[h] = \frac{1}{m} \sum_{j=1}^{m} L(y_j - h(x_j)), \tag{3}$$

which is completely defined by training sample (1). A classifier $h_0$ making no classification errors on sample (1) is called *perfect* w.r.t. this sample.

The Minimum Affine Separating Committee (MASC) combinatorial optimization problem is equivalent to the minimization problem of the functional $I_{emp}$ over the special class $\mathcal{H}_{asc}$ of piece-wise linear classifiers

$$h(x) = \operatorname{sign} \sum_{i=1}^{k} a_i \operatorname{sign}(w_i^T x - b_i) \tag{4}$$

for some positive integers $a_i$, where decisions of *weak* affine classifiers $\operatorname{sign}(w_i^T x - b_i)$ are aggregated by the simple majority voting rule. The value $n = \sum_{i=1}^{k} a_i$ is called *a length* of the classifier $h$. Any perfect committee classifier w.r.t. sample (1) of minimum length is called *minimum affine separating committee*.

Actually, the MASC problem appears to be one of mathematical formalizations of the well-known Vapnik-Chervonenkis structural risk minimization principle [6], which is of searching of the most precise classifier in the most narrow subfamily of feasible ones.

We use the traditional mathematical notation $\mathbb{N}$ and $\mathbb{R}$ for the sets of natural and real numbers, and $\mathbb{N}_k$ for the set $\{1, \ldots, k\}$. For convenience, we introduce subsets $A, B \subset \mathbb{R}^d$ consisting of $x_i$ from sample (1) and defined by the equations

$$A = \{x_i \colon y_i = 1\}, \quad B = \{x_i \colon y_i = -1\}. \tag{5}$$

In [7], it is proved the following criterion of perfect learnability in the class of affine committee classifiers.

*Theorem 4:* Class $\mathcal{H}_{asc}$ contains a perfect classifier if and only if

$$A \cap B = \varnothing. \tag{6}$$

Hereinafter, we assume training sample (1) to be *regular*, i.e.

(i) satisfying condition (6);

(ii) the set $A \cup B$ is in *general position* (see Definition 1)

*Definition 1:* A set $D \subset \mathbb{R}^d$, $|D| > d$, is said to be in general position, if, for any subset $D' \subseteq D$, $|D'| = n + 1$, the equality $\dim \operatorname{aff} D' = d$ is valid.

*Problem 2 (MASC-GP):* For a given sample (1) it is required to construct a minimum affine separating committee.

To emphasize the important special case of the MASC-GP problem, in which the dimensionality $d$ is fixed in advance, we call this problem MASC-GP(d). Complexity of the both problems is described in Theorem 5 [8] and 6 [9].

*Theorem 5:* The MASC-GP problem is strongly NP-hard.

*Theorem 6:* The MASC-GP($d$) problem is polynomially solvable for $d = 1$ and NP-hard for any fixed $d > 1$.

The state-of-the-art *Boosting Greedy Committee (BGC)* approximation algorithm for the MASC-GP(d) problem is proposed in [10]. This algorithm has the best knonwn approximation ration and a rather huge but polynomial complexity bound.For brevity, we skip its formal description but recall main properties in Theorem .

In sequel, we call an instance of the MASC-GP($d$) problem *nice* if there exists a minimum affine separating committee (4) of odd length $n$ such that, for any $t = 1, \ldots, (n-1)/2$, the following conditions

$$(w_t^T x - b_t > 0) \quad \vee \quad (w_{t+1}^T x - b_{t+1} > 0), \quad (x \in A),$$
$$(w_t^T x - b_t < 0) \quad \vee \quad (w_{t+1}^T x - b_{t+1} < 0), \quad (x \in B),$$

are valid.

*Theorem 7:* The BGC algorithm has the approximation ratio of $O(\ln(m))$ for the nice instances of the MASC-GP($d$) problem and

$$O\left(\left(\frac{m \ln m}{d}\right)^{1/2}\right),$$

otherwize. Its time-complexity is of $O(m^{3d})$ for any $d > 2$.

## B. Committee Minimal Partitions

The Integer Partition Problem (IPP) is one of the fundamental problems, which is studied in combinatorics and number theory. In this problem, for a given natural number $n$ it is required to enumerate all of ways to represent of this number us a sum of other non-negative integer numbers $n = a_1 + \ldots + a_k$. The finite sequence $(a_1, \ldots, a_k)$ is called *integer partition* of the number $n$. For brevity, the fact 'the sequence $A$ is a partition of the number $n$' is denoted by $A \vdash n$.

We consider the restricted version of the IPP; we call this problem Minimum Committee Integer Partition Problem (MC-IPP) which is closely related to the minimum committee notion considered above. In MC-IPP, for a given training sample $\xi$, a given class $\mathcal{F}$ of weak classifiers, and for a given number $n$, it is required to enumerate only such partitions $(a_1, \ldots, a_k)$, for which there exists a perfect (w.r.t. the sample $\xi$) minimum committee classifier $h$ such that

$$h(x) = \operatorname{sign} \sum_{i=1}^{k} a_i f_i(x), \quad (f_i \in \mathcal{F}). \tag{7}$$

Hereinafter, to emphasize that the classifier $h$ is defined by the partition $(a_1, \ldots, a_k)$ and the weak classifiers $f_1, \ldots, f_k$, we use the notation $h(\cdot | a_i, f_i)$.

We introduce the following partial order over the set of all partitions of the length $k$.

*Definition 2:* Suppose $A = (a_1, \ldots, a_k)$ and $B = (b_1, \ldots, b_k)$ be finite sequences of non-negative integers. The relation $A \succcurlyeq B$ is defined by the following equations

$$\sum_{i=1}^{k} a_i = n \quad \geq \quad m = \sum_{i=1}^{k} b_i, \tag{8}$$

$$(J \subseteq \mathbb{N}_k) \left( \sum_{i \in J} a_i > \frac{n}{2} \right) \iff \left( \sum_{i \in J} b_i > \frac{m}{2} \right) \tag{9}$$

Notice that, if $A \succcurlyeq B$ then, for any sample $\xi$ and for any weak classifiers $f_1, \ldots, f_k$, the committee classifiers $h(\cdot|a_i, f_i)$ and $h(\cdot|b_i, f_i)$ are perfect w.r.t. the sample $\xi$ simultaneously and the length of the latter committee does not exceeds the length of the former.

We call a sequence $A = (a_1, \ldots, a_k)$ *decreasing* if $a_1 \geq \ldots \geq a_k$. It is convenient to restrict our consideration to decreasing partitions only.

*Assertion 1:* Let sequences $A = (a_1, \ldots, a_k)$ and $B = (b_1, \ldots, b_k)$ be partitions of numbers $n$ and $m$, respectively, satisfying equation (9). If the partition $A$ is decreasing, then, for the number $p$ there exists a decreasing partition $C = (c_1, \ldots, c_k)$ such that, for $A$ and $C$ the equation (9) is valid, as well.

*Proof:* W.l.o.g., it is sufficient to verify that, if $b_1 < b_2$ then the sequence $B' = (b'_1, b'_2, \ldots, b'_k)$, for which $b'_1 = b_2$, $b'_2 = b_1$, and $b'_i = b_i$ for $i > 2$, is also satisfies (together with $A$) condition (9).

Choose any $J \subset \mathbb{N}_k$ such that

$$\sum_{i \in J} b'_i > \frac{m}{2}.$$

If $\{1, 2\} \cap J = \varnothing$ or $\{1, 2\} \subset J$ then

$$\sum_{i \in J} b_i = \sum_{i \in J} b'_i > \frac{m}{2};$$

therefore,

$$\sum_{i \in J} a_i > \frac{n}{2},$$

by condition. If $1 \in J$ and $2 \notin J$, then

$$\sum_{i \in J \setminus \{1\} \cup \{2\}} b_i = \sum_{i \in J} b'_i > \frac{m}{2}.$$

Therefore,

$$\sum_{i \in J \setminus \{1\} \cup \{2\}} a_i > \frac{n}{2},$$

by condition, and consequently

$$\sum_{i \in J} a_i \geq \sum_{i \in J \setminus \{1\} \cup \{2\}} a_i > \frac{n}{2},$$

as $a_1 \geq a_2$. Finally, if $1 \notin J$ and $2 \in J$, then $\sum_{i \in J} b_i > \sum_{i \in J} b_i$, since $b_2 > b'_2 = b_1$. Hence,

$$\sum_{i \in J} a_i > \frac{n}{2}.$$

Assertion 1 is proved. ■

Hereinafter, for any partition $A$, we assume that $A$ is decreasing. Assertion 1 suggests us to extend (using the zero-padding technique) the defined above partial order $A \succcurlyeq B$ onto sequences of unequal lengths.

*Definition 3:* Suppose, for a sequence $A$ and for any other sequence $B$, either $B \succcurlyeq A$ or the sequences $A$ and $B$ are incomparable. Then, the sequence $A$ is called *a committee minimal (c-minimal) sequence*.

In the sequel, we apply Definition 3 to partitions of natural numbers. Evidently, if committee (7) is perfect for some sample $\xi$, then the partition $(a_1, \ldots, a_k)$ is c-minimal. The inverse claim is also valid.

*Theorem 8:* Let $A = (a_1, \ldots, a_k)$ be a c-minimal partition of some number $n$. There exist a training sample $\xi$ and a class $\mathcal{F}$ of weak classifiers such that committee (8) (for some classifiers $f_1, \ldots, f_k$) is perfect w.r.t. the sample $\xi$.

Theorem 8 is announced in [11] and can be proved using the technique proposed in [12].

To define the MC-IPP, we introduce some additional notation. For an arbitrary odd number $n$ and natural number $s \leq n$, we denote by $\Lambda(n)$ and $\Lambda(n, s)$ the set of all partitions and the set of $s$-fold partitions of $n$, respectively.

*Problem 3 (MC-IPP):* For an odd number $n$ it is required to enumerate all the c-minimal elements of $\Lambda(n)$.

Further we propose the following efficiently verifiable c-minimality condition.

*Condition 1:* Let a partition $A = (a_1, \ldots, a_k$ be given. For any $1 \leq i_1 < i_2 \leq k$ there exists a subset $J \subset \mathbb{N}_k$ such that $\{i_1, i_2\} \subset J$ and $\sum_{i \in J} a_i = \lceil n/2 \rceil$.

*Lemma 1 (Necessary condition):* For any odd number $n$ and any c-minimal partition $A = (a_1, \dots, a_k) \vdash n$, Condition 1 is valid.

*Proof:* Assume by contradiction that there exist numbers $i_1 < i_2$ such that, for any $J \subset \mathbb{N}_k$ the condition

$$\left(\{i_1, i_2\} \subset J,\ \sum_{i\in J} a_i > \frac{n}{2}\right) \Rightarrow \left(\sum_{i\in J} a_i > \left\lceil\frac{n}{2}\right\rceil\right) \quad (10)$$

is valid. We consider the partition $B = (b_1, \dots, b_k)$ of the number $n-2$ defined by equations

$$b_{i_1} = a_{i_1} - 1, b_{i_2} = a_{i_2} - 1,\ b_i = a_i\ (i \in \mathbb{N}_k \setminus \{i_1, i_2\}) \quad (11)$$

and verify that $A \succcurlyeq B$. Indeed, let, for some subset $J \subset \mathbb{N}_k$, the equation $\sum_{i\in J} b_i > n/2 - 1$ be valid. If $\{i_1, i_2\} \cap J \neq \varnothing$, then $\sum_{i\in J} a_i > n/2$ by construction of the partition $B$.

Otherwise, we obtain

$$\sum_{i\in J} a_i > \frac{n}{2} - 1.$$

To this end, if

$$\sum_{i\in J} a_i < \frac{n}{2},$$

then

$$\sum_{i\in J} a_i = \left\lfloor\frac{n}{2}\right\rfloor;$$

therefore,

$$\sum_{i\notin J} a_i = \left\lceil\frac{n}{2}\right\rceil.$$

This equation contradicts to our assumption (10), since $\{i_1, i_2\} \subset \mathbb{N}_k \setminus J$. Lemma is proved. ■

Further, we prove that, to verify whether a given partition $A = (a_1, \dots, a_k)$ is c-minimal, it is sufficient to examine partitions $B = (b_1 \dots, b_l)$, for which $l \leq k$.

*Lemma 2:* (i) Suppose, $A = (a_1, \dots, a_k)$, $a_k > 0$, is not a c-minimal partition of an odd number $n$. For some number $m < n$, there exists a partition $B = (b_1, \dots, b_l) \vdash m$ such that $A \succcurlyeq B$ and $l \leq k$. (ii) If, in addition, for the partition $A$, Condition 1 is valid, then $l = k$.

*Proof:* Indeed, since the partition $A$ is not c-minimal, there exists a partition $C = (c_1, \dots, c_l) \vdash m$ such that $c_l > 0$ and $A \succcurlyeq C$. Assume, $l > k$; and introduce the following notation

$$d = \sum_{i=k+1}^{l} c_i. \quad (12)$$

Let $d$ be an even number. We consider the partition $B = (c_1, \dots, c_k) \vdash (m-d)$ and show that $A \succcurlyeq B$. For an arbitrary subset $J \subset \mathbb{N}_k$ such that

$$\sum_{i\in J} c_i > \frac{m-d}{2} > \frac{m}{2} - d,$$

we obtain

$$\sum_{i\in J\cup\{k+1,l\}} c_i > \frac{m}{2};$$

therefore,

$$\sum_{i\in J} a_i = \sum_{i\in J} a_i + \sum_{i=k+1}^{l} 0 > \frac{n}{2},$$

by choice of the partition $C$.

On the other hand, let $d$ be an odd number. To this end, we consider the partition $B = (b_1, b_2, \dots, b_k) \vdash (n-d-1)$ defined by the equations

$$b_k = c_k - 1,\ b_i = c_i\ (i \in \mathbb{N}_{k-1}).$$

As above, we define $d$ by (12). For any $J \subset \mathbb{N}_k$ such that

$$\sum_{i\in J} b_i > \frac{m-d-1}{2} > \frac{m}{2} - d,$$

we obtain

$$\sum_{i\in J\cup\{k+1,l\}} c_i > \frac{m}{2}.$$

Consequently,

$$\sum_{i\in J} a_i > \frac{n}{2}.$$

Claim (i) is proved.

To prove claim (ii), assume that, for the partition $A$, the necessary condition of c-minimality (proven in Lemma 1) is valid. In this case, we prove that, for any partition $B = (b_1, \dots, b_l)$,

$$(A \succcurlyeq B,\ b_l > 0) \Rightarrow (l \geq k).$$

Assume, by contradiction that there is a partition $B = (b_1, \dots, b_l) \vdash m$, for which $b_l > 0$ and $l < k$. Since the partition $A$ satisfies the necessary condition, there exists a subset $J \subset \mathbb{N}_k$ such that $\{1, k\} \subset J$ and

$$\sum_{i\in J} a_i = \left\lceil\frac{n}{2}\right\rceil.$$

As $A \succcurlyeq B$,

$$\sum_{i\in \mathbb{N}_l \cap J} b_i > \frac{m}{2}.$$

On the other hand,

$$\sum_{i\in \mathbb{N}_k\setminus J} a_i + a_k > \sum_{i\in \mathbb{N}_s\setminus J} a_i = \left\lfloor\frac{n}{2}\right\rfloor.$$

Therefore, $\sum_{i\in\mathbb{N}_l\setminus J} b_i > m/2$, i.e. $\sum_{i=1}^{l} b_i > m$ that contradicts to the choice of the partition $B \vdash m$. Claim (ii) and lemma are proved. ■

Further, we prove several assertions providing an approach to recurrent construction of c-minimal partitions.

*Assertion 2:* Let $A = (a_1, \dots, a_k)$ be a c-minimal partition of some odd number $n$. The partition

$$B = (a_1, \dots, a_k, 1, 1) \vdash (n+2) \quad (13)$$

is c-minimal as well.

*Proof:* As $A$ is a c-minimal partition, for this partition Lemma 1 is valid. It is easy to verify, that the same necessary condition is valid for the partition $B$ as well. Assume, by contradiction that there exists a partition $C = (c_1, \dots, c_l) \vdash m$

such that $B \not\succcurlyeq C$ and $m \leq n$. Due to claim (ii) of Lemma 2, $l = k+2$ and $c_{k+2} > 0$. By choice of the partition $C$, for any $J \subset \mathbb{N}_k$, the equation

$$\sum_{i \in J} c_i + c_{k+1} > \frac{m}{2}$$

implies

$$\sum_{i \in J} a_i + 1 > \frac{n+2}{2}, \text{ i.e. } \sum_{i \in J} a_i > \frac{n}{2}.$$

Let $c_{k+1} + c_{k+2} \equiv 0 \pmod 2$. Consider the partition

$$D = (c_1, \dots, c_k) \vdash m - (c_{k+1} + c_{k+2}).$$

Assume that, for some $J \subset \mathbb{N}_k$

$$\sum_{i \in J} c_i > \frac{m - c_{k+1} - c_{k+2}}{2}.$$

Then,

$$\sum_{i \in J} c_i > \frac{m}{2} - c_{k+1},$$

since the partition $C$ is decreasing. Therefore,

$$\sum_{i \in J} c_i + c_{k+1} > \frac{m}{2};$$

hence,

$$\sum_{i \in J} a_i + 1 > \frac{n+2}{2}, \text{ and } \sum_{i \in J} a_i > \frac{n}{2}.$$

Thus, $A \not\succcurlyeq D$ and the partition $A$ is not c-minimal, since $m - (c_{k+1} + c_{k+2}) < n$.

In the case $c_{k+1} + c_{k+2} \equiv 1 \pmod 2$, we obtain $c_k > 1$ by construction of the partition $C$. Consider the partition

$$D = (c_1, \dots, c_k - 1) \vdash m - (c_{k+1} + c_{k+2} + 1).$$

As $c_{k+1} > c_{k+2}$, then

$$\frac{c_{k+1} + c_{k+2} + 1}{2} \leq c_{k+1}.$$

It is easy to show that, for any $J \subset \mathbb{N}_k$, the condition

$$\sum_{i \in J} c_i > \frac{m - (c_{k+1} + c_{k+2} + 1)}{2}$$

implies

$$\sum_{i \in J} a_i > \frac{n}{2}.$$

Therefore, again, $A \not\succcurlyeq D$ and the partition $A$ is not c-minimal, since $m - (c_{k+1} + c_{k+2} + 1) < n$. The obtained contradiction completes the proof. ∎

*Assertion 3 (Sufficient condition):* Let $A = (a_1, \dots, a_k) \vdash n$, $a_k > 0$, be a c-minimal partition for a given odd number $n$. For some $i \in \mathbb{N}_k$, let the partition $B = (b_1, \dots, b_k, b_{k+1})$ be defined by the equation

$$b_i = a_i + 1, \; b_j = a_j \; (j \in \mathbb{N}_k \setminus \{i\}), b_{k+1} = 1. \tag{14}$$

If Condition 1 is valid for the partition $B$, then this partition is c-minimal.

*Proof:* Assume, by contradiction that there exists a partition $C = (c_1 \dots, c_{k+1} \vdash m)$ such that $B \not\succcurlyeq C$ and $m \leq n$. Define the partition $D = (d_1, \dots, d_k) \vdash m - 2c_{k+1}$ by the formulas

$$d_i = c_i - c_{k+1}, \; d_j = c_j \; (j \in \mathbb{N}_k \setminus \{i\})$$

Suppose, for some subset $J \subset \mathbb{N}_k$,

$$\sum_{j \in J} d_i > \frac{m}{2} - c_{k+1}.$$

If $i \in J$, then

$$\sum_{j \in J} c_j = \sum_{j \in J} d_j + c_{k+1} > \frac{m}{2}.$$

Therefore,

$$\sum_{j \in J} b_j + 1 = \sum_{j \in J} a_j + 2 > \frac{n+2}{2}.$$

Hence, as $a_j$ are integers and $n$ is an odd number,

$$\sum_{j \in J} a_j > \frac{n}{2}. \tag{15}$$

Else, if $i \notin J$, then

$$\sum_{j \in J \cup \{k+1\}} c_j = \sum_{j \in J} d_j + c_{k+1} > \frac{m}{2};$$

therefore

$$\sum_{j \in J \cup \{k+1\}} b_j = \sum_{j \in J} a_j + 1 > \frac{n}{2} + 1.$$

Thus, equation (15) is valid again.

Since $m - 2c_{k+1} < n$ and $A \not\succcurlyeq D$, then $A$ is not c-minimal partition. The obtained contradiction completes the proof. ∎

Assertions 2 and 3 provides an efficient algorithm for recurrent construction of c-minimal partitions. It can be easily verified that, if a c-minimal partition $B = (b_1, \dots, b_l)$ is constructed (from another c-minimal partition $A$) by equations (13) or (14), then the partitions $C = (b_1, \dots, b_l, 1, 1)$ and

$$D = (b_1, b_2, \dots, b_i + 1, \dots, b_l, 1)$$

are c-minimal as well for any $i \in \mathbb{N}_l$.

## IV. Conclusion

In the paper connections between several combinatorial optimization problems and machine learning techniques are shown. In particular,

(i) approximability of metric and square Euclidean Minimum Weighted Clique Problem closely related to Minimum Sum of Squares clustering method are observed;

(ii) computational complexity and approximability issues of Minimum Affine Separating Committee Problem related to optimal learning procedures in the class of piece-wise linear majority classifiers are listed;

(iii) new results establishing the connection between classic number theory and combinatorics and structural risk minimization machine learning principle are presented.

## Acknowledgment

This research was supported by Russian Science Foundation grant no. 14-11-00085.


## References


[1] B. Schölkopf and A. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2002.

[2] L. G. Valiant, "A theory of the learnable," *Communications of the ACM*, vol. 27, no. 11, pp. 1134–1142, 1984.

[3] R. Schapire and Y. Freund, *Boosting: Foundations and algorithms*. MIT Press, 2012.

[4] Y. Freund, "Boosting a weak learning algorithm by majority," *Information and Computation*, vol. 121, pp. 256–285, 1995. [Online]. Available: http://libgen.org/scimag/index.php?doi=10.1006/inco.1995.1136

[5] I. Eremin, E. Gimadi, A. Kelmanov, A. Pyatkin, and M. Y. Khachai, "2-approximation algorithm for finding a clique with minimum weight of vertices and edges," *Proceedings of the Steklov Institute of Mathematics*, vol. 284, no. 1, pp. 87–95, 2014.

[6] V. Vapnik, *Statistical Learning Theory*. Wiley, 1998.

[7] V. Mazurov, "Committees of inequalities systems and the pattern recognition problem," *Kibernetika*, vol. 3, pp. 140–146, 1971.

[8] M. Khachai, "Computational and approximational complexity of combinatorial problems related to the committee polyhedral separability of finite sets," *Pattern Recognition and Image Analysis*, vol. 18, no. 2, pp. 236–242, 2008.

[9] M. Khachay and M. Poberii, "Complexity and approximability of committee polyhedral separability of sets in general position," *Informatica*, vol. 20, no. 2, pp. 217–234, 2009.

[10] M. Khachai and M. Poberii, "Scheme of boosting in the problems of combinatorial optimization induced by the collective training algorithms," *Automation and Remote Control*, vol. 75, no. 4, pp. 657–667, 2014.

[11] M. Khachai, "On one combinatorial problem concerned with the notion of minimal committee," *Pattern Recognition and Image Analysis*, vol. 11, no. 1, pp. 45–46, 2001.

[12] M. Khachay, "Estimate of the number of members in the minimal committee of a system of linear inequalities," *Computational Mathematics and Mathematical Physics*, vol. 37, no. 11, pp. 1356–1361, 1997.