Comput Optim Appl
DOI 10.1007/s10589-008-9202-9

Concave programming for minimizing the zero-norm over polyhedral sets

F. Rinaldi · F. Schoen · M. Sciandrone

Received: 20 March 2008 / Revised: 18 August 2008

Abstract Given a non empty polyhedral set, we consider the problem of finding a vector belonging to it and having the minimum number of nonzero components, i.e., a feasible vector with minimum zero-norm. This combinatorial optimization problem is NP-Hard and arises in various fields such as machine learning, pattern recognition, signal processing. One of the contributions of this paper is to propose two new smooth approximations of the zero-norm function, where the approximating functions are separable and concave. In this paper we first formally prove the equivalence between the approximating problems and the original nonsmooth problem. To this aim, we preliminarily state in a general setting theoretical conditions sufficient to guarantee the equivalence between pairs of problems. Moreover we also define an effective and efficient version of the Frank-Wolfe algorithm for the minimization of concave separable functions over polyhedral sets in which variables which are null at an iteration are eliminated for all the following ones, with significant savings in computational time, and we prove the global convergence of the method. Finally, we report the numerical results on test problems showing both the usefulness of the new concave formulations and the efficiency in terms of computational time of the implemented minimization algorithm.

Keywords Zero-norm · Concave optimization · Frank-Wolfe algorithm · Feature selection

F. Rinaldi
Istituto di Analisi dei Sistemi ed Informatica, Consiglio Nazionale delle Ricerche, Viale Manzoni 30, 00185 Roma, Italy
e-mail: francesco.rinaldi@iasi.cnr.it

F. Schoen (✉) · M. Sciandrone
Dipartimento di Sistemi e Informatica, Via di S. Marta 3, 50139 Firenze, Italy
e-mail: fabio.schoen@unifi.it

M. Sciandrone
e-mail: sciandro@dsi.unifi.it

1 Introduction

Given a polyhedral set, we consider the problem of finding a vector belonging to it and having the minimum number of nonzero components. Formally, the problem is

$$\begin{aligned} \min_{x \in R^n} \quad & \|x\|_0 \\ & x \in P \end{aligned} \tag{1}$$

where $\|x\|_0$ is the zero-norm of x defined as $\|x\|_0 = \text{card}\{x_i : x_i \neq 0\}$, $P \subset R^n$ is a non empty polyhedral set.

This combinatorial optimization problem is NP-Hard [2], and arises in various fields such as machine learning (see, e.g., [9]), pattern recognition (see, e.g., [11]), signal processing (see, e.g., [4–6]).

In order to make the problem tractable, the simplest approach can be that of replacing the zero-norm, which is a nonconvex discontinuous function, by the ℓ_1 norm thus obtaining the linear programming problem

$$\begin{aligned} \min_{x,y \in R^n} \quad & \sum_{i=1}^{n} y_i \\ & x \in P \\ & -y_i \leq x_i \leq y_i \quad i = 1, \ldots, n, \end{aligned} \tag{2}$$

which can be efficiently solved even when the dimension of the problem is very large. Under suitable assumptions on the polyhedral set P (defined by an underdetermined linear system of equations) it is possible to prove that a solution of (1) can be obtained by solving (2) (see, e.g., [8]). However, these assumptions may be not satisfied in many cases, and some experiments concerning machine learning problems and reported in [3] show that a concave optimization-based approach performs better than that based on the employment of the ℓ_1 norm.

In order to illustrate the idea underlying the concave approach, we observe that the objective function of problem (1) can be written as follows

$$\|x\|_0 = \sum_{i=1}^{n} s(|x_i|)$$

where $s : R \to R^+$ is the *step function* such that $s(t) = 1$ for $t > 0$ and $s(t) = 0$ for $t \leq 0$. The nonlinear approach experimented in [3] was originally proposed in [10], and is based on the idea of replacing the discontinuous step function by a continuously differentiable concave function $v(t) = 1 - e^{-\alpha t}$, with $\alpha > 0$, thus obtaining a problem of the form

$$\begin{aligned} \min_{x,y \in R^n} \quad & \sum_{i=1}^{n} (1 - e^{-\alpha y_i}) \\ & x \in P \\ & -y_i \leq x_i \leq y_i \quad i = 1, \ldots, n. \end{aligned} \tag{3}$$

The replacement of (1) by the smooth concave problem (3) is well-motivated (see [10]) both from a theoretical and a computational point of view:

- for sufficiently high values of the parameter α there exists a vertex solution of (3) which provides a solution of the original problem (1), and in this sense the approximating problem (3) is equivalent to the given nonsmooth problem (1);
- the Frank-Wolfe algorithm [7] with unitary stepsize is guaranteed to converge to a vertex stationary point of (3) in a finite number of iterations (this convergence result was proved for a general class of concave programming problems); thus the algorithm requires the solution of a finite sequence of linear programs for computing a stationary point of (3), and this may be quite advantageous from a computational point of view.

A similar concave optimization-based approach has been proposed in [12], where the idea is that of using the logarithm function instead of the step function, and this leads to a concave smooth problem of the form

$$\begin{aligned} &\min_{x,y\in R^n} \quad \sum_{i=1}^{n} \ln(\epsilon + y_i) \\ &x \in P \\ &-y_i \le x_i \le y_i \quad i = 1, \dots, n, \end{aligned} \tag{4}$$

with $0 < \epsilon \ll 1$. Formulation (4) is practically motivated by the fact that, due to the form of the logarithm function, it is better to increase one variable y_i while setting to zero another one rather than doing some compromise between both, and this should facilitate the computation of a sparse solution. A relation of (4) with the minimization of the zero-norm has been given in [12], and similarly to [10], the Frank-Wolfe algorithm with unitary stepsize has been applied to solve (4), and good computational results have been obtained.

In this paper, in Sect. 2 we derive new results on the equivalence, in a sense to be made more precise later, between a specific optimization problem and a parameterized family of problems. This analysis allows us to derive, within a general framework, results about two previously known families of approximations schemes for the zero-norm problem. Then, in Sect. 3 we introduce two new families of approximation problems for which, thanks to the theory developed in Sect. 2, it is possible to obtain equivalence results. In Sect. 4 after a brief review of the well known Frank-Wolfe method, we derive some new theoretical results which have an important impact on the computational efficiency of the method when applied to concave optimization over polyhedra. In particular we prove that once the algorithm sets a variable to zero, it will not change this variable any more. This result suggests the definition of a version of the method that eliminates the variables set to zero, thus allowing for a dimensionality reduction which greatly increments the speed of the procedure. We formally prove the global convergence of this modified version of the Frank-Wolfe method. Finally, in Sect. 5 we report the numerical results on test problems showing both the usefulness of the new concave formulations and the efficiency in terms of computational time of the implemented minimization algorithm.

2 Results on the equivalence between problems

In this section we state general conditions sufficient to ensure that a problem depending on a vector of parameters is equivalent to a given (unspecified) problem. Consider the problem

$$\min_{x \in T} \quad g(x) \tag{5}$$

where $g : R^n \to R$, $T \subseteq R^n$, and assume that it admits solutions. Let $G^\star$ be the set of such solutions.

Let $f(\cdot, u) : R^n \to R$ be a function depending on a vector of parameters $u \in U \subseteq R^m$. For any $u \in U$, consider the following problem

$$\min_{x \in T} \quad f(x, u) \tag{6}$$

Assumption 1 *There exists a finite set $S^\star \subset R^n$ having the property that, for any $u \in U$, a point $x(u) \in S^\star$ exists such that*

$$x(u) \in \arg\min_{x \in T} f(x, u). \tag{7}$$

Theorem 1 *Let $\{u^k\} \subset U$ be an infinite sequence such that for every $\tilde{x} \in T \setminus G^\star$ and every $x^\star \in G^\star$, for all but finitely many indices k we have:*

$$f(\tilde{x}, u^k) > f(x^\star, u^k). \tag{8}$$

Then, under Assumption 1, there exists a finite index $\bar{k}$ such that, for any $k \geq \bar{k}$, problem (6), with $u = u^k$, has a solution x^k that also solves the original problem (5).

Proof Let $x^\star \in G^\star$ be a solution of (5). In order to prove the thesis, by contradiction let us assume that there exists a subsequence $\{u^k\}_K$ such that, for all $k \in K$, denoting by x^k a point in $S^\star$ such that

$$x^k \in \arg\min_{x \in T} f(x, u^k), \tag{9}$$

we have

$$g(x^k) > g(x^\star). \tag{10}$$

Since $S^\star$ is finite, we can extract a further subsequence such that $x^k = \bar{x}$ for all $k \in K$, and hence, from (10), we can write

$$g(\bar{x}) > g(x^\star). \tag{11}$$

Thus $\bar{x} \in T \setminus G^\star$ and, as a consequence,

$$f(\bar{x}, u^k) > f(x^\star, u^k) \tag{12}$$

for all k sufficiently large. But this contradicts (9). □

Using the above theorem we can state the next proposition.

Proposition 1 *Let* $\{u^k\} \subset U$ *be an infinite sequence such that*

$$\lim_{k\to\infty} \frac{f(\tilde{x}, u^k) - f(x^\star, u^k)}{a + |f(x^\star, u^k)|} = C \cdot [g(\tilde{x}) - g(x^\star)] \quad \forall \tilde{x} \in T, x^\star \in G^\star \tag{13}$$

with $a \geq 0$ *and* $C > 0$. *Under Assumption* 1, *there exists a finite index* $\bar{k}$ *such that, for any* $k \geq \bar{k}$, *problem* (6), *with* $u = u^k$, *has a solution* x^k *that also solves the original problem* (5).

Proof If $\tilde{x} \in T \setminus G^\star$ then the right hand side in (13) is strictly positive. From this it follows that, for k large enough, also $f(\tilde{x}, u^k) - f(x^\star, u^k)$ will be strictly positive. □

As immediate consequence of Proposition 1 we have the following result.

Corollary 1 *Let* $\{u^k\} \subset U$ *be an infinite sequence such that*

$$\lim_{k\to\infty} f(x, u^k) = g(x) \quad \forall x \in T. \tag{14}$$

Under Assumption 1, *there exists a finite index* $\bar{k}$ *such that, for any* $k \geq \bar{k}$, *problem* (6), *with* $u = u^k$, *has a solution* x^k *that also solves the original problem* (5).

Under additional assumptions on the feasible set T and on the objective function $f(x, u)$ we can prove the following results.

Proposition 2 *Suppose that the feasible set* T *is a polyhedral set and that it admits a vertex. Assume that, for any* $u \in U$, *the objective function of* (6) *is concave, continuously differentiable, and bounded below on* T. *Let* $\{u^k\} \subset U$ *be an infinite sequence such that*

$$\lim_{k\to\infty} \frac{f(\tilde{x}, u^k) - f(x^\star, u^k)}{a + |f(x^\star, u^k)|} = C \cdot [g(\tilde{x}) - g(x^\star)] \quad \forall \tilde{x} \in T, x^\star \in G^\star \tag{15}$$

with $a \geq 0$ *and* $C > 0$. *There exists a finite index* $\bar{k}$ *such that, for any* $k \geq \bar{k}$, *problem* (6), *with* $u = u^k$, *has a solution* x^k *that also solves the original problem* (5).

Proof Let $S^\star$ be the set of vertices of T. Since the objective function of (6) is concave, continuously differentiable, and bounded below on T, it follows that $S^\star$ satisfies Assumption 1, and hence the thesis follows from Proposition 1. □

Corollary 2 *Suppose that the feasible set* T *is a polyhedral set and that it admits a vertex. Assume that, for any* $u \in U$, *the objective function of* (6) *is concave, continuously differentiable, and bounded below on* T. *Let* $\{u^k\} \subset U$ *be an infinite sequence such that*

$$\lim_{k\to\infty} f(x, u^k) = g(x) \quad \forall x \in T. \tag{16}$$

There exists a finite index $\bar{k}$ *such that, for any* $k \geq \bar{k}$, *problem* (6), *with* $u = u^k$, *has a solution* x^k *that also solves the original problem* (5).

3 Concave formulations equivalent to the zero-norm problem

In this section we define two concave smooth problems depending on some parameters, and we show (using the general results stated in the preceding section) that these problems, for suitable values of their parameters, are equivalent to the original nonsmooth problem (1), which for our convenience is rewritten as follows:

$$\begin{aligned} \min_{x\in R^n, y\in R^n} \quad & \|y\|_0 \\ & x \in P \\ & -y_i \le x_i \le y_i \quad i = 1, \dots, n \end{aligned} \tag{17}$$

We state the following assumption.

Assumption 2 *The polyhedral set P has at least a vertex.*

We denote by T the feasible set of (17), i.e.,

$$T = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in R^{2n} : x \in P, -y_i \le x_i \le y_i \ i = 1, \dots, n \right\}. \tag{18}$$

Assumption 2 implies that the polyhedral feasible set T has at least a vertex.

We introduce two concave formulations related to the ideas developed in [10] and [12], respectively.

Formulation 1

$$\begin{aligned} \min_{x\in R^n, y\in R^n} \quad & \sum_{i=1}^{n} (y_i + \epsilon)^p \\ & x \in P \\ & -y_i \le x_i \le y_i \quad i = 1, \dots, n \end{aligned} \tag{19}$$

with $0 < p < 1$, and $0 < \epsilon$.

We observe that:

- given p and ϵ, the objective function is concave, continuously differentiable, bounded below on the feasible set;
- $\lim_{p\to 0} \sum_{i=1}^{n} y_i^p = \|y\|_0$, so that the objective function can be view as a smooth approximation of the zero-norm.

The following proposition shows the equivalence between the approximating problem (19) and the zero-norm problem (17).

Proposition 3 *There exist values $\bar{p} > 0$, $\bar{\epsilon} > 0$, $\bar{\gamma} > 0$ such that, for any pair $(p, \epsilon)^T \in R^2_+$ and satisfying*

$$\begin{aligned} & p \le \bar{p} \\ & \epsilon \le \bar{\epsilon} \\ & \epsilon^p \le \bar{\gamma}, \end{aligned} \tag{20}$$

problem (19) *has a vertex solution* $(x(p,\epsilon), y(p,\epsilon))^T$ *which is also solution of the original problem* (17).

Proof In order to prove the thesis, assume by contradiction that there exists a sequence $\{(p^k, \epsilon^k, \gamma^k)^T\}$ converging to $(0,0,0)^T$, with

$$(\epsilon^k)^{p^k} \le \gamma^k, \tag{21}$$

and such that, any vertex solution of (19), with $p = p^k$ and $\epsilon = \epsilon^k$, is not a solution of (17).

Set $z = (x, y)^T$, $u = (p, \epsilon)^T$, $g(z) = \|y\|_0$, $f(z, u) = \sum_{i=1}^n (y_i + \epsilon)^p$. Problems (17) and (19) can be written as follows

$$\min_{z \in T} \quad g(z) \tag{22}$$

$$\min_{z \in T} \quad f(z, u) \tag{23}$$

where T is defined in (18). From (21), as $\gamma^k \to 0$, we can write

$$\lim_{k \to \infty} (\epsilon^k)^{p^k} = 0. \tag{24}$$

Let $\{u^k\} = \{(p^k, \epsilon^k)^T\}$ be the sequence convergent to $(0,0)^T$ and satisfying condition (24). Since for any $y \in R^+$ we have

$$\lim_{k \to \infty} (y_i + \epsilon^k)^{p^k} = \begin{cases} 1 & \text{if } y_i > 0, \\ 0 & \text{if } y_i = 0 \end{cases}$$

we obtain

$$\lim_{k \to \infty} f(z, u^k) = g(z) \quad \forall z \in T. \tag{25}$$

For any $u \in U$ the objective function of (23) is concave, continuously differentiable, and bounded below on T, so that, recalling (25), the assumptions of Corollary 2 hold and hence, for any k sufficiently large there exists a vertex solution $(x^k, y^k)^T$ which is also a solution of (22), in contradiction with our initial assumption. □

Formulation 2

$$\begin{aligned} \min_{x \in R^n, y \in R^n} \quad & -\sum_{i=1}^n (y_i + \epsilon)^{-p} \\ & x \in P \\ & -y_i \le x_i \le y_i \quad i = 1, \dots, n \end{aligned} \tag{26}$$

with $1 \le p$, and $0 < \epsilon$.

We observe that:

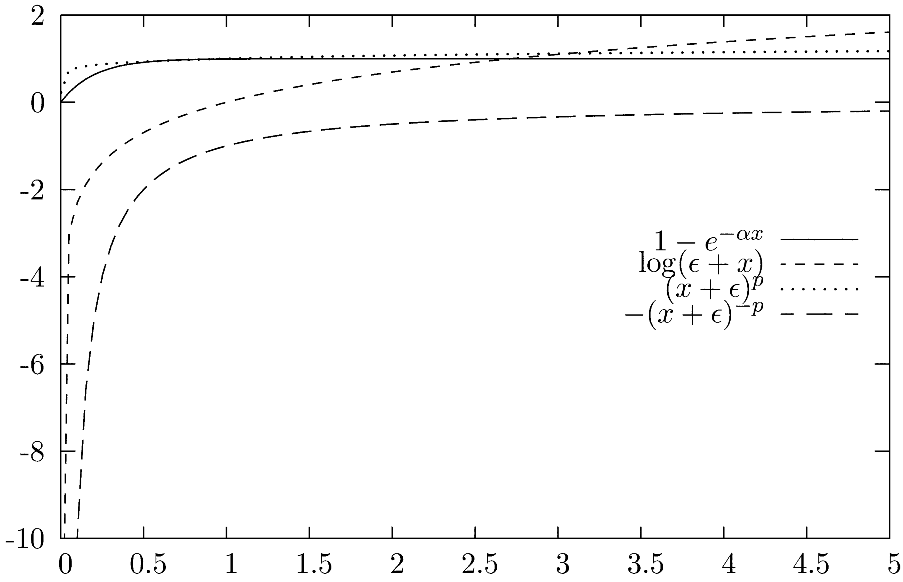

Fig. 1 Graph of functions (3) with $\alpha = 5$, (4) with $\varepsilon = 10^{-9}$, (19) with $\varepsilon = 10^{-9}$, $p = 0.1$, (26) with $\varepsilon = 10^{-9}$, $p = 1$

– given p and ϵ, the objective function is concave, continuously differentiable, bounded below on the feasible set;
– similarly to the logarithm functions appearing in problem (4), the functions $-(y_i + \epsilon)^{-p}$ favor sparse vectors rather than points having many small nonzero components; indeed, when a variable is set to zero the decrease of the function is strong compared to the increase for a larger value of another variable;
– differently from the logarithm functions of problem (4), the functions $-(y_i + \epsilon)^{-p}$ are bounded above for positive values of the independent variables, and this may be a useful additional feature for finding sparse solutions.

For easier reference, in Fig. 1 we report the graphs of the four concave functions analyzed in this paper.

The equivalence between problem (26) and the original problem (17) is formally proved below.

Proposition 4 *Assume that problem* (17) *admits a solution* $y^\star$ *such that* $\|y^\star\|_0 < n$. *There exists a value* $\bar{\epsilon} > 0$ *such that, for any* $\epsilon \in (0, \bar{\epsilon}]$, *problem* (26) *has a vertex solution* $(x(\epsilon), y(\epsilon))^T$ *which is also solution of the original problem* (17).

Proof In order to prove the thesis, assume by contradiction that there exists a sequence $\{\epsilon^k\}$ converging to zero and such that, any vertex solution of (26), with $\epsilon = \epsilon^k$, is not a solution of (17).

Set $z=(x,y)^T$, $u=\epsilon$, $g(z)=\|y\|_0$, $f(z,u)=-\sum_{i=1}^n (y_i+u)^{-p}$. Problems (17) and (26) can be written as follows

$$\min_{z\in T} \quad g(z) \tag{27}$$

$$\min_{z\in T} \quad f(z,u) \tag{28}$$

where T is defined in (18). Let $\{u^k\}=\{\epsilon^k\}$ be the sequence convergent to 0. For any $z\in T$ we have

$$f(z,u)=-\sum_{i:y_i=0} u^{-p}-\sum_{i:y_i\neq 0}(y_i+u)^{-p}=-(n-\|y\|_0)u^{-p}-\sum_{i:y_i\neq 0}(y_i+u)^{-p},$$

so that, recalling that $u^k\to 0$ for $k\to\infty$, we can write for each $\tilde{z}\in T$ and for each $z^\star\in G^\star$ (being $G^\star$ the set of optimal solutions for problem (17))

$$\begin{aligned}
&\lim_{k\to\infty}\frac{f(\tilde{z},u^k)-f(z^\star,u^k)}{|f(z^\star,u^k)|}\\
&\quad=\lim_{k\to\infty}\frac{-(n-\|\tilde{y}\|_0)(u^k)^{-p}-\sum_{i:\tilde{y}_i\neq 0}(\tilde{y}_i+u^k)^{-p}+(n-\|y_i^\star\|_0)(u^k)^{-p}+\sum_{i:y_i^\star\neq 0}(y_i^\star+u^k)^{-p}}{|-(n-\|y^\star\|_0)(u^k)^{-p}-\sum_{i:y_i^\star\neq 0}(y_i^\star+u^k)^{-p}|}\\
&\quad=\frac{\|\tilde{y}\|_0-\|y^\star\|_0}{n-\|y^\star\|_0}=C\cdot[g(\tilde{z})-g(z^\star)].
\end{aligned} \tag{29}$$

For any $u\in U=R_+$ the objective function of (28) is concave, continuously differentiable, and bounded below on T, so that, recalling (29), the assumptions of Proposition 2 hold (by setting a equal to zero) and hence, for any k sufficiently large there exists a vertex solution $(x^k,y^k)^T$ which is also a solution of (27), in contradiction with our initial assumption. □

We terminate the section by showing that the general results of Sect. 2 allow us to prove the equivalence between the smooth concave problems (3) and (4) and the given nonsmooth problem (17). We remark that the theoretical equivalence between (3) and (17) was proved in [10], while the equivalence between (4) and (17) was not formally proved.

Proposition 5 *There exists a value* $\bar{\alpha}>0$ *such that, for any* $\alpha\geq\bar{\alpha}$, *problem* (3) *has a vertex solution* $(x(\alpha),y(\alpha))^T$ *which is also solution of the original problem* (17).

Proof In order to prove the thesis, assume by contradiction that there exists a sequence $\{\alpha^k\}$ such that $\alpha^k\to\infty$, and any vertex solution of (3), with $\alpha=\alpha^k$, is not a solution of (17).

Set $z=(x,y)^T$, $u=\alpha$, $g(z)=\|y\|_0$, $f(z,u)=\sum_{i=1}^n(1-e^{-uy_i})$ and consider the problems

$$\min_{z\in T} \quad g(z) \tag{30}$$

$$\min_{z \in T} \quad f(z, u) \tag{31}$$

where T is defined in (18). Let $\{u^k\} = \{\alpha^k\}$ be the sequence convergent to $+\infty$. Since for any $y \in R^+$ we have

$$\lim_{k\to\infty} \left(1 - e^{-u^k y_i}\right) = \begin{cases} 1 & \text{if } y_i > 0 \\ 0 & \text{if } y_i = 0 \end{cases}$$

we obtain

$$\lim_{k\to\infty} f(z, u^k) = g(z) \quad \forall z \in T. \tag{32}$$

For any $u \in U = R_+$ the objective function of (31) is concave, continuously differentiable, and bounded below on T, so that, recalling (32), the assumptions of Corollary 2 hold and hence, for any k sufficiently large there exists a vertex solution $(x^k, y^k)^T$ which is also a solution of (30), in contradiction with our initial assumption. □

Proposition 6 *Assume that problem* (17) *admits a solution* $y^\star$ *such that* $\|y^\star\|_0 < n$. *There exists a value* $\bar{\epsilon} > 0$ *such that, for any* $\epsilon \in (0, \bar{\epsilon}]$, *problem* (4) *has a vertex solution* $(x(\epsilon), y(\epsilon))^T$ *which is also solution of the original problem* (17).

Proof In order to prove the thesis, assume by contradiction that there exists a sequence $\{\epsilon^k\}$ such that $\epsilon^k \to 0$, and any vertex solution of (4), with $\epsilon = \epsilon^k$, is not a solution of (17).

Set $z = (x, y)^T$, $u = \epsilon$, $g(z) = \|y\|_0$, $f(z, u) = \sum_{i=1}^n \log(y_i + u)$ and consider the problems

$$\min_{z \in T} \quad g(z) \tag{33}$$

$$\min_{z \in T} \quad f(z, u) \tag{34}$$

where T is defined in (18). Let $\{u^k\} = \{\epsilon^k\}$ be the sequence convergent to 0. For any $z \in T$ we have

$$f(z, u) = \sum_{i:y_i=0} \log u + \sum_{i:y_i \neq 0} \log(y_i + u) = (n - \|y\|_0) \log u + \sum_{i:y_i \neq 0} \log(y_i + u),$$

so that, recalling that $u^k \to 0$ for $k \to \infty$, we can write for each $\tilde{z} \in T$, $z^\star \in G^\star$ (being $G^\star$ the set of optimal solutions for problem (17))

$$\begin{aligned} &\lim_{k\to\infty} \frac{f(\tilde{z}, u^k) - f(z^\star, u^k)}{|f(z^\star, u^k)|} \\ &\quad = \lim_{k\to\infty} \frac{(n - \|\tilde{y}\|_0) \log u^k + \sum_{i:\tilde{y}_i \neq 0} \log(\tilde{y}_i + u^k) - (n - \|y^\star\|_0) \log u^k - \sum_{i:y_i^\star \neq 0} \log(y_i^\star + u^k)}{|(n - \|y^\star\|_0) \log u^k + \sum_{i:y_i^\star \neq 0} \log(y_i^\star + u^k)|} \\ &\quad = \frac{\|\tilde{y}\|_0 - \|y^\star\|_0}{n - \|y^\star\|_0} = C \cdot [g(\tilde{z}) - g(z^\star)] \end{aligned} \tag{35}$$

For any $u \in U = R_+$ the objective function of (34) is concave, continuously differentiable, and bounded below on T, so that, recalling (35), the assumptions of Proposition 2 hold (by setting a equal to zero) and hence for any k sufficiently large there exists a vertex solution $(x^k, y^k)^T$ which is also a solution of (33), in contradiction with our initial assumption. □

4 The Frank-Wolfe algorithm for the minimization of concave separable functions

Let us consider the problem

$$\min_{x \in P} \quad f(x) \tag{36}$$

where $P \subset R^n$ is a non empty polyhedral set, $f : R^n \to R$ is a concave, continuously differentiable function, bounded below on P.

The Frank-Wolfe algorithm with unitary stepsize can be described as follows.

Frank-Wolfe—Unitary Stepsize (**FW1**) Algorithm

1. Let $x^0 \in R^n$ be the starting point;
2. For $k = 0, 1, \dots,$
   if $x^k \notin \arg\min_{x \in P} \nabla f(x^k)^T x$ then compute a vertex solution x^{k+1} of

$$\min_{x \in P} \quad \nabla f(x^k)^T x \tag{37}$$

   else exit.

The algorithm involves only the solution of linear programming problems, and the following result, proved in [10], shows that the algorithm generates a finite sequence and that it terminates at a stationary point.

Proposition 7 *The Frank-Wolfe algorithm with unitary stepsize converges to a vertex stationary point of problem* (36) *in a finite number of iterations.*

Now consider the problem

$$\begin{aligned} \min \quad & f(x) = \sum_{j=1}^{n} f_j(x_j) \\ & x \in P \\ & x_i \geq 0, \ i \in I \subseteq \{1, \dots, n\} \end{aligned} \tag{38}$$

where $f_j : R \to R$, for $j = 1, \dots, n$ are concave, continuously differentiable functions. We assume that f is bounded below on P.

We observe that problem (38) includes as particular cases the concave programming problems presented in the preceding section.

The next proposition shows that, under suitable conditions on the concave functions f_j, the algorithm does not change a nonnegative variable once that it has been fixed to zero.

Proposition 8 *Let* $\{x^0, x^1, \ldots, x^h\}$ *be any finite sequence generated by the Frank-Wolfe algorithm with unitary stepsize. There exists a value* M *such that, if* $i \in I$ *and* $f_i^{'}(0) \geq M$*, then we have that*

$$x_i^k = 0 \quad \text{implies} \quad x_i^{k+1} = \ldots = x_i^h = 0.$$

Proof At each iteration k of the Frank-Wolfe algorithm the linear problem to be solved is

$$\begin{aligned} \min \quad & \sum_{j: x_j^k \neq 0} f_j'(x_j^k)\, x_j + \sum_{j \notin I: x_j^k = 0} f_j'(0) x_j + \sum_{j \in I: x_j^k = 0} f_j'(0) x_j \\ & x \in P \\ & x_i \geq 0, \ i \in I \subseteq \{1, \ldots, n\} \end{aligned} \tag{39}$$

Let x^{k+1} be a vertex solution of (39). For any $i \in I$ such that $x_i^k = 0$, by (ii) of Proposition 10 it follows that there exists a value M^k such that, if $f_i^{'}(0) \geq M^k$, then we have $x_i^{k+1} = 0$. Thus, if $i \in I$, $x_i^k = 0$ and $f_i^{'}(0) \geq M^k$, then we obtain $x_i^{k+1} = x_i^k = 0$. Letting

$$M = \max_{0 \leq k \leq h} \{M^k\},$$

and assuming

$$f_i^{'}(0) \geq M$$

the thesis follows by induction. □

On the basis of Proposition 8 we can define the following version of the Frank-Wolfe algorithm with unitary stepsize, where the linear problems to be solved are of reduced dimension. In particular, whenever a variable is set to zero at an iteration, the method eliminates this variable for all the following ones.

We denote by Ω the feasible set of problem (38), i.e.,

$$\Omega = \{x \in R^n : x \in P, x_i \geq 0, i \in I\}.$$

Frank-Wolfe—Unitary Stepsize—Reduced Dimension (**FW1-RD**) Algorithm

1. Let $x^0 \in R^n$ be the starting point;
2. For $k = 0, 1, \ldots,$ let $I^k = \{i \in I : x_i^k = 0\}$, $P^k = \{x \in \Omega : x_i = 0\ \forall i \in I^k\}$
   if $x^k \notin \arg\min_{x \in P^k} \nabla f(x^k)^T x$ then compute a vertex solution x^{k+1} of

$$\min_{x \in P^k} \quad \nabla f(x^k)^T x \tag{40}$$

   else exit.

Note that the linear programming problem (40) is equivalent to a linear problem of dimension $n - |I^k|$, and that $I^k \subseteq I^{k+1}$, so that the linear problems to be solved are of nonincreasing dimensions. This yields obvious advantages (shown in the next section) in terms of computational time. Since Algorithm FW1-RD is different from the standard Frank-Wolfe method, its convergence properties cannot be derived from the known result given by Proposition 7. By exploiting the result stated in the appendix, we can formally prove the finite convergence of the algorithm at a stationary point.

Proposition 9 *There exists a value M such that, if $f_j^{'}(0) \geq M$ for $j = 1, \ldots, n$, then Algorithm* FW1-RD *converges to a vertex stationary point of problem* (38) *in a finite number of iterations.*

Proof Since f is a concave differentiable function and is bounded below on Ω, we can write

$$-\infty < \inf_{x \in \Omega} f(x) - f(x^k) \leq f(x) - f(x^k) \leq \nabla f(x^k)^T (x - x^k), \quad \forall\, x \in \Omega.$$

Therefore, as $P^k \subseteq \Omega$, it follows that $\nabla f(x^k)^T x$ is bounded below on the polyhedral set P^k and hence problem (40) admits a vertex solution x^{k+1}, so that Step 2 is well-defined.
We observe that the number of polyhedral sets P^k is finite and hence the number of vertex points generated by the algorithm is finite.

Now we show that $x^k \notin \arg\min_{x \in P^k} \nabla f(x^k)^T x$ implies $f(x^{k+1}) < f(x^k)$. Indeed, in this case we have $\nabla f(x^k)^T (x^{k+1} - x^k) < 0$, and hence, recalling the assumptions on f, we can write

$$f(x^{k+1}) \leq f(x^k) + \nabla f(x^k)^T (x^{k+1} - x^k) < f(x^k). \tag{41}$$

Since the number of points visited by the algorithm is finite, from (41) we get that the algorithm terminates in a finite number k of iterations with a point $x^k \in \arg\min_{x \in P^k} \nabla f(x^k)^T x$. We prove that x^k is a stationary point. Indeed, x^k is a vertex solution of

$$\begin{aligned} \min \quad & \sum_{j: x_j^k \neq 0} f_j'(x_j^k)\, x_j + \sum_{j \notin I^k : x_j^k = 0} f_j'(0)\, x_j \\ & x \in \Omega \\ & x_i = 0, \ i \in I^k \end{aligned}$$

and by (i) of Proposition 10 it follows that there exists a value M such that, if $f_j^{'}(0) \geq M$ then x^k is a solution of

$$\begin{aligned} \min \quad & \sum_{j: x_j^k \neq 0} f_j'(x_j^k)\, x_j + \sum_{j \notin I^k : x_j^k = 0} f_j'(0)\, x_j + \sum_{j \in I^k : x_j^k = 0} f_j'(0) x_j \\ & x \in \Omega \end{aligned} \tag{42}$$

Therefore we have

$$\nabla f(x^k)^T x^k \leq \nabla f(x^k)^T x \quad \forall x \in \Omega,$$

and this proves that x^k is a stationary point of problem (38). □

Concerning the separable concave objective functions of problems (3), (4), (19), (26), we have for $j = 1, \ldots, n$

- $f_j(y_j;\alpha) = 1 - e^{-\alpha y_j}$ and $f_j^{'}(0) = \alpha$;
- $f_j(y_j;\epsilon) = \ln(y_j + \epsilon)$ and $f_j^{'}(0) = 1/\epsilon$;
- $f_j(y_j;\epsilon, p) = (y_j + \epsilon)^p$ and $f_j^{'}(0) = p(\epsilon)^{p-1}$ with $0 < p < 1$;
- $f_j(y_j;\epsilon, p) = -(y_j + \epsilon)^{-p}$ and $f_j^{'}(0) = p(\epsilon)^{-p-1}$ with $1 \leq p$;

Therefore, the assumption of Proposition 9 holds for suitable values of the parameters of the above concave functions, so that Algorithm FW1-RD can be applied to solve problems (3), (4), (19), (26). The results obtained on computational experiments will be presented in the next section.

5 Computational experiments

In our computational experiments we have considered feature selection problems of linear classification models. Given two linearly separable sets of points in a n-dimensional feature space, the problem is that of finding the hyperplane that separates the two sets and utilizes as few of the feature as possible. Formally, given two linearly separable sets

$$S1 = \{u^i \in R^n, i = 1, \ldots, p\} \qquad S2 = \{v^j \in R^n, j = 1, \ldots, q\},$$

the problem is

$$\begin{aligned} \min_{w \in R^n, \theta \in R} \quad & \|w\|_0 \\ & w^T u^i + \theta \geq 1 \quad i = 1, \ldots, p \\ & w^T v^j + \theta \leq -1 \quad j = 1, \ldots, q \end{aligned} \tag{43}$$

Thus, according to the notation adopted in the paper, the problems we used in our experiments take the form

$$\begin{aligned} \min_{x \in R^n, \theta \in R} \quad & \|x\|_0 \\ & A \begin{pmatrix} x \\ \theta \end{pmatrix} \geq e \end{aligned} \tag{44}$$

where $A \in R^{m \times (n+1)}$, $e \in R^m$ is a vector of ones.

We remark that the aim of the experiments has been that of evaluating the effectiveness of various formulations in finding sparse vectors (possibly the sparsest vectors) belonging to polyhedral sets. As said above, the class of problems (44) considered in the experimentation derives from a specific machine learning problem, that is the feature selection problem of linear classifier models. Such a machine learning problem would require to investigate other important issues concerning, for instance, the generalization capability of the linear classifier model determined. This aspect will not be considered here, since it deserves particular attention and will be the object of a future work specifically devoted to the study and the experimentation of feature selection techniques.

We observe that the mixed integer linear programming problem

$$\begin{aligned}
&\min_{x\in R^n,\theta\in R,\delta\in\{0,1\}^n} \quad \sum_{i=1}^{n}\delta_i \\
&A\begin{pmatrix}x\\ \theta\end{pmatrix}\geq e \\
&-M\delta_i \leq x_i \leq M\delta_i \quad i=1,\dots,n \\
&\delta_i\in\{0,1\} \quad i=1,\dots,n
\end{aligned} \tag{45}$$

is equivalent to problem (44) for sufficiently high values of M. Thus, for relatively small dimensional test problems we can determine an optimal solution of (44) by solving (45) by means of an exact method.

Test problems **P-random**. For several values of n and m we randomly generated the matrix A. In particular, each instance of (44) was generated as follows: we randomly defined an hyperplane in a n_1-dimensional space, and we randomly determined m_1 points u^i in an half-space (corresponding to labels $+1$) and other m_2 points v^j in the other half-space (corresponding to labels -1), for a total number of $m=m_1+m_2$ points. We added to each of these vectors a number n_2 of random components, thus obtaining two linearly separable sets, of cardinality m_1 and m_2 respectively, in the space of dimension $n=n_1+n_2$. In this way, the resulting problem (44) had the optimal objective function value less or equal than $n_1<n$.

Colon cancer [1]. The colon cancer dataset contains 22 normal and 40 colon cancer tissues described by 2000 genes expression values extracted from DNA microarray data.

Catalysis. In Catalysis Dataset targets represent the presence (or absence) of catalytic activity of a protein. Inputs are gene expression levels of the genes encoding those proteins. This version of the database was prepared for the Pascal 2004 Evaluating Predictive Uncertainty Challenge. The data are available at http://predict.kyb.tuebingen.mpg.de/pages/home.php.

Nova. This dataset consists of 1754 articles collected from 20 different newsgroups. There are 499 articles related to politics or religion topics and 1255 articles related to other topics. Input variables use a bag-of-words representation with a vocabulary of approximately 17000 words. This version of the database was prepared for the WCCI 2006 challenge on performance prediction. The data are available at http://clopinet.com/isabelle/Projects/modelselect/.

Experiments For each problem we performed experiments using:

- formulation (2), denoted by ℓ_1;
- formulation (3), denoted by *exp*, with $\alpha=5$;
- formulation (4), denoted by *log*, with $\epsilon=10^{-9}$;
- formulation (19), denoted by Formulation 1, with $\epsilon=10^{-9}$ and $p=0.001$;
- formulation (26), denoted Formulation 2, with $\epsilon=10^{-9}$ and $p=1$.

Table 1 Comparison on P-random problems (average zero-norm value/best zero-norm value/percentage of best values attained)

P-random	m	n	$\|x^*\|_0$	l_1	**Exp**	**Log**	**Form. I**	**Form. II**
1	20	10	2	3	3.0/3/100	2.1/**2**/93	2.1/**2**/93	2.0/**2**/97
2	20	10	3	4	4.0/4/100	3.6/**3**/66	3.6/**3**/66	3.9/**3**/45
3	40	20	3	8	8.0/8/100	6.3/4/9	6.2/4/9	5.7/**3**/6
4	40	20	4	10	10.0/10/100	7.7/5/1	7.7/5/1	6.5/5/15
5	60	30	6	12	12.0/12/100	10.0/8/2	10.0/8/2	8.8/**6**/3
6	60	30	7	14	13.9/13/3	10.9/8/1	11.0/8/1	9.7/**7**/6
7	80	40	6	14	14.0/14/100	10.4/7/1	10.4/7/1	9.4/**6**/3
8	80	40	9	24	23.4/22/14	16.4/12/1	16.4/12/1	14.1/11/4
9	100	50	8	19	19.0/19/100	15.1/11/1	15.2/11/1	13.0/**8**/2
10	100	50	10	28	28.0/28/100	18.5/14/3	18.5/14/2	16.0/12/6

Table 2 Comparison on three benchmark problems (average zero-norm value/best zero-norm value/percentage of best values attained)

Problem	m	n	l_1	**Exp**	**Log**	**Form. I**	**Form. II**
Colon Cancer	62	2000	57	8.5/**6**/10	13.8/7/1	13.7/7/1	9.4/**6**/3
Catalysis	873	617	422	199.3/184/1	222.1/201/1	221.0/201/1	189.8/**173**/1
Nova	1754	16 969	448	168.5/147/2	127.0/**105**/1	126.7/**105**/1	131.9/114/1

We applied the Frank-Wolfe algorithm for solving the instances of (3), while we used Algorithm FW1-RD, that is the version of the Frank-Wolfe algorithm presented in the preceding section, for solving problems (4), (19), (26). The reason for which we employed the standard version of the Frank-Wolfe algorithm, instead of Algorithm FW1-RD, for solving the instances of (3) is that the chosen value $\alpha = 5$, suggested in [3], did not seem sufficiently high to ensure that the assumptions of Proposition 9 were satisfied. We used 100 random initial points for all the problems.

Implementation details The instances of problem (45) were solved by means of CPLEX (8.0). Algorithms FW1 and FW1-RD were implemented in C using GLPK (4.9) as solver of the linear programming problems. The experiments were carried out on Intel Pentium 4 3.2 GHz 512 MB RAM.

Results The results obtained on P-random problems and on the other three test problems are shown in Tables 1 and 2 respectively, where we report

- the number m of constraints, the number n of variables;
- for formulation ℓ_1, the zero-norm of the optimal solution attained;
- for each nonlinear concave formulation:
  - the average of the zero-norm value of the stationary points determined;
  - the best zero-norm value of those stationary points;
  - percentage of runs where the best zero-norm value was attained.

Table 3 Comparison using Formulation 1 between the two versions of the Frank-Wolfe algorithm in terms of CPU-time (seconds)

Problem	FW1	FW1-RD
Colon Cancer	225	24
Catalysis	2776	465
Nova	10 448	1003

In Table 1, which concerns relatively small dimensional problems, we also report the optimal value $\|x^*\|_0$ determined by solving (45).

From Table 1 we can observe that the best results are obtained by means of Formulation 2. Indeed, in seven problems over ten, a simple multi-start strategy applied to Formulation 2 allowed us to attain the certificated optimal solution. We may note that the results obtained by means of formulations *log* and Formulation 1 are comparable, and clearly better than those corresponding to formulations ℓ_1 and *exp*.

The results obtained on problems Colon cancer, Catalysis, and Nova are reported in Table 2, where we can observe that the multi-start strategy applied to the nonlinear concave formulations performed clearly better than the approach based on the minimization of the ℓ_1 norm. Furthermore, we can note that the best results on problem Colon Cancer were obtained by *exp* and Formulation 1, the best results on problem Catalysis were obtained by Formulation 2, while the best results on problem Nova were obtained by *log* and Formulation 1.

Summarizing, the computational experiments confirm the validity of the concave-based approach for the minimization of the zero-norm over a polyhedral set, and show that the concave formulations here proposed are valid alternatives to known formulations. Indeed, Formulation 1 and Formulation 2 attained the best results in 3 tests over 13 and 9 tests over 13 respectively. We remark that a wider availability of efficient formulations is important since it can facilitate the search of sparse enough solutions for different classes of problems.

Finally, in order to assess the differences in terms of computational time between the standard Frank-Wolfe (FW1) algorithm and the version of the algorithm presented in the preceding section and denoted by Algorithm FW1-RD, we report in Table 3 the results obtained by the two algorithms on the three benchmark problems using Formulation 1. As we might expect, the differences are remarkable and show the usefulness of Algorithm FW1-RD. Further experiments not here reported and performed using the other concave formulations point out the same differences between the two algorithms in terms of computational time. In all the tests we did not detect differences between the two algorithms in terms of computed solution.

6 Conclusions

In this work we have considered the general hard problem of minimizing the zero-norm over polyhedral sets, which arises in different important fields, such as machine learning and signal processing. Following the concave optimization-based approach, we have proposed two new smooth concave formulations and we have formally proved the equivalence of these and other formulations with the original non-smooth problem. The main contributions of this paper are both theoretical and computational. From the theoretical point of view, we have been able to introduce some general results on approximability for concave optimization problems and we obtained an important characterization of the behavior of the Frank-Wolfe algorithm which has, as we could confirm in computational experiments, a dramatic influence on the efficiency of the method. The computational evidence we report suggests a speed-up in the range 5 to 10 when using the variable fixing variant of the Franke-Wolfe method in place of the traditional one. This very high speed-up might prove to be extremely beneficial when multiple runs of the algorithm are performed, e.g. in a Multistart method. Apart from the great improvement in efficiency, the computational experiments also evidenced that the new formulations are valid alternatives to known formulations, as in most cases they allowed us to compute highly sparse solutions. We remark that a wider availability of efficient formulations is important since it can facilitate the search of sparse enough solutions for different classes of problems.

Future work will be devoted to develop global optimization algorithms for the minimization of concave separable functions over polyhedral sets and to define suitable techniques for the feature selection problem.

Appendix

For convenience of the reader we report a known result (and its proof) that we have used to derive some new convergence results of the Frank-Wolfe method and of the modified version we have presented.

Proposition 10 *Consider the linear programming problems*

$$\begin{aligned} \min \quad & c^T x \\ & Ax \geq b \\ & Hx = d \end{aligned} \tag{46}$$

$$\begin{aligned} \min \quad & c^T x + Me^T z \\ & Ax + Qz \geq b \\ & Hx + Sz = d \\ & z \geq 0 \end{aligned} \tag{47}$$

where $c \in R^n$, $e \in R^{n_z}$ *is a vector of ones*, $b \in R^m$, $d \in R^p$, $A \in R^{m\times n}$, $H \in R^{p\times n}$, $Q \in R^{m\times n_z}$, $S \in R^{p\times n_z}$. *Assume that problem* (46) *admits a solution* $x^\star$. *Then, there exists a value* M_0 *such that for all* $M \geq M_0$ *we have that*

(i) *the vector* $(x^\star, 0)^T$ *is a solution of* (47);
(ii) *if* $(\bar{x}, \bar{z})^T$ *is a solution of* (47), *then* $\bar{z} = 0$ *and* $\bar{x}$ *is a solution of* (46).

Proof (i) Since $x^\star$ is a solution of problem (46) we have that the dual problem

$$\begin{aligned} \max \quad & b^T\lambda + d^T\mu \\ & A^T\lambda + H^T\mu = c \\ & \lambda \geq 0 \end{aligned} \tag{48}$$

admits a solution $(\lambda^\star, \mu^\star)^T \in R^{m+p}$ and we have

$$c^T x^\star = b^T\lambda^\star + d^T\mu^\star. \tag{49}$$

Now consider problem (47) and its dual

$$\begin{aligned} \max \quad & b^T\lambda + d^T\mu \\ & A^T\lambda + H^T\mu = c \\ & Q^T\lambda + S^T\mu \leq Me \\ & \lambda \geq 0. \end{aligned} \tag{50}$$

The vector $(x^\star, 0)^T$ is a feasible point for (47), and for M sufficiently large the vector $(\lambda^\star, \mu^\star)^T$ is a feasible point for (50). Thus, from (49) the assertion is proved. Furthermore, we can also conclude that $(\lambda^\star, \mu^\star)^T$ is a solution of (50) for M sufficiently large.

(ii) By contradiction let us assume that there exist a sequence of positive scalars $\{M^k\}$, with $M^k \to \infty$ for $k \to \infty$, and a corresponding sequence of vectors $\{(x^k, z^k)^T\}$ such that $z^k \neq 0$, and $(x^k, z^k)^T$ is solution of (47) when $M = M^k$. We can then define an infinite subset K such that, for all $k \in K$

- we have $z_i^k > 0$ for some index $i \in \{1, \dots, n_z\}$;
- the vector $(\lambda^\star, \mu^\star)^T$ is a solution of (50) when $M = M^k$.

Then, using the complementarity conditions we can write

$$\left(e_i^T Q^T\lambda^\star + e_i^T S^T\mu^\star - M^k\right) = 0 \quad \forall k \in K,$$

which contradicts the fact that $M^k \to \infty$. □

References

1. Alon, U., Barkai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., Levine, A.J.: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissue probed by oligonucleotide arrays. Proc. Natl. Acad. Sc. **96**, 6745–6750 (1999)
2. Amaldi, E., Kann, V.: On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems. Theor. Comput. Sci. **209**, 237–260 (1998)
3. Bradley, P.S., Mangasarian, O.L.: Feature selection via concave minimization and support vector machines. In: Shavlik, J. (ed.) Machine Learning Proceedings of the Fifteenth International Conference (ICML '98), pp. 82–90. Morgan Kaufmann, San Francisco (1998)
4. Bruckstein, A.M., Donoho, D.L., Elad, M.: From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM Rev. (2008, to appear)

5. Chen, S.S., Donoho, D.L., Saunders, M.A.: Atomic decomposition basis pursuit. SIAM Rev. **43**, 129–159 (2001)
6. Donoho, D.L., Elad, M., Temlyakov, V.N.: Stable recovery of sparse overcomplete representation in the presence of noise. IEEE Trans. Inf. Theory. **52**, 6–18 (2006)
7. Frank, M., Wolfe, P.: An algorithm for quadratic programming. Nav. Res. Logist. Q. **3**, 95–110 (1956)
8. Gribonval, R., Nielsen, M.: Sparse representation in union of bases. IEEE Trans. Inf. Theory. **49**, 3320–3325 (2003)
9. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**, 1157–1182 (2003)
10. Mangasarian, O.L.: Machine learning via polyhedral concave minimization. In: Fischer, H., Riedmueller, B., Schaeffler, S. (eds.) Applied Mathematics and Parallel Computing—Festschrift for Klaus Ritter, pp. 175–188. Physica, Heidelberg (1996)
11. Warmack, R.E., Gonzalez, R.C.: An algorithm for optimal solution of linear inequalities and its application to pattern recognition. IEEE Trans. Comput. **22**, 1065–1075 (1973)
12. Weston, J., Elisseef, A., Scholkopf, B.: Use of the zero-norm with linear models and kernel model. J. Mach. Learn. Res. **3**, 1439–1461 (2003)