Minimum-Support Solutions of Polyhedral Concave Programs *

O. L. Mangasarian†

This paper is dedicated to the memory of Karl-Heinz Elster

Abstract

Motivated by the successful application of mathematical programming techniques to difficult machine learning problems, we seek solutions of concave minimization problems over polyhedral sets with a minimum number of nonzero components. We prove that if such problems have a solution, they have a vertex solution with a minimal number of zeros. This includes linear programs and general linear complementarity problems. A smooth concave exponential approximation to a step function solves the minimum-support problem exactly for a finite value of the smoothing parameter. A fast finite linear-programming-based iterative method terminates at a stationary point, which for many important real world problems provides very useful answers. Utilizing the complementarity property of linear programs and linear complementarity problems, an upper bound on the number of nonzeros can be obtained by solving a single convex minimization problem on a polyhedral set.

1 Introduction

This work is prompted by applications of mathematical programming to machine learning and other problems where a solution to a mathematical program is desired with as many of its components equal to zero as possible [14, 3, 15, 2]. For example in determining a separating plane in the n-dimensional real space R^n that separates as best as possible two given sets $\mathcal{A}$ and $\mathcal{B}$ from each other, it is often desirable to achieve this separation in as small dimensional subspace of R^n as possible. This leads to better "generalization" [22], that is a plane that will discriminate better between new elements of $\mathcal{A}$ and $\mathcal{B}$ that were not involved in determining the original separating plane [21, 14, 3]. We are thus interested in obtaining *minimum-support* solutions to mathematical programs, that is solutions with as many components equal to zero as possible. Since there is virtually no previous work on minimum-support solutions of mathematical programs, it is one of the primary purposes of this paper to bring to the attention of the mathematical programming community the importance and the richness of such problems from the point of view of machine learning [14, 3, 2] and data mining [15, 1]. A related area in these fields that has received recent wide attention is that of support vector machines [20, 7, 4] where instead of suppressing solution components, as we are doing here,

*Mathematical Programming Technical Report 97-05, April 1997. Revised March 1998. This material is based on research supported by National Science Foundation Grant CCR-9322479 and Air Force Office of Scientific Research Grant F49620-97-1-0326.

†Computer Sciences Department, University of Wisconsin, 1210 West Dayton Street, Madison, WI 53706, *olvi@cs.wisc.edu*. The author gratefully acknowledges the gracious hospitality of the Mathematics Department of the University of California at San Diego during his sabbatical leave January-May 1997.

emphasis is placed on obtaining the relatively few active constraints with positive multipliers that determine a solution.

The class of problems that we shall consider in this paper and which subsumes many of the machine learning problems is the following one:

$$\min_{x \in S} \; f(x), \tag{1}$$

where $f : R^n \longrightarrow R$ is a concave function on R^n that is bounded below on the polyhedral set S. This problem includes linear programs of course, but also includes NP-hard problems, such as the general linear complementarity problem [6, 5] and others. We shall therefore be concerned with existence of solutions to problem (1) with minimum number of nonzero elements, and methods to for computing such solutions. It will turn out (Corollaries 2.4 and 2.5) that solvable linear programs and general linear complementarity problems always have a vertex of their original feasible region as a solution with the minimum-support property. In Section 2 we give our principal result, Theorem 2.3, which establishes the existence of a minimum-support vertex solution for problem (1) and shows how to compute it as a solution of a smooth exponential penalty problem (13) for finite values of both the penalty and exponential parameters. Section 3 describes a finite Successive Linearization Algorithm 3.1 for obtaining a stationary point of the smooth penalty problem (13). In Section 4 we utilize the complementarity property of dual solutions of linear programs and of linear complementarity problems to give bounds on the cardinalities of nonzero elements of minimum-support solutions of linear programs (Theorem 4.2) and general linear complementarity problems (Theorem 4.3).

1.1 Notation

A word about our notation and background material. All vectors will be column vectors unless transposed to a row vector by a prime superscript $'$. For a vector x in the n-dimensional real space R^n, $|x|$ will denote a vector of absolute values of components $x_i,\ i = 1, \ldots, n$ of x. The step function x_* of $x \in R^n$ is defined as a vector in R^n of zeros and ones, with ones corresponding to positive components of x and zeros corresponding to nonpositive components. The notation min $\{x, y\}$ applied to vectors x and y in R^n will denote a vector with components that are minima of corresponding components of x and y. The scalar product of two vectors x and y in the n-dimensional real space will be denoted by $x'y$. For a linear program $\min\limits_{x \in X} c'x$ the notation arg vertex $\min\limits_{x \in X} c'x$, will denote the set of vertex solutions of the linear program. For $x \in R^n$, the norm $\|x\|_2$ will denote the 2-norm: $(x'x)^{\frac{1}{2}}$, while $\|x\|_1$ will denote the 1-norm: $\sum\limits_{i=1}^{n} |x_i|$. For an $m \times n$ matrix A, A_i will denote the ith row of A and A_{ij} will denote the element in row i and column j. The identity matrix in a real space of arbitrary dimension will be denoted by I, while a column vector of ones of arbitrary dimension will be denoted by e. The base of the natural logarithm will be denoted by ε, and for $y \in R^m$, ε^{-y} will denote a vector in R^m with component $\varepsilon^{-y_i},\ i = 1, \ldots, m$. For a function $f : R^n \to R$ that is concave on R^n, the supergradient $\partial f(x)$ of f at x is a vector in R^n satisfying

$$f(y) - f(x) \leq \partial f(x)(y - x) \tag{2}$$

for any $y \in R^n$. The set $D(f(x))$ of supergradients of f at the point x is nonempty, convex, compact and reduces to the ordinary gradient $\nabla f(x)$, when f is differentiable at x [18, 19]. For a point $x \in R^n$, $card(x)$ will denote the cardinality of the nonzero components of x, that is the number of the nonzero components of x.

2 Minimum-Support Vertex Solutions

We begin this section with two preliminary lemmas that will enable us to derive our principal result, the existence of a minimum-support vertex solution to the polyhedral concave minimization problem.

Lemma 2.1 Vertex Projection *Let T be a closed convex set in $R^n \times R^k$. If $(\bar{x}, \bar{y})$ is a vertex of T, then $\bar{x}$ is a vertex of $T(\bar{y})$ where*

$$T(\bar{y}) := \{x \mid (x, \bar{y}) \in T\}. \tag{3}$$

Proof The point $(\bar{x}, \bar{y})$ is a vertex of T if and only if

$$\left\langle \begin{array}{c} (x^1, y^1) \in T,\ (x^2, y^2) \in T \\ (\bar{x}, \bar{y}) = ((1-\lambda)x^1 + \lambda x^2,\ (1-\lambda)y^1 + \lambda y^2) \\ 0 < \lambda < 1 \end{array} \right\rangle \Longrightarrow (x^1, y^1) = (x^2, y^2), \tag{4}$$

from which follows that $\bar{x}$ is a vertex of $T(\bar{y})$ if we take $y^1 = y^2 = \bar{y}$. $\diamondsuit$

The following lemma, also needed in the proof of our principal result, shows that the concave minimization problem (1) penalized by the cardinality-minimizing term $e'|x|_*$ has a vertex solution. The proof which utilizes a smooth exponential to approximate the step function also leads to a computational approach for solving the problem.

Lemma 2.2 Vertex Solution of Penalty Problem *Let $f : R^n \longrightarrow R$ be a concave function on R^n that is bounded below on the polyhedral set $S \subset R^n$, and let S have no straight lines going to infinity in both directions. Then the problem*

$$\min_{x \in S}\ f(x) + \beta e'|x|_*, \quad \beta > 0, \tag{5}$$

has a vertex of S as a solution which is also a vertex solution of:

$$\min_{(x,y) \in T} f(x) + \beta e'(e - \varepsilon^{-\alpha y}), \tag{6}$$

where

$$T := \{(x, y) \mid x \in S,\ y \in R^n,\ -y \leq x \leq y\}. \tag{7}$$

for some $\alpha \geq \alpha_0(\beta) > 0$.

Proof Since, for $\alpha > 0$,

$$(e - \varepsilon^{-\alpha|x|}) \leq |x|_*, \tag{8}$$

it follows that

$$\min_{(x,y) \in T} f(x) + \beta e'(e - \varepsilon^{-\alpha y}) = \min_{x \in S} f(x) + \beta e'(e - \varepsilon^{-\alpha|x|}) \leq \inf_{x \in S} f(x) + \beta e'|x|_*. \tag{9}$$

Since T has a finite number of vertices, one vertex, say $(\bar{x}, \bar{y})$, will repeatedly solve

$$\min_{(x,y) \in T} f(x) + \beta e'(e - \varepsilon^{-\alpha_i y}), \tag{10}$$

for some sequence of positive numbers $\{\alpha_0, \alpha_1, \ldots\} \uparrow \infty$. This follows from the fact that a concave function bounded below on a polyhedral set not containing lines going to infinity in both directions attains a minimum at a vertex of the polyhedral set [19, Corollaries 32.3.3 & 32.3.4]. Hence for $\alpha_i \geq \alpha_0(\beta)$:

$$\begin{array}{rcl} f(\bar{x}) + \beta e'(e - \varepsilon^{-\alpha_i|\bar{x}|}) & = & f(\bar{x}) + \beta e'(e - \varepsilon^{-\alpha_i \bar{y}}) \\ & = & \min\limits_{(x,y)\in T} f(x) + \beta e'(e - \varepsilon^{-\alpha_i y}) \\ & = & \min\limits_{x\in S} f(x) + \beta e'(e - \varepsilon^{-\alpha_i|x|}) \\ & \leq & \inf\limits_{x\in S} f(x) + \beta e'|x|_*, \end{array} \tag{11}$$

where the last inequality follows from (8). Letting $i \longrightarrow \infty$ gives:

$$f(\bar{x}) + \beta e'|\bar{x}|_* \leq \inf_{x\in S} f(x) + \beta e'|x|_*. \tag{12}$$

Since $(\bar{x}, \bar{y})$ is a vertex of T, it follows by Lemma 2.1 that $\bar{x}$ is a vertex of S and hence $\bar{x}$ is a vertex solution of (5). By (11) $\bar{x}$ is a vertex solution of (6) for $\alpha \in \{\alpha_i\} \geq \alpha_0 = \alpha_0(\beta)$. $\diamondsuit$

We establish now our principal result, that of the existence of a minimum-support vertex solution to our original concave minimization problem, via a sufficient penalty function argument similar to that used in [8, Theorem 4.1],[11, Theorem 2.3], for a finite value of the penalty parameter. This result will also be useful in specifying a finite successive linear approximation algorithm for obtaining a stationary point of the NP-hard minimum-support problem. Related finite penalty parameter results are also given in [13, 17].

Theorem 2.3 Existence of Exact Vertex Solution for Finite Values of Smoothing and Penalty Parameters *Let $f : R^n \longrightarrow R$ be a concave function on R^n that is bounded below on the polyhedral set $S \subset R^n$, and let S have no straight lines going to infinity in both directions. Then problem (1) has a minimum-support solution x which is a vertex of S and such that x and some $y \in R^n$ constitute a vertex solution of:*

$$\min_{(x,y)\in T} f(x) + \beta e'(e - \varepsilon^{-\alpha y}),\ \beta \leq \beta_0,\ \alpha \geq \alpha_0(\beta), \tag{13}$$

for some $\beta_0 > 0$ and $\alpha_0(\beta) > 0$, where T is defined in (7).

Proof By [19, Corollary 32.3.3] problem (1) has a solution because the concave function f is bounded below on the polyhedral set S. Furthermore, by [19, Corollary 32.3.4] it has a vertex solution because S contains no straight lines that go to infinity in both directions. Applying Lemma 2.2 we conclude that since S has a finite number of vertices, one vertex, say $\hat{x}$, will repeatedly solve the penalized problem (5) for β belonging to some decreasing sequence of positive numbers: $\{\beta_0, \beta_1, \ldots\} \downarrow 0$. Hence for $\beta_i \leq \beta_0$:

$$\hat{x} \in \arg\ \min_{x\in S}\ f(x) + \beta_i e'|x|_* = \arg\ \min_{x\in S}\ f(x) - f(\tilde{x}) + \beta_i e'|x|_*, \tag{14}$$

where $\tilde{x}$ is defined as a solution of (1), that is

$$\tilde{x} \in S \text{ and } f(\tilde{x}) = \min_{x\in S} f(x). \tag{15}$$

First we show by contradiction that

$$f(\hat{x}) - f(\tilde{x}) \leq 0\ \ (\text{or equivalently } f(\hat{x}) - f(\tilde{x}) = 0). \tag{16}$$

For if not, then for

$$\frac{1}{\beta_i} > \max\{\frac{e'|\tilde{x}|_* - e'|\hat{x}|_*}{f(\hat{x}) - f(\tilde{x})},\ \frac{1}{\beta_0}\}, \tag{17}$$

we get the contradiction

$$e'|\tilde{x}|_* = e'|\tilde{x}|_* + \frac{1}{\beta_i}(f(\tilde{x}) - f(\tilde{x})) \geq e'|\hat{x}|_* + \frac{1}{\beta_i}(f(\hat{x}) - f(\tilde{x})) > e'|\hat{x}|_* + e'|\tilde{x}|_* - e'|\hat{x}|_*, \tag{18}$$

where the first inequality follows from (14) and the last inequality from (17). Hence (16) holds. We now show that $e'|\hat{x}|_*$ is a minimum over the set of solutions of (1), that is $\hat{x}$ has a minimal number of zeros besides being optimal for problem (1). We have for $\beta_i \leq \beta_0$ and all optimal x, that is $x \in S$ and $f(x) \leq f(\tilde{x})$:

$$e'|\hat{x}|_* = e'|\hat{x}|_* + \frac{1}{\beta_i}(f(\hat{x}) - f(\tilde{x})) \leq e'|x|_* + \frac{1}{\beta_i}(f(x) - f(\tilde{x})) \leq e'|x|_*, \tag{19}$$

where the first inequality follows from (14) and the second inequality from $f(x) - f(\tilde{x}) \leq 0$. Hence $\hat{x}$ is an optimal vertex solution with a minimal number of zeros. The last statement of the theorem follows by applying Lemma 2.2 to the penalized problem (13) for $\beta \leq \beta_0$. $\diamondsuit$

As an immediate consequence of this theorem we can establish the existence of a minimum-support vertex solution to a linear program in standard form because it satisfies all the hypotheses of the above theorem.

Corollary 2.4 Minimum-Support Vertex Solution of Linear Programs *Let the linear program*

$$\min_{x \in S_1} c'x \text{ where } S_1 = \{x \mid Ax \geq b,\ x \geq 0\} \tag{20}$$

have a solution, then it has a minimum-support solution which is a vertex of S_1 and such that x and some y constitute a vertex solution of:

$$\min_{(x,y)}\{c'x + \beta e'(e - \varepsilon^{-\alpha y}) \mid Ax \geq b,\ x \geq 0,\ -y \leq x \leq y\},\ \beta \leq \beta_0,\ \alpha \geq \alpha_0(\beta), \tag{21}$$

for some $\beta_0 > 0$ and $\alpha_0(\beta) > 0$.

For the general linear complementarity problem

$$0 \leq x \ \perp\ Mx + q \geq 0, \tag{22}$$

where M is a given matrix in $R^{n\times n}$, it is easy to show [9, Lemma 1] that the problem is equivalent to the following concave minimization problem:

$$0 = \min_x\ \{e' \min\{x,\ Mx+q\} \mid Mx + q \geq 0,\ x \geq 0\}. \tag{23}$$

Again because problem (22) fits the hypotheses of our principal Theorem 2.3, the following result follows immediately.

Corollary 2.5 Minimum-Support Vertex Solution of General Linear Complementarity Problems *Any solvable general linear complementarity problem has a minimum-support solution which is a vertex of its feasible region $\{x \mid Mx + q \geq 0,\ x \geq 0\}$ and such that x and some $y \in R^n$ constitute a vertex of solution of:*

$$\min_{(x,y)}\{e' \min\{x,\ Mx+q\} + \beta e'(e - \varepsilon^{-\alpha y}) \mid Mx + q \geq 0,\ x \geq 0,\ -y \leq x \leq y\},\ \beta \leq \beta_0,\ \alpha \geq \alpha_0(\beta), \tag{24}$$

for some $\beta_0 > 0$ and $\alpha_0(\beta) > 0$.

Note that since the minimum of two linear functions is a concave function, the objective function of (24) is concave but not differentiable on R^n. We turn our attention to computational methods for finding minimum-support solutions. The methods are based on the penalty approach formulation described above that leads to minimizing a concave function on a polyhedral set.

3 Computational Algorithms

We shall employ the finitely terminating fast successive linearization algorithm (SLA) proposed in [14] for differentiable concave function minimization and [16] for nondifferentiable concave function minimization on polyhedral sets. Since the linear complementarity minimum-support penalty reformulation (24) has a nondifferentiable concave objective function, we shall give the more general SLA version of [16].

We consider the penalty formulation (13) and apply the SLA algorithm of [16] to it. This gives the following algorithm.

Algorithm 3.1 Successive Linearization Algorithm (SLA) *Choose $\alpha > 0$, $\beta > 0$ and start with a random $x^0 \in R^n$ and $y^0 = |x^0|$. Having (x^i, y^i) determine (x^{i+1}, y^{i+1}) as follows:*

$$(x^{i+1}, y^{i+1}) \in \text{arg vertex} \min_{x \in T} \partial f(x^i)(x - x^i) + \alpha\beta\varepsilon^{-\alpha y^{i'}}(y - y^i), \tag{25}$$

where T is defined in (7). Stop if

$$(x^i, y^i) \in T \text{ and } \partial f(x^i)(x - x^i) + \alpha\beta\varepsilon^{-\alpha y^{i'}}(y - y^i) = 0$$

.

By Theorem 3 of [16] we can establish the following finite termination result of the SLA 3.1 at a stationary point satisfying the minimum principle necessary optimality condition.

Theorem 3.2 SLA Finite Termination Theorem *Let f, a concave function on R^n, be bounded below on S, and let S be polyhedral and not containing straight lines that go to infinity in both directions. The SLA 3.1 generates a finite sequence of feasible iterates $\{(x^1, y^1), (x^2, y^2), \ldots, (x^{\bar{i}}, y^{\bar{i}})\}$ of strictly decreasing objective function values of (13), such that $(x^{\bar{i}}, y^{\bar{i}})$ satisfies the minimum principle necessary optimality condition:*

$$\partial f(x^{\bar{i}})(x - x^{\bar{i}}) + \alpha\beta\varepsilon^{-\alpha y^{\bar{i}'}}(y - y^{\bar{i}}), \;\; \forall (x, y) \in T. \tag{26}$$

4 Upper-Bounding of Cardinality via Complementarity

In this section of the paper we use the complementarity condition present in solutions of a pair of dual linear programs and in complementarity problems to obtain upper bounds on the cardinality of the nonzero elements of a minimum-support solution by solving a concave maximization problem in the dual variable complementary to the variable that we are trying to obtain a minimum-support solution for. We begin with the following simple lemma towards that purpose.

Lemma 4.1 Bounding Cardinality of Nonzero Elements *Let $x \in R^n$, $y \in R^n$. Then:*

$$\left\langle \begin{array}{c} 0 \le x \perp w \ge 0, \; x + w > 0 \; (i.e. nondegenerate) \implies \inf\limits_x \; e'x_* + \sup\limits_w \; e'w_* = n \\ 0 \le x \perp w \ge 0, \; x + w \not> 0 \; (i.e. degenerate) \implies \inf\limits_x \; e'x_* + \sup\limits_w \; e'w_* \le n \end{array} \right\rangle \tag{27}$$

Proof In the nondegenerate case when $x + w > 0$ we have:

$$x_* + w_* = e \implies \inf_x \ e'x_* = n + \inf_w \ -e'w_* \implies \inf_x \ e'x_* = n - \sup \ e'w_*. \tag{28}$$

In the degenerate case when $x + w \not> 0$ we have:

$$x_* + w_* \leq e; \implies \inf_x \ e'x_* \leq n + \inf_w \ -e'w_* \implies \inf_x \ e'x_* \leq n - \sup \ e'w_*. \tag{29}$$

$\diamondsuit$

By using the above lemma, the inequality (8) and the characterization of bounded solutions of monotone linear complementarity problems and linear programs of [10] we can obtain the following upper-bounds on the cardinality of the nonzero elements of solutions by solving a concave maximization problem. If the solution set is nondegenerate, then these upper bounds equal the actual cardinality.

Theorem 4.2 Upper Bounding Nonzero Elements of LP Solutions *Let $\bar{S}_1$ be the nonempty solution set of the linear program (20), and let $\bar{U}_1$ be the corresponding nonempty solution set of its dual:*

$$\max_{u \in U_1} b'u \text{ where } U_1 = \{u \mid A'u \leq c, \ u \geq 0\}. \tag{30}$$

Then for any $\alpha > 0$, it follows that

$$\min_{x \in \bar{S}_1} \ e'x_* \ \leq n - \sup_{u \in \bar{U}_1} e'(A'u - c)_* \ \leq \ \inf_{u \in \bar{U}_1} \ e'\varepsilon^{-\alpha(-A'u+c)}. \tag{31}$$

In case of a nondegenerate primal-dual solution set, the first inequality of (31) is an equality. In case of boundedness of the dual optimal set $\bar{U}_1$ which is equivalent to [10, Theorem 3.1]:

$$\{x \mid Ax \geq b, \ x > 0\} \neq \emptyset, \tag{32}$$

then the inf of (31) can be replaced by a min.

Proof The first inequality of (31) follows from the second implication of Lemma 4.1, and the second inequality follows from inequality (8). When the primal-dual solution set is nondegenerate, then by the first implication of Lemma 4.1, the first inequality can be replaced by an equality. If (32) holds, then by Theorem 3.1 of [10], the optimal dual solution set $\bar{U}_1$ is bounded. Hence the continuous function $e'\varepsilon^{-\alpha(-A'u+c)}$ which is bounded below on bounded compact set $\bar{U}_1$, attains its infimum on $\bar{U}_1$. Hence the inf in (31) can be replaced by a minimum. $\diamondsuit$

We consider now the linear complementarity problem and establish similarly an upper bound for the cardinality of its minimum-support solution.

Theorem 4.3 Upper Bounding Nonzero Elements of LCP Solutions *Let $\bar{S}_2$ be the nonempty solution set of the linear complementarity problem (22). Then for any $\alpha > 0$, it follows that*

$$\min_{x \in \bar{S}_2} \ e'x_* \ \leq n - \sup_{x \in \bar{S}_2} e'(Mx + q)_* \ \leq \ \inf_{x \in \bar{S}_2} \ e'\varepsilon^{-\alpha(Mx+q)}. \tag{33}$$

In case of a nondegenerate solution set, that is

$$0 \leq x \ \perp \ Mx + q \geq 0 \implies x + Mx + q > 0, \tag{34}$$

the first inequality of (33) is an equality. In case of boundedness of the solution set $\bar{S}_2$, that is for some $\gamma > 0$,

$$0 \leq x \ \perp \ Mx + q \geq 0 \implies e'x \leq \gamma, \tag{35}$$

then the inf of (33) can be replaced by a min.

Proof The first inequality follows from the second implication of Lemma 4.1 and the complementarity condition $0 \leq x \perp Mx + q \geq 0$. The second inequality follows from inequality (8). If the solution set is nondegenerate, then by the first implication of Lemma 4.1, the first inequality of (33) can be replaced by an equality. If (35) holds, then the closed solution set $\bar{S}_2$ is bounded and the continuous function $e'\varepsilon^{-\alpha(Mx+q)}$ attains its infimum on $\bar{S}_2$. Hence the inf in (33) can be replaced by a minimum. $\diamondsuit$

If we make use of the polyhedral characterization of the solution set of a monotone linear complementarity problem [12, Corollary 2], then the solution set $\bar{S}_2$ in the above theorem can be explicitly given in terms of any solution point $\bar{x}$ as follows.

Corollary 4.4 Upper Bounding Nonzero Elements of Minimum-Support Solution of Monotone LCP *Let M be positive semidefinite and let $\bar{x}$ be any element of the solution set $\bar{S}_2$ of the LCP (22). Then Theorem 4.3 holds with $\bar{S}_2$ given by the following explicit form:*

$$\bar{S}_2 = \{x \mid Mx + q \geq 0,\ x \geq 0,\ q'(x - \bar{x}) = 0,\ (M + M')(x - \bar{x}) = 0\}, \tag{36}$$

By solving the above convex program with a polyhedral constraint we are able to give an upper bound (or the exact number in case of nondegeneracy) on the cardinality of nonzero elements in a minimum-support solution of a monotone linear complementarity problem.

5 Summary & Conclusion

Motivated by machine learning principles where "simplest is best", we have established the existence of a minimum-support solution to concave minimization problems on polyhedral sets as a vertex of their feasible region. This class of problems includes linear programs and general linear complementarity problems. Our existence result is based on a constructive approach that approximates the step function by a concave exponential from below. This leads to a computational algorithm that consists of solving a finite number of linear programs and terminates at a stationary point. This approach has been quite successful in machine learning applications. Finally we used the complementarity property inherent in linear programs and linear complementarity problems to bound the cardinality of the nonzero elements of minimum-support solution by solving a single convex program. The upper bound is exact if the solution set is nondegenerate.

It is hoped that the results presented in this work will lead to further investigation and application of polyhedral concave minimization problems. These problems, even though NP-hard, can yield through their easily computed stationary points, significant results for many real world problems.

References

[1] P. S. Bradley, Usama M. Fayyad, and O. L. Mangasarian. Data mining: Overview and optimization opportunities. Technical Report 98-01, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, January 1998. INFORMS Journal on Computing, submitted. Available from: ftp://ftp.cs.wisc.edu/math-prog/tech-reports/998-01.ps.Z.

[2] P. S. Bradley, O. L. Mangasarian, and J. B. Rosen. Parsimonious least norm approximation. Technical Report 97-03, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, March 1997. *Computational Optimization and Applications*, to appear. ftp://ftp.cs.wisc.edu/math-prog/tech-reports/97-03.ps.Z.

[3] P. S. Bradley, O. L. Mangasarian, and W. N. Street. Feature selection via mathematical programming. INFORMS Journal on Computing, 1998. To appear. Available at ftp://ftp.cs.wisc.edu/math-prog/tech-reports/95-21.ps.Z.

[4] C. J. C. Burges. Simplified support vector decision rules. In L. Saita, editor, *Machine Learning–Proceedings of the THirteenth International Conference (ICML '96)–Bari, Italy July 3-6, 1996*, pages 71–77, San Francisco, CA, 1996. Morgan Kaufmann.

[5] S.-J. Chung. NP-completeness of the linear complementarity problem. *Journal of Optimization Theory and Applications*, 60:393–399, 1989.

[6] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem.* Academic Press, New York, 1992.

[7] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik. Support vector regression machines. In M. C. Mozer, M. I. Jordan, and T. Petsche, editors, *Advances in Neural Information Processing Systems -9-*, pages 155–161, Cambridge, MA, 1997. MIT Press.

[8] S.-P. Han and O. L. Mangasarian. Exact penalty functions in nonlinear programming. *Mathematical Programming*, 17:251–269, 1979.

[9] O. L. Mangasarian. Characterization of linear complementarity problems as linear programs. *Mathematical Programming Study*, 7:74–87, 1978.

[10] O. L. Mangasarian. Simple computable bounds for solutions of linear complementarity problems and linear programs. *Mathematical Programming Study*, 25:1–12, 1985.

[11] O. L. Mangasarian. Some applications of penalty functions in mathematical programming. In R. Conti, E. De Giorgi, and F. Giannessi, editors, *Optimization and Related Fields*, pages 307–329. Springer–Verlag, Heidelberg, 1986. Lecture Notes in Mathematics 1190.

[12] O. L. Mangasarian. A simple characterization of solution sets of convex programs. *Operations Research Letters*, 7(1):21–26, 1988.

[13] O. L. Mangasarian. Misclassification minimization. *Journal of Global Optimization*, 5:309–323, 1994.

[14] O. L. Mangasarian. Machine learning via polyhedral concave minimization. In H. Fischer, B. Riedmueller, and S. Schaeffler, editors, *Applied Mathematics and Parallel Computing - Festschrift for Klaus Ritter*, pages 175–188. Physica-Verlag A Springer-Verlag Company, Heidelberg, 1996. Available at ftp://ftp.cs.wisc.edu/math-prog/tech-reports/95-20.ps.Z.

[15] O. L. Mangasarian. Mathematical programming in data mining. *Data Mining and Knowledge Discovery*, 1(2):183–201, 1997. Available at ftp://ftp.cs.wisc.edu/math-prog/tech-reports/96-05.ps.Z.

[16] O. L. Mangasarian. Solution of general linear complementarity problems via nondifferentiable concave minimization. *Acta Mathematica Vietnamica*, 22(1):199–205, 1997. Available at ftp://ftp.cs.wisc.edu/math-prog/tech-reports/96-10.ps.Z.

[17] O. L. Mangasarian and J.-S. Pang. Exact penalty functions for mathematical programs with linear complementarity constraints. *Optimization*, 42:1–8, 1997. Available from: ftp://ftp.cs.wisc.edu/math-prog/tech-reports/96-06.ps.Z.

[18] B. T. Polyak. *Introduction to Optimization.* Optimization Software, Inc., Publications Division, New York, 1987.

[19] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.

[20] V. N. Vapnik. *The Nature of Statistical Learning Theory.* Springer, New York, 1995.

[21] D. H. Wolpert. On the connection between in-sample testing and generalization error. *Complex Systems*, 6:47–94, 1992.

[22] D. H. Wolpert, editor. *The Mathematics of Generalization*, Reading, MA, 1995. Addison-Wesley.