Annals of Operations Research, 25 (1990) 291–300

ON THE EQUIVALENCE BETWEEN SOME DISCRETE AND CONTINUOUS OPTIMIZATION PROBLEMS *

Fabio TARDELLA

Istituto di Elaborazione della Informazione, Consiglio Nazionale delle Ricerche, Via Santa Maria, 46, 56100 Pisa, Italy

Abstract

The simplex algorithm for linear programming is based on the well-known equivalence between the problem of maximizing a linear function f on a polyhedron P and the problem of maximizing f over the set V_P of all vertices of P. The equivalence between these two problems is also exploited by some methods for maximizing a convex or quasi-convex function on a polyhedron.

In this paper we determine some very general conditions under which the problem of maximizing f over P is equivalent, in some sense, to the problem of maximizing f over V_P. In particular, we show that these two problems are equivalent when f is convex or quasi-convex on all the line segments contained in P and parallel to some edge of P.

In the case where P is a box our results extend a well-known result of Rosenberg for 0–1 problems. Furthermore, when P is a box or a simplex, we determine some classes of functions that can be maximized in polynomial time over P.

1. Introduction

Consider the continuous optimization problem

$$\text{(CP)} \qquad \begin{cases} \max f(x) \\ x \in P \end{cases}$$

and the discrete problem

$$\text{(DP)} \qquad \begin{cases} \max f(x) \\ x \in V_P, \end{cases}$$

where P is a polyhedron in $\mathbb{R}^n$, V_p is the set of vertices of P and f is a real-valued function on P. In some cases problems CP and DP are known to be equivalent, i.e. one or all the solutions of CP can be obtained from the solutions of DP and vice versa. Clearly, when CP and DP are equivalent one could use

* This paper has been partially written while the author was visiting the Rutgers Center for Operations Research (RUTCOR). The support of the Air Force grants AFORS-89-0512 and AFORS-90-0008 is gratefully acknowledged.

both continuous and discrete optimization techniques to solve them. Typical examples of this fact are provided by simplex and interior point methods for linear programming or by the methods employed for maximizing a convex function (or, equivalently, minimizing a concave function) over a polyhedron (see [11] for a survey of these methods).

Let $P_{\max}$ denote the set of global maximum points of f over P. We say that CP and DP are:

- *Strongly equivalent*, if $P_{\max}$ is a face of P.
- *Equivalent*, if $P_{\max}$ is a union of faces of P or, equivalently, if whenever $x \in P_{\max}$, the smallest face of P containing x is contained in $P_{\max}$.
- *Weakly equivalent*, if $P_{\max} \cap V_P \neq \emptyset$, whenever $P_{\max} \neq \emptyset$.

Clearly, strong equivalence implies equivalence which, in turn, implies weak equivalence. Furthermore, it is well known that CP and DP are strongly equivalent, equivalent or weakly equivalent when f is linear, convex or quasi-convex, respectively (see Murty [10], Rockafellar [12] and Greenberg et al. [3]).

Let $v(\mathrm{CP})$ and $v(\mathrm{DP})$ denote the optimal values of CP and DP respectively. It is easy to see that weak equivalence is sufficient to guarantee that $v(\mathrm{CP}) = v(\mathrm{DP})$. Furthermore, when P is an integral polyhedron, i.e. all the vertices of P are integer, weak equivalence between CP and DP trivially implies that at least one optimal solution for the problem of maximizing f over all the integer points in P lies in V_P.

In this paper we establish some very general conditions under which these three types of equivalence hold. Exploiting the equivalence between CP and DP we then deduce the polynomiality of some classes of optimization problems.

Let us introduce some definitions and notations which will be used in the sequel. Given a set $X \subset \mathbb{R}^n$, $\dim(X)$ denotes the dimension of the smallest affine manifold containing X, while $\operatorname{ri} X$, $\operatorname{rbd} X$ and $\operatorname{co} X$ denote the relative interior, the relative boundary and the convex hull of X respectively (see Rockafellar [12]). Let X be a convex set in $\mathbb{R}^n$. A function $f: X \to \mathbb{R}$ is called (strictly) quasi-convex (see Avriel et al. [1]) if for every x^1, $x^2 \in X$ and for every $\alpha \in [0, 1]$

$$f(\alpha x^1 + (1-\alpha)x^2) \leqslant \max\{f(x^1), f(x^2)\}$$

$$\left(\text{resp. } f(\alpha x^1 + (1-\alpha)x^2) < \max\{f(x^1), f(x^2)\}\right).$$

Similarly, a function f is called (strictly) convex if for every x^1, $x^2 \in X$ and for every $\alpha \in [0, 1]$

$$f(\alpha x^1 + (1-\alpha)x^2) \leqslant \alpha f(x^1) + (1-\alpha) f(x^2)$$

$$\left(\text{resp. } f(\alpha x^1 + (1-\alpha)x^2) < \alpha f(x^1) + (1-\alpha) f(x^2)\right).$$

If $-f$ is (strictly) convex or (strictly) quasi-convex f is called (strictly) concave or (strictly) quasi-concave respectively.

2. The maximum principle

It is well-known from Harmonic Analysis (see [9]) that all global maximum points of a subharmonic function f on a domain D must lie on the boundary of D unless f is constant. This property, called the maximum principle, is enjoyed also by other classes of functions. In this section we will determine some classes of functions satisfying the maximum principle on some or all the faces of a polyhedron.

DEFINITION 2.1

Let X be a subset of $\mathbb{R}^n$ and let f be a function from X into $\mathbb{R}$. We say that f satisfies the *Maximum Principle* (MP for short) on X iff, whenever $x^* \in X$ and $f(x) \leqslant f(x^*)$ for every x in X, either $f(x) = f(x^*)$ for every x in X or $x^* \in \operatorname{rbd} X$. We say that f satisfies the *Weak Maximum Principle* (WMP for short) on X iff, whenever $f(x) \leqslant f(x^*)$ for every x in X, there exist $x' \in \operatorname{rbd} X$ such that $f(x') = f(x^*)$.

Note that convex functions satisfy MP on every convex set, whereas quasi-convex functions satisfy WMP on every convex compact set.

Let P be a polyhedron in $\mathbb{R}^n$ and let $s^1, \ldots, s^q$ be a set of vectors in $\mathbb{R}^n$ such that every edge of P is parallel to some s^i. Furthermore, for every $x \in P$ and for every $s \in \mathbb{R}^n$ define

$$P_s(x) = \{ z \in P : z = x + \lambda s,\ \lambda \in \mathbb{R} \}.$$

Given any face F of P define also

$$I_F = \{ i : s^i \text{ is parallel to some edge of } F \}.$$

PROPOSITION 2.1

If $f: X \to \mathbb{R}$ is convex on $P_{s^h}(x)$, for every $h = 1, \ldots, q$ and for every $x \in P$, then f satisfies MP on every face of P.

Proof

Let F be a face of P and assume that there is a point $x \in \operatorname{ri} F$ that globally maximizes f over F. Given any point $y \in F$, we can then find a finite sequence $\{x^i\}_{0 \leqslant i \leqslant r}$ such that

$$x^0 = x,\ x^r = y,\ x^i \in \operatorname{ri} P_{s^{h_i}}(x^i),\ x^{i+1} \in P_{s^{h_i}}(x^i),$$
$$h_i \in \{1, \ldots, q\},\ i = 0, \ldots, r-1.$$

Observe that $x^0 \in \operatorname{ri} P_{s^{h_0}}(x^0)$ is a global maximum point for f over $P_{s^{h_0}}(x^0)$. Then, since f is convex over $P_{s^{h_0}}(x^0)$, f must be constant on $P_{s^{h_0}}(x^0)$. Hence, $f(x^0) = f(x^1)$ and x^1 is also a global maximum point for f over F. Proceeding by induction we then get

$$f(x) = f(x^0) = f(x^1) = \cdots = f(x^r) = f(y).$$

Hence f is constant over F.

Remark 2.1

Note that when f is twice continuously differentiable, the assumption of the previous proposition is equivalent to

$$(s^h)^T H_f(x) s^h \geqslant 0, \quad \forall x \in P \text{ and } \forall h = 1, \ldots, q, \tag{1}$$

where $H_f(x)$ is the Hessian matrix of f at x.

PROPOSITION 2.2

Let $f: P \to \mathbb{R}$, where P is a bounded polyhedron. If for every $x \in P$ there exists $y \in P$, $y \neq x$, such that, setting $s = y - x$, f is (strictly) quasi-convex on $P_s(x)$, then f satisfies WMP (MP) on P.

Proof

Let us assume that there exists a global maximum point $x^0 \in \operatorname{ri} P$ for f over P. Let $y \in P$ be such that f is (strictly) quasi-convex on $P_s(x^0)$, where $s = y - x$. Note that the extreme points x^1 and x^2 of $P_s(x^0)$ belong to $\operatorname{rbd} P$. Furthermore, by the quasi-convexity of f we have

$$f(x^0) \leqslant \max\{f(x^1), f(x^2)\}.$$

Hence, either x^1 or x^2 is a global maximum point for f over P. If f is strictly quasi-convex on $P_s(x^0)$, we have

$$f(x^0) < \max\{f(x^1), f(x^2)\},$$

contradicting the global maximality of x^0.

The following corollary is a trivial consequence of proposition 2.2.

COROLLARY 2.2.1

Let $f: P \to \mathbb{R}$, where P is a bounded polyhedron. Assume that f is (strictly) quasi-convex on $P_{s^h}(x)$, $\forall x \in P$ and $\forall h \in H \subset \{1, \ldots, q\}$. Then f satisfies WMP (MP) on all the faces F of P such that $I_F \cap H \neq \emptyset$.

COROLLARY 2.2.2

Let $f: P \to \mathbb{R}$, where P is a bounded polyhedron with p edges. Assume that f is (strictly) quasi-convex on $P_{s^h}(x)$, $\forall x \in P$ and $\forall h \in H \subset \{1, \ldots, q\}$. For every $h \in H$ denote d_h the number of edges of P parallel to s^h and let $d = \sum_{h \in H} d_h$. Then f satisfies WMP (MP) on all faces of dimension m such that

$$m(m+1)/2 > p - d. \tag{2}$$

Proof

The number of edges of an m-dimensional face F of P is not smaller than $m(m+1)/2$, the number of edges of an m-dimensional simplex (see Brondsted

[2]). Hence, from (2) we deduce that there exists an edge of F parallel to s^h, for some $h \in H$. The conclusion then follows from corollary 2.2.1.

Remark 2.2

Note that the conclusions of proposition 2.2 and corollary 2.2.1 remain valid also when P is unbounded but has at least one vertex, if the (strict) quasi-convexity assumption is replaced by (strict) convexity. Indeed, in this case, in the proof of proposition 2.2 $P_s(x^0)$ may be unbounded on one side. If this is the case the proof may be modified as follows. Consider x^1, the only extreme point of $P_s(x^0)$ and take any point $x^2 \in P_s(x^0)$ such that $x^0 = \alpha x^1 + (1-\alpha)x^2$ for some $\alpha \in [0, 1]$. If f is convex we have

$$f(x^0) \leqslant \alpha f(x^1) + (1-\alpha)f(x^2) \leqslant \alpha f(x^0) + (1-\alpha)f(x^0) = f(x^0)$$

and hence

$$f(x^0) = f(x^1) = f(x^2).$$

Since $x^1 \in \operatorname{rbd} P$, this concludes the proof of this case. If f is strictly convex we have the contradiction

$$f(x^0) < \alpha f(x^1) + (1-\alpha)f(x^2) \leqslant f(x^0).$$

The previous results require the knowledge of some or all edges of P. When the edges of P are not explicitly known we can still guarantee the validity of MP or WMP by making more restrictive assumptions on f.

PROPOSITION 2.3

Let $f: P \to \mathbb{R}$ be twice continuously differentiable. If at any $x \in P$ the Hessian matrix $H_f(x)$ of f has at least m positive eigenvalues, then f satisfies MP on every face F of P such that $\dim(F) > n - m$.

Proof

Let F be a face of P with $\dim(F) > n - m$. Then, for any point $x \in \operatorname{ri} F$ it is possible to find $y \in F$ such that $(y-x)^{\mathrm{T}} H_f(x)(y-x) > 0$. Hence, no point in the relative interior of F can satisfy the second order necessary conditions for maximality. Therefore f satisfies MP on F.

DEFINITION 2.2

Let f be a function from a convex set $X \subset \mathbb{R}^n$ into $\mathbb{R}$. We say that f is m-convex (m-quasi-convex) over X, for $1 \leqslant m \leqslant n$, if

$$f\big(\alpha x^1 + (1-\alpha)x^2\big) \leqslant \alpha f(x^1) + (1-\alpha)f(x^2)$$

$$\big(\text{resp. } f\big(\alpha x^1 + (1-\alpha)x^2\big) \leqslant \max\{f(x^1), f(x^2)\}\big),$$

for every $\alpha \in [0, 1]$ and for every x^1, $x^2 \in X$ such that $x_i = y_i$ for at least $n - m$ indices i in $\{1, \ldots, n\}$.

Note that n-(quasi-)convexity coincides with ordinary (quasi-)convexity. Furthermore, if f is m-(quasi-)convex, then it is also m'-(quasi-)convex, for every $m' < m$. The converse is not true in general: the function $f(x, y) = x^2 + y^2 - 4xy$ is 1-convex but is neither 2-convex nor 2-quasi-convex.

It is a straightforward consequence of the second order characterization of convexity that a twice continuously differentiable function f on a convex open set Ω is m-convex over Ω if and only if all the m-dimensional principal submatrices of $H_f(x)$ are positive semidefinite for every x in Ω. In particular, f is 1-convex if and only if all the diagonal elements of $H_f(x)$ are nonnegative for every x in Ω. Note also that 1-convexity may be viewed as convexity with respect to each variable separately.

Let us now assume that the polyhedron P is expressed in one of the following ways:

$$P = \{x \in \mathbb{R}^n : Ax = b,\ \alpha \leqslant x \leqslant \beta\} \tag{3}$$

or

$$P = \{x \in \mathbb{R}^n : Ax \leqslant b,\ \alpha \leqslant x \leqslant \beta\}, \tag{4}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, α, $\beta \in \mathbb{R}^n$, $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\beta = (\beta_1, \ldots, \beta_n)$, $-\infty \leqslant \alpha_i < \beta_i \leqslant +\infty$ and $\min\{|\alpha_i|, |\beta_i|\} < +\infty$. Note that the assumptions on α and β imply that, if $P \neq \emptyset$, P has at least one vertex.

PROPOSITION 2.4

Let P be a polyhedron defined by (3) or (4) and assume that A has rank $m-1$. If f is m-(quasi-)convex over P, then f is (quasi-)convex over $P_{s^h}(x)$ for every $h = 1, \ldots, q$ and for every $x \in P$.

Proof

Let $s = (s_1, \ldots, s_n)$ be a vector parallel to an edge of P. Observe that every edge of P lies in the intersection of $n-1$ linearly independent hyperplanes taken from among those defining P in (3) or (4). Since $\operatorname{rank}(A) = m-1$, we have $s_i = 0$ for at least $n-m$ indices i. From m-(quasi-)convexity of f we then derive that f is (quasi-)convex over the sets $P_s(x)$ for every $x \in P$.

The following corollaries are a straightforward consequence of propositions 2.1 and 2.4.

COROLLARY 2.4.1

Under the assumptions of proposition 2.4, f satisfies MP (WMP) on all the faces of P.

COROLLARY 2.4.2

If $P = \{x \in \mathbb{R}^n : \alpha \leqslant x \leqslant \beta\}$ and f is 1-(quasi-)convex over P, then f satisfies MP (WMP) on all the faces of P.

3. The set of global maximum points of a function over a polyhedron

In this section we present some rather general conditions for the equivalence between problems CP and DP. Let Ω denote a set of faces of a polyhedron P and let Ω^c be the set of all faces of P that are not contained in Ω.

PROPOSITION 3.1

Let f be a real function on P and assume that f satisfies WMP for every face $F \in \Omega$. Then the set $P_{\max}$ of global maximum points of f over P satisfies the following relation

$$P_{\max} \cap \left(V_P \cup \bigcup_{F \in \Omega^c} F \right) \neq \emptyset, \tag{5}$$

when $P_{\max} \neq \emptyset$. Furthermore, if f satisfies MP for every $F \in \Omega$, then there exists $\Omega' \subset \Omega$ such that

$$\bigcup_{F \in \Omega'} F \subset P_{\max} \subset \bigcup_{F \in \Omega' \cup \Omega^c} F. \tag{6}$$

Proof

Assume that f satisfies WMP for every $F \in \Omega$ and that $P_{\max} \neq \emptyset$. Let F^* denote a face of P of minimal dimension among those satisfying the relation $P_{\max} \cap F \neq \emptyset$. If $F^* \in \Omega^c$ or $\dim(F^*) = 0$, then (5) holds. Hence assume that $F^* \in \Omega$ and $\dim(F^*) > 0$ and let $x \in P_{\max} \cap F^*$. By the WMP one can then find $y \in \operatorname{rbd} F^*$ such that $y \in P_{\max}$, contradicting the minimality assumption on F^*. Assume now that f satisfies MP for every $F \in \Omega$ and, for every $x \in P$, denote F_x the smallest face of P containing x. Note that, by the MP, if $x \in P_{\max}$ we have $F_x \subset P_{\max}$. Then, setting $\Omega' = \{ F \in \Omega : F = F_x,\ x \in P_{\max} \}$, the first inclusion of (6) holds. The second inclusion is a trivial consequence of the definition of Ω' and of the following one

$$P_{\max} \subset \bigcup_{x \in P_{\max}} F_x.$$

COROLLARY 3.1.1

If f satisfies MP (WMP) on all the faces of P, then problems CP and DP are (weakly) equivalent.

COROLLARY 3.1.2

If f is quasi-concave on P and satisfies MP on all the faces of P, then problems CP and DP are strongly equivalent.

Proof

By proposition 3.1, $P_{\max}$ is a union of faces of P. On the other hand, from the quasi-concavity of f it follows that $P_{\max}$ is convex (see [1]). Hence $P_{\max}$ must be a face of P.

The following straightforward consequence of proposition 3.1 extends a result of Hager et al. [5].

COROLLARY 3.1.3

If f satisfies WMP on all the faces of P of dimension greater than m, then at least one point of $P_{\max}$ lies in an m-dimensional face of P.

4. The box case

In this section we will specialize the previous results to the case where P is a box, while in the next section we shall deal with the case where P is a simplex. Let us then assume that P is a box, i.e. P is defined by (3) with $A = 0$, $b = 0$ and $-\infty < \alpha_i < \beta_i < +\infty$ for $i = 1, \dots, n$. In this case, it is easily seen that a set of vectors parallel to all the edges of P is $e^1, \dots, e^n$, the canonical basis of $\mathbb{R}^n$. Furthermore, by making the simple change of variables

$$y = D(x - \alpha),$$

where $D = \operatorname{diag}((\beta_1 - \alpha_1)^{-1}, \dots, (\beta_n - \alpha_n)^{-1})$, P may be transformed into the 0–1 hypercube $Q = [0, 1]^n$ and hence V_P coincides with the 0–1 discrete hypercube $B = \{0, 1\}^n$. In the sequel we will then assume, without loss of generality, that $P = [0, 1]^n$. Note that in this case problem DP is an unconstrained 0–1 program. Hence, when CP and DP are equivalent in some sense, the methods of pseudo-Boolean programming (see [6–8]) may be employed to solve the continuous problem CP and, conversely, the methods of continuous optimization can be used to solve DP. Rosenberg [13] has shown that CP and DP are equivalent, in the box case, when f is a polynomial which is linear with respect to each variable. This result is generalized in the next proposition, which is a straightforward consequence of corollaries 2.4.2, 3.1.1 and 3.1.2.

PROPOSITION 4.1

If f is 1-(quasi-)convex on P, then problems CP and DP are (weakly) equivalent. Furthermore, if f is 1-convex and quasi-concave on P, problems CP and DP are strongly equivalent.

A well-known result in 0–1 optimization is the fact that the problem of maximizing a supermodular function over the discrete 0–1 hypercube can be solved in polynomial time (see Groetschel et al. [4]). In view of the above remarks, this result can be extended to the continuous case as follows.

PROPOSITION 4.2

If f is 1-quasi-convex on P and the restriction of f to V_P is supermodular then problem CP can be solved in polynomial time.

Remark 4.1

The assumptions of proposition 4.2 are satisfied, e.g., when f is twice continuously differentiable and the second order derivatives $f_{ij}(x)$ are nonnegative for every i, $j=1,\dots,n$ and for every $x \in P$. Indeed, in this case f is clearly 1-convex. Furthermore, Topkis [14] has shown that the nonnegativity of the mixed partial derivatives $f_{ij}(x)$ for $i \neq j$ implies the supermodularity of f on P and hence, a fortiori, on V_P.

Remark 4.2

In view of the remarks of section 1, if α and β are integer and the assumptions of proposition 4.2 hold, then at least one optimal solution for the integer program

$$(\mathrm{IP}) \qquad \begin{cases} \max f(x) \\ \alpha \leqslant x \leqslant \beta, \quad x \text{ integer} \end{cases}$$

must be attained at a vertex. Hence problem IP can also be solved in polynomial time.

5. The simplex case

In this section we discuss the equivalence of CP and DP in the case where P is expressed in the following way

$$P = \{x \in \mathbb{R}^n : x \geqslant 0,\ a^{\mathrm{T}}x \leqslant b\}. \tag{7}$$

Note that in this case the cardinality of V_P is not greater than n. Hence problem DP is trivial. Therefore, when CP and DP are equivalent, problem CP can also be solved immediately. Note also that if $b > 0$ and $a_i > 0$ for $i = 1,\dots,n$, P is a simplex. It is easy to check that a set of vectors parallel to all edges of P is $S = \{s^{ij}\}_{i,j=1,\dots,n}$, where $s^{ij} = a_i e^i - a_j e^j$, if $i \neq j$, and $s^{ii} = e^i$, for $i = 1,\dots,n$.

PROPOSITION 5.1

Let f be twice continuously differentiable and assume that

(i) $f_{ii}(x) \geqslant 0, \quad \forall x \in P,\ i,\ j = 1,\dots,n$;
(ii) $a_i^2 f_{ii}(x) + a_j^2 f_{jj}(x) - 2a_i a_j f_{ij}(x) \geqslant 0, \quad \forall x \in P,\ i,\ j = 1,\dots,n,\ i \neq j.$

Then CP and DP are equivalent.

Proof

Note that conditions (i) and (ii) are equivalent to $(s^{ij})^{\mathrm{T}} H_f(x)(s^{ij}) \geqslant 0$, for every $x \in P$. Hence, the proof follows from remark 2.1, proposition 2.1 and corollary 3.1.1.

References

[1] M. Avriel, W.E. Diewert, S. Schaible and I. Zang, *Generalized concavity* (Plenum, New York, 1988).
[2] A. Brondsted, *An Introduction to Convex Polytopes* (Springer, Berlin, 1983).
[3] H.J. Greenberg and W.P. Pierskalla, A review of quasi-convex functions, Oper. Res. 19 (1971) 1553–1570.
[4] M. Groetschel, L. Lovasz and A. Schrijver, The ellipsoid method and its consequences in combinatorial optimization, Combinatorica 1 (1981) 169–197.
[5] W.W. Hager, P.M. Pardalos, I.M. Roussos and H.D. Sahinoglou, Active constraints in optimization, Technical report, Penn State University, Computer Science Department (1988) to appear in JOTA.
[6] P.L. Hammer and S. Rudeanu, *Boolean Methods in Operations Research and Related Areas* (Springer, Berlin, 1968).
[7] P. Hansen, Methods of nonlinear 0–1 programming, Ann. Discrete Math. 5 (1979) 53–70.
[8] P. Hansen, B. Jaumard and V. Mathon, Constrained nonlinear 0–1 programming, Rutgers Research Report 47–89 (1989).
[9] W.K. Hayman and P.B. Kennedy, *Subharmonic Functions*, vol. 1 (Academic Press, London, 1976).
[10] K.G. Murty, *Linear Programming* (Wiley, New York, 1983).
[11] P.M. Pardalos and J.B. Rosen, Methods for global concave minimization: a bibliografic survey, SIAM Rev. 18 (1986) 367–379.
[12] R.T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970).
[13] I. Rosenberg, 0–1 optimization and nonlinear programming, Rev. Française Inform. Recherche Operationelle 6 (1972) 95–97.
[14] D.M. Topkis, Minimizing a submodular function on a lattice, Oper. Res. 26 (1978) 305–321.