Application of Fenchel's Duality Theorem to Penalty Methods in Convex Programming

By

Hisashi MINE* and Masao FUKUSHIMA*

(Received September 30, 1977)

Abstract

This paper studies a new class of sequential unconstrained optimization methods, called the conjugate penalty method, for solving convex programming problems. The validity of the method is based on Fenchel's duality theorem. It is shown that, under certain conditions, conjugate penalty founctins are uniformly bounded on a neighborhood of a point which is an optimum of Fenchel's dual problem.

1. Introduction

Recently, authors have reported a number of methods for solving nonlinear programming problems by transforming each constrained optimization problem into unconstrained optimization problems[2,4,5,7,14]. A characteristic underlying those methods is that a solution of the original problem can be obtained as a limit of sequential solutions to transformed unconstrained problems. Among those methods, the sequential unconstrained minimization techniques,[4] commonly abbreviated to SUMT, have been used in practice. They, sometimes called penalty methods, reduce the computational process to unconstrained minimization of a transformed function, called a penalty function, combining the objective function, the constraint functions and one or more praameters.

In this paper, we study a sequential unconstrained optimization method, proposed first in[9], for solving convex programming problems. The method can be regarded as dual to the ordinary penalty methods. The theoretical validity of the method is based upon the well-known Fenchel's duality theorem[3,6,11,12] in the theory of convex analysis. Since it utilizes conjugate convex and concave functions, we call it the conjugate penalty method.[9]

* Department of Applied Mathematics and Physics

Computational difficulties, such as ill-conditioning of matrices involved in ordinary penalty function methods, are caused by the fact that optimal solutions of the penalty functions lie in steep-sided valleys as the sequential minimizations proceed. Those circumstances seem unavoidable as long as ordinary penalty functions are concerned. On the other hand, one can exclude such unfavorable circumstances by employing conjugate penalty functions when some conditions on the problem are satisfied. Namely, the behavior of the conjugate penalty functions near the solution is expected to be mild as sequential maximizations proceed.

In section 2, we show a duality between two extremum problems derived from a general convex programming problem. In section 3, we define the conjugate penalty method and prove convergence. In section 4, several conditions are given, under which the conjugate penalty functions are well-behaved.

In the remainder of this section, we summarize some notations in convex analysis that will be needed in the subsequent sections. The readers might refer to Rockafellar[12] for the details. Some of them may also be found elsewhere.[6,13,15]

For a convex set C in R^n, we denote the interior and the relative interior of C by int C and ri C, respectively. Throughout this paper all functions are understood to be extended-real-valued. Let f be a convex (or concave) function on R^n. The epigraph and the effective domain of f are denoted by epi f and dom f, respectively, and the subdifferential of f at x, which is the set of all subgradients of f at x, is denoted by $\partial f(x)$. In particular, if functions are differentiable in the ordinary sense, then the subgradients reduce, of course, to the gradients, and the usual notation $\nabla f(x)$ is used.

Asymptotic properties of convex sets and convex (concave) functions are particularly important in the development of this paper. The notion of 'recession' will be useful in formulating various growth conditions that specify some behavior of sets and functions at infinity. A recession cone of a convex set C, denoted by 0^+C, is the set consisting of the zero vector and all directions of recession of C. The recession function of a convex (concave) function f is defined as a function whose epigraph is the recession cone of epi f in R^{n+1}. We denote it as $f0^+$. The set of all vectors y, such that $(f0^+)(y)\leqq$ $(\geqq)0$ for convex (concave) f, is called the recession cone of f, and such vectors are called directions of recession of f.

Although we make no distinction about symbols, e.g.*, ∂, 0^+ etc., for convex functions nor for concave functions, the meanings will be clear from the context.

A fundamental and beautiful duality theorem proved by Fenchel[3] is one of the splendid results about the theory of convex analysis. This theorem plays a central role in the development of this paper.

Fenchel's Duality Theorem Let f and g be a convex function and a concave function on R^n, respectively. If ri (dom f) $\cap$ ri (dom g) $\neq \phi$, then

$$\inf_{x \in R^n} \{f(x) - g(x)\} = \sup_{y \in R^n} \{g^*(y) - f^*(y)\},$$

where the supremum is attained at some y.

2. A general convex program and Fenchel duality

Consider the following general convex programming problem:

(P) minimize $f(x)$ over $x \in C$,

where f is a closed convex function on R^n, and C is a non-empty closed convex set in R^n. The convex programming problem (P) is equivalent to the following 'unconstrained' problem:

(1) minimize $U(x) \triangleq f(x) - \gamma_C(x)$ over $x \in R^n$,

where γ_C is the indicator function of C defined by

$$\gamma_C(x) = 0 \text{ if } x \in C, = -\infty \quad \text{if } x \notin C.^\dagger$$

Obviously, γ_C is a closed concave function on R^n.

Throughout this paper, we assume the following:

A–1. The finite minimum of U is uniquely attained at $\bar{x}$. Namely $\bar{x}$ is the unique minimum of problem (P);

A–2. ri(dom f) and ri C have a point in common.

In order to guarantee the existence of a minimum (possibly not unique), we may suppose that f and C have no direction of recession in common [12, Th. 27.3]. Some other conditions for the existence of minima are found in [1]. The latter assumption is automatically satisfied when f is finite everywhere, i.e. dom $f = R^n$, and ri C is nonempty, as is almost the case in practical problems.

Let V be a closed concave function defined by

(2) $V(y) \triangleq \gamma_C^*(y) - f^*(y)$,

where f^* and γ_C^* are conjugates of f and γ_C, respectively. The function γ_C^* is the negative of the support function of C [12, p. 28] and, hence, a positively homogeneous closed concave function.

The following lemma is derived from Fenchel's duality theorem.

Lemma 1. Let U and V be defined by (1) and (2), respectively. If assumptions A–1 and A–2 are satisfied, then

$$f(\bar{x}) = \min_{x \in R^n} U(x) = \sup_{y \in R^n} V(y),$$

† γ_C is the negative of the indicator function in [12, p. 28].

where the supremum is attained. Moreover, the maximum set of V is $\partial f(\bar{x}) \cap \partial\gamma c(\bar{x})$, and conversely, for each maximum $\bar{y}$ of V, the set $\partial f^*(\bar{y}) \cap \partial\gamma c^*(\bar{y})$ is a singleton $\{\bar{x}\}$.

Proof. The first part is immediate from Fenchel's duality theorem with $g(x)=\gamma c(x)$. Now prove the latter half. We can show that the following conditions are equivalent;

(i) $f(\bar{x})-\gamma c(\bar{x})=\gamma c^*(\bar{y})-f^*(\bar{y})$;

(ii) $f(\bar{x})+f^*(\bar{y})=\langle\bar{x},\bar{y}\rangle=\gamma c(\bar{x})+\gamma c^*(\bar{y})$;

(iii) $\bar{y}\in\partial f(\bar{x}) \cap \partial\gamma c(\bar{x})$;

(iv) $\bar{x}\in\partial f^*(\bar{y}) \cap \partial\gamma c^*(\bar{y})$. (see [12, Th. 23.5])

Let x' be an arbitrary point in $\partial f^*(\bar{y}) \cap \partial\gamma c^*(\bar{y})$, then obviously $f(x')-\gamma c\ (x')=f(\bar{x})-\gamma c(\bar{x})$. Thus, $\partial f^*(\bar{y}) \cap \partial\gamma c^*(\bar{y})$ must be a singleton $\{\bar{x}\}$ by the uniqueness.

The minimum set $\partial f(\bar{x}) \cap \partial\gamma c(\bar{x})$ of V is clearly a closed convex set. In particular, when f is difierentiable, the maximum set of V is fairly simplified by the following.

Lemma 2. Let all requirements in Lemma 1 be satisfied. In addition, if f is differentiable at $\bar{x}$, then the supremum of V is uniquely attained at $\bar{y}=\nabla f(\bar{x})$.

Proof. Since $\partial f(\bar{x})=\{\nabla f(\bar{x})\}$ and the maximum set of V is non-empty by Lemma 1, it is necessarily $\{\nabla f(\bar{x})\}$.

3. A conjugate penalty method

There are a number of methods that solve problem (P) by transforming it into a sequence of unconstrained problems of the form:

(P_k) minimize $U_k(x) \triangleq f(x)-h_k(x)$ over $x\in R^n$.

For $k=1, 2, \ldots$, each auxiliary problem (P_k) is solved and the optimal solution to problem (P) is obtained as a limit of a sequence of the optimal solutions to problems (P_k). According to the types of h_k, functions U_k are classified into several classes, e.g. barrier functions, loss functions, etc.[4,5,7,14] Here, we call those functions generically **penalty functions.** The peanlty functions U_k should be constructed so that:

(i) for every k, there exists a (unique) x_k that minimizes U_k over R^n;

(ii) x_k converges to $\bar{x}$ as $k\to\infty$, and the limit of $U_k(x_k)$ is the minimum value of (P).

In the following, let appropriate conditions be implicitly assumed so that the properties (i) and (ii) above are fulfilled. Such conditions can be found, for example, in [8]. Moreover we assume the following:

A–3. Each h_k is a closed concave function with ri(dom f) $\cap$ ri(dom h_k) $\neq \phi$;

A–4. int(dom f) $\neq \phi$ and f is differentiable on int(dom f);

A–5. Every x_k, a minimum of U_k, belongs to int(dom f).

Define the **conjugate penalty functions** V_k on R^n by

(3) $$V_k(y) \triangleq h_k^*(y) - f^*(y),$$

where f^* and h_k^* are the conjugates of f and h_k, respectively. Now consider a sequences of problems;

(Q_k) maximize $V_k(y)$ over $y \in R^n$.

Since f^* and h_k^* are convex and concave respectively, each problem (Q_k) is to find an unconstrained maximum of the concave function V_k.

It is now possible to define a conjugate penalty method for solving problem (P) in a manner quite similar to that in ordinary penalty methods. Specifically, we try to solve problem (P) by successive maximizations of the conjugate penalty functions V_k, $k=1, 2, \ldots\ldots$ Therefore, the method may be regarded as one of the SUMT procedures.

The following theorem proves convergence of the conjugate penalty method.

Theorem 1. Let V_k be defined by (3). If assumptions A–1 through A–5 are satisfied, then there exists for every k a unique maximum y_k of V_k and $y_k = \nabla f(x_k)$, where x_k is a minimum of problem (P_k). Moreover the y_k and $V_k(y_k)$ converge to $\nabla f(\bar{x})$ and $f(\bar{x})$, respectively, as $k \to \infty$.

Proof. By A–3, the existence of a maximum y_k of V_k follows from Fenchel's duality theorem. By the differentiability of f, it follows from a similar argument in Lemma 2 that the y_k is unique and is equal to $\nabla f(x_k)$. As x_k converges to $\bar{x}$ and the mapping ∇f is continuous on int(dom f), y_k also converges to $\nabla f(\bar{x})$. Since $U_k(x_k)$ converges to $f(\bar{x})$, the convergence of $V_k(y_k)$ to $f(\bar{x})$ follows immediately from the relation

$$U_k(x_k) = \min_x U_k(x) = \max_y V_k(y) = V_k(y_k).$$

It may be remarked that the conjugate penalty functions are really defined on the dual space of R^n, which is identified with R^n. Thus for optimization problems in more general spaces, it may be possible to consider conjugate penalty functions on the dual spaces.

4. Advantage of the conjugate penalty method

Difficulties in computing ordinary penalty functions result mainly from the fact that the penalty function U_k grows extremely steep-valleyed near the minimum of the problem as k increases.[14] Since U_k should converge in a certain sense to the function U, the reason for such irregularity may be that the minimum $\bar{x}$ generally lies on the boundary of dom U. A convergence property of ordinary penalty functions for general convex programs is studied in [8].

As regards the function V, however, the maxima of V may be in int(dom f) even when the minimum of U is on the boundary of dom U. In fact, this is true for a certain class of problems. In those problems the conjugate penalty functions V_k are expected to be uniformly bounded on a neighborhood of the maxima of V. Therefore, we may bypass the difficulty inherent to ordinary penalty methods by employing the conjugate penalty functions V_k to solve problem (P).

In this section, we study conditions on problem (P) for the maxima of V interior to dom V. The necessary and sufficient condition for $\bar{y} \in$ int(dom V) is stated in the following

Theorem 2. Assume that assumptions A-1 and A-2 are satisfied and that f is differentiable at $\bar{x}$. Then, $\bar{y} \in$ int(dom V), if and only if the following two conditions are simultaneously satisfied:

(a) $\qquad (f0^+)(x) > \langle \nabla f(\bar{x}), x \rangle \qquad$ for every $x \neq 0$;

(b) $\qquad \langle \nabla f(\bar{x}), x \rangle > 0 \qquad$ for every $x \in 0^+C$ and $x \neq 0$;

where $\bar{x}$ is the minimum of U and $\bar{y}$ is the maximum of V.

Proof. First, note that $\bar{y} \in$ int(dom V) if and only if $\bar{y} \in$ int(dom f^*) and $\bar{y} \in$ int(dom γc^*) simultaneously. It follows from [12, Cor. 13.3.4.(c)] that $\bar{y} \in$ int (dom γc^*) if and only if $(f0^+)(x) - \langle \bar{y}, x \rangle > 0$ for every $x \neq 0$. This is exactly the condition (a) since $\bar{y} = \nabla f(\bar{x})$ by Lemma 2. On the other hand, taking account of the concavity of γc, the necessary and sufficient condition for $\bar{y} \in$ int(dom γc^*) is

$$(\gamma c0^+)(x) - \langle \bar{y}, x \rangle < 0 \text{ for every } x \neq 0,$$

which reduces to

$$\langle \bar{y}, x \rangle > 0 \text{ for every nonzero } x \in 0^+C,$$

because $(\gamma c0^+)(x) = 0$ when $x \in 0^+C$, $= -\infty$ otherwise.

Condition (a) in Theorem 2 can be geometrically interpreted as follows; The hyperplane $z = \langle \nabla f(\bar{x}), x - \bar{x} \rangle + f(\bar{x})$ in R^{n+1} supports epi f at $\bar{x}$ but the set of points at which the hyperplane contacts with epi f is bounded. On the other hand, condition (b) says that either C is compact, i.e. $0^+C = \{0\}$, or there is no halfline orthogonal to $\nabla f(\bar{x})$, which emanates from $\bar{x}$ and is contained in C.

In general, the negative of the polar of the convex cone generated by dom γc^* is the recession cone of γc. Dually, the negative of the polar of the recession cone of γc is the closure of the convex cone generated by dom γc^*[12, Th. 14.2]. While the recession cone of γc is 0^+C and the convex cone generated by dom γc^* is dom γc^* itself, because γc^* is a positively homogeneous closed concave function. Consequently, dom γc^* is the negative of the polar of the cone 0^+C.

In particular, if V is finite everywhere, i.e. dom $V=R^n$, then of course $\bar{y} \in$ int(dom V) holds. The following theorem states a necessary and sufficient condition for V everywhere finite.

Theorem 3. V is finite everywhere if and only if the following conditions are simultaneously satisfied:

(a) f is co-finite, i.e. $(f0^+)(x)=+\infty$ for every $x \neq 0$;

(b) C is compact, i.e. $0^+C=\{0\}$.

Proof. Note that V is finite everywhere if and only if both f^* and γc^* are finite everywhere. From [12, Cors. 13.3.1 and 13.3.2], the theorem follows immediately. The co-finiteness of f implies that epi f contains no non-vertical halflines. This condition is satisfied, of course, if dom f is compact.

Inequality constraints Now consider the problem

(P′) minimize $f(x)$
subject to $g_i(x) \geqq 0,\ i \in I$,

where f and $-g_i$, $i \in I$, are closed convex functions on R^n, and I is an arbitrary index set. The convex programming problem (P′) frequently encountered in practice is a typical case of the (abstract) program (P), in which the constraint set C is specified by the system of inequalities $C=\{x \in R^n;\ g_i(x) \geqq 0,\ i \in I\}$. Note that C may be rewritten as $C=\{x \in R^n;\ g(x) \geqq 0\}$, where g is a closed concave function defined by $g(x)=\inf_{i \in I} g_i(x)$.

In order to state the results obtained in the earlier part of this section, we should represent the function γc^* in terms of the constraint functions g_i. By virtue of [12, Ths. 13.5 and 16.5], γc^* is the closure of the positively homogeneous concave function generated by g^* or $\mathrm{cl}(\mathrm{conv}_{i \in I}\, g_i^*)$, where cl of a function is a function whose epigraph is the closure of the function, and conv of functions is a function whose epigraph is the convex hull of the functions. Namely,

$$\gamma c^*(y)=\mathrm{cl}\,\{\sup_{\lambda \geqq 0} g^*\lambda\}(y)$$
$$=\mathrm{cl}\,\{\sup_{\lambda \geqq 0} \mathrm{cl}(\mathrm{conv}_{i \in I}\, g_i^*\lambda)\}(y).$$

However, such expressions are somewhat complicated and impractical.

In the following, we derive a simple sufficient condition on f and g_i that assures condition (b) in Theorem 2. Then such a condition, if it exists, together with condition (a) in Theorem 2 will imply $\bar{y} \in$ int(dom V).

For simplicity, we make the following assumption on (P′):

A–6. $I=\{1, 2, \ldots, m\}$ and $f, g_1, \ldots, g_m$ are differentiable at $\bar{x}$, the unique minimum of (P′).

Then it is easy to see that $0^+C=\bigcap_{i=1}^{m} 0^+C_i$, since $C=\bigcap_{i=1}^{m} C_i$, where $C_i=\{x;\, g_i(x)\geqq 0\}$. In general, for any concave function g, the direction of recession z satisfies the inequality

$$\langle \nabla g(x),\, z\rangle \geqq 0$$

for every x at which g is differentiable [10, p. 383]. Therefore, since each constraint function g_i is differentiable at $\bar{x}$, $x\in 0^+C=\bigcap_{i=1}^{m} 0^+C_i$ implies

$$\langle \nabla g_i(\bar{x}), x\rangle \geqq 0, \qquad i=1, \ldots, m.$$

In other words, 0^+C is contained by the negative of the polar of the cone generated by $\{\nabla g_1(\bar{x}), \ldots, \nabla g_m(\bar{x})\}$. From this, we have a sufficient condition that guarantees condition (b) in Theorem 2.

Theorem 4. Suppose assumption A–6 is satisfied. If every non-zero vector x such that

$$\langle \nabla g_i(\bar{x}),\, x\rangle \geqq 0, \qquad i=1,\, \ldots,\, m,$$

satisfies the inequality

$$\langle \nabla f(\bar{x}),\, x\rangle > 0,$$

then the same inequality holds for any non-zero vector $x\in 0^+C$.

Proof. Immediate from

$$0^+C\subset \{x;\, \langle \nabla g_i(\bar{x}),\, x\rangle \geqq 0, \qquad i=1,\, \ldots,\, m\}$$

It may be noted that the condition in Theorem 4 is fairly strong. In fact, there are problems in which some vector x, such that $\langle \nabla g_i(\bar{x}),\, x\rangle \geqq 0$ for **all** i, does not satisfy $\langle \nabla f(\bar{x}), x\rangle > 0$, while condition (b) in Theorem 2 is satisfied. For instance, consider the problem

$$\begin{array}{rl} \text{minimize} & (x_1+1)^2+x_2^2 \\ \text{subject to} & x_1-x_2^2 \geqq 0 \\ \text{and} & 1-x_1-x_2 \geqq 0. \end{array}$$

Obviously, the solution is $\bar{x}=(\bar{x}_1, \bar{x}_2)=(0, 0)$ and $\nabla f(\bar{x})=(2, 0)$, $\nabla g_1(\bar{x})=(1, 0)$, $\nabla g_2(\bar{x})=(-1, -1)$. Taking $x=(0, -1)$, we see that the condition in Theorem 4 is not met. However, this problem satisfies condition (b) in Theorem 2, because the constraint set is compact.

5. Example

Let us consider the problem

(4) minimize $f(x)=\frac{1}{2}\langle x, Ax\rangle+\langle b, x\rangle$

subject to $g(x)=r-\langle x, Dx\rangle \geqq 0$,

where A and D are positive definite n×n matrices, b is an n vector and r is a positive number.

The logarithmic penalty method[4] is to solve a sequence of problems of the form

(5) minimize $U_k(x)=f(x)-t_k \log g(x)$ over $x \in R^n$,

where $\{t_k\}$ is a strictly decreasing sequence of positive numbers converging to zero. Put

$$h(x)=\log g(x)$$

and for $k=1, 2, \ldots$,

$$h_k(x)=t_k h(x).$$

Corresponding to (5), we can define the conjugate penalty function as

(6)
$$\begin{aligned} V_k(y)&=h_k^*(y)-f^*(y) \\ &=t_k h^*(y/t_k)-f^*(y). \end{aligned}$$

For problem (4), by direct calculation, we have

(7)
$$f^*(y)=\frac{1}{2}\langle y-b, A^{-1}(y-b)\rangle$$

and

(8)
$$h^*(y)=1-(1+r\langle y, D^{-1}y\rangle)^{\frac{1}{2}}+\log\frac{1+(1+r\langle y, D^{-1}y\rangle)^{\frac{1}{2}}}{2tr}.$$

Substituting (7) and (8) into (6), we have

$$\begin{aligned} V_k(y)=&\log[t_k+(t_k+r\langle y, D^{-1}y\rangle)^{1/2}]-(t_k^2+r\langle y, D^{-1}y\rangle)^{1/2} \\ &-\frac{1}{2}\langle y-b, A^{-1}(y-b)\rangle+t_k(1-\log 2rt_k). \end{aligned}$$

Finally, we mention briefly the convergence rate of the conjugate penalty functions. It is known[5] that for logarithmic penalty functions

$$\|x_k-\bar{x}\|=O(t_k).$$

Since

$$\nabla f(x)-\nabla f(z)=A(x-z) \qquad \text{for all } x \text{ and } z,$$

we have by Theorem 1

$$\begin{aligned}\|y_k-\bar{y}\| &= \|\nabla f(x_k)-\nabla f(\bar{x})\| \\ &\leqq \|A\|\cdot\|x_k-\bar{x}\|.\end{aligned}$$

Similarly,

$$\begin{aligned}\|x_k-\bar{x}\| &= \|\nabla f^*(y_k)-\nabla f^*(\bar{y})\| \\ &\leqq \|A^{-1}\|\cdot\|y_k-\bar{y}\|.\end{aligned}$$

Therefore we, can conclude that x_k and y_k converge at the same rate, namely,

$$\|y_k-\bar{y}\|=O(t_k).$$

In fact, a similar statement is valid for more general problems under assumption of the Lipschitz continuity of ∇f and ∇f^* on neighborhoods of $\bar{x}$ and $\bar{y}$, respectively, (see[9].)

6. Conclusion

We have presented a new class of sequential unconstrained optimization techniques for the solution of convex programming problems. The method has an advantage over ordinary penalty function methods in that it will circumvent unfavorable boundary properties of ordinary penalty functions, as far as the conditions in Theorem 2 are satisfied. Those conditions are met for some classes of problems that are often encountered in practice. For example, a strictly convex objective function and strictly concave constraint functions will form a problem which satisfies both conditions. However, we mention that the present method has a drawback in its practical implementation, because it is not an easy matter to obtain a terse expression of conjugates for any convex and concave functions. This approach may be particularly attractive for the type of problems such as problem (4), in which functions have their conjugates in a simple closed form.

References

1) R. A. Abrams and A. Ben-Israel, "Optimality conditions and recession cones," Operations Research 23 (1975) 549–553.
2) D. P. Bertsekas, "Multiplier methods: A survey," in: Proceedings IFAC 6th Triennial World Congress, Boston/Cambridge, 1975.
3) W. Fenchel, "Convex cones, sets and functions," mimeographed lecture notes, Princeton University (1951).
4) A. V. Fiacco and G. P. McCormick, Nonlinear programming: sequential unconstrained minimization techniques, Wiley, New York, 1968.
5) F. A. Lootsma, "A survey of methods for solving constrained minimization problems via unconstrained minimization," in: F. A. Lootsma, ed., Numerical methods for non-linear optimization, Academic Press, London, 1972, pp. 313–347.
6) D. G. Luenburger, Optimization by vector space methods, Wiley, New Yrok, 1969.

7) D. G. Luenburger, Introduction to linear and nonlinear programming, Addison-Wesley, Reading, 1973.
8) H. Mine and M. Fukushima, "Penalty function theory for general convex programming problems," J. Optimization Theory Appl. 24 (1978).
9) H. Mine, K. Ohno and M. Fukushima, "A 'conjugate' interior penalty method for certain convex programs," SIAM J. Control and Opt. 15 (1977) 747–755.
10) R. T. Rockafellar, "Helly's theorem and minima of convex functions," Duke Mathematical Journal 32 (1965) 381–397.
11) R. T. Rockafellar, "Extension of Fenchel's duality theorem for convex functions," Duke Mathematical Journal 33 (1966) 81–89.
12) R. T. Rockafellar, Convex analysis, Princeton University Press, Princeton, N. J., 1970.
13) R. T. Rokcafellar, Conjugate duality and optimization, Society for Industrial and Applied Mathematics, Philadelphia, Penn., 1974.
14) D. M. Ryan, "Penalty and barrier functions," in: P. E. Gill and W. Murray, eds., Numerical methods for constrained optimization, Academic Press, London, 1974, pp. 175–190.
15) J. Stoer and C. Witzgall, Convexity and optimizatioln in finite dimensions I, Springer-Verlag, Berlin, 1970.