ISSN (Print): 2476-8316
ISSN (Online): 2635-3490
Dutse Journal of Pure and Applied Sciences (DUJOPAS), *Vol. 4 No. 2 December 2018*

Dual and Feasible Directional Algorithm for the Solution of Nonlinear Constrained Optimization Problems

Nagwai, A.P[1] & Modi, B[1]*
[1]Gombe State University, Department of Mathematics,
Faculty of Science.
modi_pas@yahoo.co.uk

Abstract
An algorithm for the solution of nonlinearly constrained optimization problems is proposed by minimizing a sequence of convex quadratic problems with essentially no negativity constraints. The quadratic subproblems are solved by principal pivoting or other fast quadratic methods. A new method for preventing jamming is proposed in this paper. Also, a general convergence theorem for optimization algorithms is presented using the concept of a general necessary optimality function and incorporating the anti-jamming feature. All algorithms work under a choice of a number of step size selection methods.

Keywords: Optimization, linearity, Convergence, Algorithm, Convex Quadratic Problem, Dual

1.0 INTRODUCTION
In this work as seen in Section 2.2, an algorithm for solving nonlinear programming problems with nonlinear constraints as presented in Problem 2.1 was developed. At each step, a convex quadratic problem with linear constraints as seen in equations2.1.1a-c, were solved by principal pivoting (Cottle , 1968 and Cottle and Dantzig, 1968) or any fast quadratic programming algorithm (Bartels et al, 1970 and Stoer,1971). Although no rates of convergence are given, super linear order effects could be incorporated by appropriate choice of a matrix H(x) appearing in the quadratic sub problems. For instance, this matrix can be taken to as the inverse of a Hessian matrix or can be updated by methods similar to those of the variable matrix methods (Broyden,1969 and Stoer, 1971). Preliminary computational results are encouraging and indicate that by appropriate choice of $H(x_1)$ convergence can be increased appreciably.

A new anti-jamming procedure which is a modification and improvement of the Topkis-Veinott (Topkis and Veinott, 1967) procedure was developed. It is simpler than the Zoutendijk (Zoutendijk, 1960) and Zangwill (Zangwill, 1969)procedure in that one fixed ε tolerance is used throughout the algorithm.

The quadratic sub-problems are transformed dual problems to the feasible direction sub-problems of Topkis-Veinott (Topkis and Veinott, 1967) with the modified anti-jamming procedure. The Topkis-Veinottsub-problems have quadratic constraints and hence are not computationally tractable. The sub problem presented in this paper has a quadratic objective function and linear constraints, and can be solved by fast quadratic programming algorithms.

**Author for Correspondence*

The convergence of the algorithm is established by establishing a general theorem for the convergence of algorithms as seen in Section 3. This is achieved by using the concept of a general optimality function. The anti-jamming procedure is included in this general theorem together with various step size selection methods.

2.0 METHODOLOGY

2.1 THE ALGORITHM

The constrained minimization problem in Equation 2.1 was considered thus:
min $f(x)$, $X = \{x|\ x\in R^n,\ g(x) \leqq 0\}$……………………………….………...............................(2.1a)

Where $R^n$is the n-dimensional real Euclidean space,*f*is a function on R^n into the real's and g is a function on R^n into R^m. The following assumptions shall be imposed throughout Section 2.

Assumption 1

$f\epsilon\ C^1$…………………………………………..(2.1b)

Assumption 2

Each component g_j of g has a lipschitz continuous gradient that is
$\|\nabla g_j\ (y) - \nabla g_j\ (x)\ \| \leqq K\ \|\ y - x\|$, for x, y∈ X,……………………….…..….........................(2.1c)
Where, j=1… m.

Assumption 3

For some positive number k, where $\|X\|$ denotes the Euclidean norm $(XX)^{1/2}$.
$\|\nabla\ f(x)\| \leqq \alpha$, $\|\nabla g_j\ (x)\| \leqq \alpha$ for all $x \in X$,……………………..…...(2.1d)

Withsome positive number α, the algorithm was divided into two parts: 1) direction finding part and 2) step size partthus:

2.1.1 Direction Finding: At $X_{i,}$ solve the quadratic programming problem with linear constraints thus:

$-\theta\ (X_i, I(x_i))$: = min $\{\|\eta\nabla f(x_i) + (\eta,\ u_{I(x_i)}) \geqq 0$
$\eta+eu_{I(x_i)}=1+u_{I(x_i)}\nabla gI(x_i)(x_i)\|^2 H\ (x_i)^{-u}I(x_i)^{g}I(x_i)^{(x_i)}\}$ ....(2.1.1a)

and denote its solution by (η_i, u_i), were $\eta \in R$, $u \in R^m$, e is a vector of ones, such that:
$U_I\nabla g_I(x) = \sum_{j\varepsilon I} u_i\nabla g_i(x)$, $uIg_i(x) = \sum_{j\in I} u_i\ g_i(x)$,……………………................................(2.1.1b)

$$\begin{cases} I(x) = \{j| -\varepsilon \leqq g_j(\times) \leqq 0 \\ e = \text{any positive number, fixed throughout algorithm.} \end{cases}$$

and:

$$\begin{cases} H(\times_i) \text{ is any continuous, symmetric, uniformly positive definite} \\ n \times n \\ \text{matrix, that is} \\ M_1\ \|\ z\ \|^2 \leqq z\ H(\times)z \leqq\ M_2\ \ \|\ z\ \|^2\ \text{ for all } z \in R^n, \times \in \times, \text{and some numbers} \\ M_{2,}M_1 > 0 \end{cases}$$ ……….(2.1.1c)

$\|z\|_H^2$ denotes zHz. The symbol $u_{I(x_i)}$ denotes a vector with component indices in $I(x_i)$.Similarly,$g_{I(x_i)}$ denotes a vector with component indices in $I(x_i)$. For convenience we refer to $g_{I(x)}$as the ε-active constraints at x, and the remaining constraints as the ε-inactive constraints atx.

The direction P_i at x_i is defined by:

$$P_i = v_i H(x_i) q_i \quad \text{……………………………………............. ..(2.1.1d)}$$

Where,

$$q_i = -2\left(\eta_i \nabla f(x_i) + u_i \nabla g_{I(x_i)}(x_i)\right) \quad \text{…………………………..(2.1.1e)}$$

$$v_i = \max\left\{1, \frac{1}{2}, \frac{1}{4}, \ldots,\right\} \text{ such that } x_i + \mu p_i \in X \text{ for } 0 \leqq \mu \leqq 1 \quad \text{……….......(2.1.1f)}$$

2.1.2 Step Size: If either $\nabla f(x_i)p_i = 0$ **or** $q_i = 0$ terminate (x_i is stationary). If neither, use any of the following step size methods to find $x_{i+1} = x_i + \lambda_i p_i$ thus:

(Minimizationalong p_i)

$f(x_i + \lambda_i p_i) = \min f(x_i + \lambda p_i)\ 0 \leqq \lambda \leqq 1$……………………..(2.1.2a)

or

$f(x_i + \lambda_i p_i) = \min f(x_i + \lambda p_i)\ x_i + \lambda p_i \in X$……………………….....................................(2.1.2b)

(Armijo, 1966) λ_i= maximum $\{1, \frac{1}{2}, \frac{1}{4}, \ldots\}$ such that :

$f(x_i) - f(x_i + \lambda_i p_i) \geqq -\lambda_1^2 \nabla f(x_i) p_i$………………...(2.1.2c)

In the step-size method,*f*is assumed to have a lipschitz continuous gradient as seen in Section 2.

(Goldstein,1967) $\lambda_i = \begin{cases} 1, & \text{if } Y_i(1) \geqq p \\ \hat{\lambda}_i \text{ such that } p \leq 1 - p, & \text{if } Y_i(\lambda) < p \end{cases}$

Where,$Y_i(\lambda) = \frac{f(x_i) - f(x_i + \lambda p_i)}{-\lambda \nabla f(x_i) p_i}$……………………………………..……....................(2.1.2d)

and*p* is any fixed number satisfying the condition: o $< p \leqq \frac{1}{2}$. For this step-size method,*f* is assumed to be bounded below on X. However, before stating the convergence result for the algorithm in Section 2.1.2, weneed to define the concept of a stationary point.

2.1.3 Stationary Point:A point $\bar{x} \in x$ is said to be stationary if, $\theta(\bar{x}, \mathbf{I}(\bar{x})) = \mathbf{0}$, where θ is defined by the Equation in 2.1.1a. The significance of such a point follows from the result in Section 2.1.4.

2.1.4 Necessary and Sufficient Optimality Theorem (Necessity): If $\bar{x}$ solves the minimization problem stated in Section 2.1, then $\theta(\bar{x}, \mathbf{I}(\bar{x})) = \mathbf{0}$.

If in addition, a constraint qualification such as the Arrow-Hurwicz-Uzawa, Karlin, or Slater constraint qualification is satisfied at $\bar{x}$ (Mangasarian, 1970. pp. 102-105) then the Kuhn-Tucker condition are also satisfied at $\bar{x}$ that is:

$\nabla f(\bar{x}) + \bar{u}\nabla g(\bar{x}) = 0, \bar{u}g(\bar{x}) = 0, g(\bar{x}) \leqq 0, \bar{u} \geqq 0$. (Sufficiency) if $\bar{x} \in X$, $\theta(\bar{x}, I(\bar{x})) = 0, \bar{\eta} > 0$ (where $(\bar{\eta}, \bar{u})$ is the solution defined in 2.1.1awith $x_i = \bar{x}$………………………………...(2.1.4a)

Where, f is pseudo-convex at $\overline{x}$ and g_j, j = 1,…,m, are quasi-convex at $\overline{x}$, then $\overline{x}$ is a global solution of Section 2. 1.

Furthermore, Theorem 2.1.4indicates that a stationary point is a desirable point to have, since it is a Kuhn-Tucker point and under additional conditions it is also a global solution.
This is precisely what the Algorithm 2.1.1 - 2.1.2leads to as indicated by the convergence result discussed in Section 2.1.5.

2.1.5 Convergence Theorem:
Either the sequence $\{X_i\}$ generated by the Algorithm 2.1.1 - 2.1.2terminates at a stationary point $\overline{x_i}$, that is $\theta\{x\bar{I}, I(X\bar{\imath})$ =0 or each accumulation point $\overline{x}$ is stationary, that is $\theta(\overline{x}, I(\overline{x}))$=0.

We remark here that the Kuhn-Tucker constraint qualification doesnot guarantee that the Kuhn-Tucker conditions hold at an $\hat{x}$ for which $\theta(\hat{x}, I(\hat{x}))$ =0. For example the problem minimize {-X_1 | $X_2 \geqq x_1^3$, $X_2 \leqq x_1^2$} has a solution at (1,1), but at the origin $\theta(0, I(0))$ the Kuhn-Tucker constraint qualification is satisfied but not the Kuhn-Tucker conditions. Such examples are ruled out if we assume the original form of the Arrow-Hurwicz-Uzawa constraint qualification, which states that there exists a z $\in R^n$, such that $\nabla g_i(x)x>0$ for $i\in\{i \mid g_i(x)=0\}$.

We now make a few remarks about the algorithm and in particular about the quadratic minimization problem stated in Section 2.1.1, condition2.1.1a. Firstly, we observe that the tolerance $\varepsilon > 0$ is fixed once and for all during the algorithm. If it is set to $\varepsilon = \infty$, then:
I(x)={1,2,…,m}, and we have an anti-jamming procedure similar to that of Topkis and Veinott(Topkis and Veinott, 1967). It is more efficient to take ε some positive number so that we exclude ε-inactive constraints at x_i from Equation 2.1.1a.

For the unconstrained case, I(x) is empty and there remains no direction finding problems stated in Equationdefined in 2.1.1abecause the only remaining variable η in 2.1.1abecomes η=1 and $q_i = 2\nabla f(x_i)$ according to condition2.1.1f, the direction p_i is given by $p_i = -2H(x_i)\nabla f(x_i)$. If we take $H(x_i)$=I, we get Cauchy's method (Ortega and Rheinbolt, 1970) and if we take $H(x_i)=\nabla^2 f(x_i)^{-1}$, which is the *nxn* inverse Hessian matrix of second partial derivatives, we get the damped Newton method (Ortega and Rheinbolt, 1970), which has a quadratic rate of convergence. However any matrix $H(x_i)$ satisfying condition 2.1.1cwill also work. For later use and for comparison purposes, it is convenient to state a quadratic programming problem, the dual of which is problem 2.1.1a.

2.1.6 Quadratically Constrained Problem:

$\varphi(x_i, I(x_i))$: = min (δ, s)

$$\left\{\delta \middle| \begin{array}{c} \nabla f(x_i)s+\frac{1}{4}s\,H(x_i)^{-1}s\leqq\delta \\ g_{j\,(x_i)}+\nabla g_j(x_i)s+\frac{1}{4}s\,H(x_i)^{-1}s\leqq\delta, j\in I(x_i) \end{array}\right. \quad\ldots\ldots(2.1.6a)$$

Where, $I(x) = \{j \mid -\varepsilon \leqq g_j(x) \leqq 0\}$, and the variable s is related to the variable q of condition 2.1.1c by:

$$s=H(x_i)q \quad\ldots\ldots(2.1.6b)$$

We observe that condition 2.1.6a is a quadratically constrained problem, whereas 2.1.1a is a linearly constrained quadratic programming problem. This makes the problem stated in 2.1.1aconsiderably easier to solve, by principal pivoting methods (Cottle, 1968 and Cottle and

Dantzig, 1968), or any of the recent fast quadratic programming algorithms (Bartels et al, 1970 and Stoer, 1971): Problems 2.1.1aand 2.1.6are related through the duality theorem discussed in Section 2.10.

2.1.7 Duality Theorem: If $(\boldsymbol{\eta_I}, \mathbf{u_i})$ solves the problem in 2.1.1athen $\mathbf{s_i}$ = θ (x_i, I$\mathbf{x_i}$) and $\mathbf{s_i}$=-2H($\mathbf{x_i}$) $(\boldsymbol{\eta_{i,}}\boldsymbol{\nabla}\mathbf{f}(\times_i) + \mathbf{u_i}\boldsymbol{\nabla}\mathbf{g}(\mathbf{x_i}))$ solve 2.1.6. Conversely if $(\boldsymbol{\delta_i},\mathbf{s_i})$ solves 2.1.6. Then some $(\boldsymbol{\eta_I}, \mathbf{u_i})$ satisfying $\boldsymbol{\eta_I}, \boldsymbol{\nabla}\mathrm{f}(\mathbf{x_i}) + \mathbf{u_i}\boldsymbol{\nabla}\mathrm{g}(\mathbf{x_i}) = -\frac{1}{2}\, \mathbf{H}(\times_i)^{-1}s_i$ solves 2.1.1a. In both cases θ(x_i, I(x_i)) =4(x_i, I(x_i)).

We point out here that the problem 2.1.6 is related to the Topkis-Veinott (Topkis and Veinott, 1967) sub-problem in the following way. If the quadratic term in *s* is removed from the constraints inI(x_i) and if ε= ∞ then 2.1.6becomes the Topkis-Veinott direction finding problem given in their Theorem 3 of (Topkis and Veinott, 1967). We feel that our formulation in 2.1.1ahas two advantages over their formulation. Firstly,we consider only a subset of the constraints, that is, the $\mathcal{E}$-active set, whereas they considered all the constraints. Secondly, the formulation in 2.1.6 is impractical for taking into account quadratic terms, whereasthe problem in 2.1.1a can accommodate such effects. We also point out the difference between our fixed-ε technique and the ε-halving method of zoutendijk(Zoutendijk, 1960) and Zangwill (Zangwill, 1969) thus:
If the ε-active constraints at each x are divided into concave and non-concave constraints, that is,
$I(x) = I_1(x) U I_2(x), I_1(x)=\{ j \mid -\varepsilon \leqq g_j(x) \leqq 0, g_j \text{is concave}\}, I_2(x)= \{ J \mid -\varepsilon \leqq g_j(x) \leqq 0, g_j \text{ is not concave}\}$.
Hence, the problem in 2.1.1acan be replaced by:
$-\theta(x_i, I(x_i))$:=minimum $(\eta, u_i(x_i)) \geqq 0$

$$\eta + eu_{I_2}(x_i)=1$$

$\{\|\eta\nabla f(x_i) + u_{I(x_i)}(x_i)\, \nabla g_{I(x_i)}(x_i)\ \|^2 H(x_i)^{-u}\, I(x_i)^{g} I(x_i)(x_i)\}$……….....................…(2.1.7a)
and the problem in 2.1.6 by:
$\varphi(x_i, I(x_i))$: = min(δ ,s)

$$\begin{cases} \nabla f(\times_i\, s + \frac{1}{4}\, s\, H(X_i)^{-1}\, s \ \leqq\ \delta \\ \delta|g_j(\times_i)\ + \nabla g_j(\times_i)s\ + \frac{1}{4}\, s\, H(\times_i)^{-1}\, s\ \leqq\ \delta, j \in\ I_2(\times_i) \\ g_j(\times_i) + \nabla g_j(\times_i)s \qquad \leqq 0, j\ \in\ I_1(\times_i) \end{cases} \quad\ldots\ldots\ldots(2.1.7b)$$

All the above theorems hold for 2.1.1aand 2.1.6 now replaced by 2.1.7a and 2.1.7b respectively provided we make the additional assumption that $\nabla g_j(\times_i)$, $j \in I_1(\times_i)$are uniformly positively linearly independent, that is,$\|u_{I_1(x)}\nabla g_{I_1(x)}(\times)\ \|^2 \geqq\ \omega\ \|\ u_{I_1(x)}\ \|^2$for some:

$$\omega\ > 0\ and\ all\ u_{I_1(x)} \geqq 0.$$

Problem 2.1.7asimplifies further, if we consider the case when $I_{2(X_i)}$is empty. Then η = 1 and we have:
$-\theta(x_i, I(x_i))$: = minimum
$u_{I_i}(x_i) \geqq 0$
$\{\|\nabla f(\times_i) + u_{I_i}(x_i)\nabla g_{I_{1(x_i)}}(x_i)\quad \|^2\ H(x_i)^{-u_{I_1}}\ (x_i)^{g_{I_i}}\ (x_i)(x_i\ )\}$…………………….……...(2.1.7c)
We remark further that if in 2.1.7bwe had that in each constraint in which $H(\times_i)^{-1}$ appears, $H(\times_i)^{-1}$ is different matrix, say twice the Hessian of each function, as would be the case if we are taking quadratic approximations of f and g, then the corresponding dual problem 2.1.7a would have an $(\times_i)$ given by:

$$H(\times_i)\ =\ \frac{1}{2}(\eta\ \nabla^2\ f(\times_i)\ +\sum_{j\in I_2(\times_i)} u_j\nabla^2 g_j(\times_i)(\times_i)^{-1}\quad \ldots\ldots(2.1.7d)$$

Problem 2.1.7a would then be highly impractical to solve, except for the case when I_2 (x_i) is empty, in which case η=1 and we essentially have the counterpart of the constrained damped Newton method. However knowing the form of the matrix H(x_i) given above which results in a Newton method, should help in constructing updating schemes, such as the variable metric method and others (Flectcher and Powell, 1963) for constrained optimization.

A preliminary code using principal pivoting has been written by Toby J. Teorey and is being tested now. On Colville's problem, Number 1 (Colville, 1968) with H(x_i) = I, the time was -0057standard units, better than any reported method. On test problem number 1 (Colville, 1968) again with H(x_i) = I the time was -0202 standard units, again better than any reported method.

3.0 RESULT AND DISCUSSION

3.1 ALGORITHM CONVERGENCE THEOREM

In this section, we give a theorem which can be utilized to prove a wide class of optimization algorithms both constrained and unconstrained. We did not strive for as broad a generality as that of the Topkis-Veinott (Topkis-Veinott,1967), Zangwill (Zangwill, 1969) or Polak (Polak, 1971) convergence algorithms. However, because our result includes a new fixed-ε anti-jamming technique, and because it uses the novel concept of a general necessary optimality condition, we have included it here rather than appealing to more general results.

3.1.1. Algorithm Convergence Theorem:

Consider the problem min f(x), where X = {x | x ∈ R^n, g(x)≦ 0} f is a function x ∈ X with continuous first derivatives from R^n into R, and g is a continuous function from R^n into R^m. We define a general optimality function θ(x, I (x)) on X, not necessarily the same function of 2.1.1a, where I(x)= {j | -ε ≦ g_i (x) ≦ 0} for some arbitrary but fixed ε > 0 as follows:

3.1.1 x∈ X, implies that θ(x, I(x)) ≦0.
3.1.2 $\bar{x}$ Solves min f (x), implies that θ($\bar{x}$, **I**($\bar{x}$)) = 0.
3.1.3 For each x ∈ X and fixed J ⊂ {1, 2… m}, θ(x, J) is continuous in x.
3.1.4 x∈ X, J⊂ L ⊂ {1, 2… m} implies that θ(x, J) ≦ θ(x, L).
Consider the following algorithm. Start with anyx0∈ X, and with x_i,determine x_{i+1} as follows:

3.1.5 Direction Finding: Choose any direction p_i∈ P where P is some compact set in R^n such that:

a) X_i + μp_i ∈ X for all μ, 0 ≦ μ ≦ 1.
b) -∇ f(x_i) p_i≧ ϭ (-θ (x_i, I(x_i))

Where ϭ any increasing continuous function mapping (0,∞) into itself and such that ϭ (0) = 0.

3.1.6 Step Size:If $\nabla f(\mathbf{x_i})p_i$ =0 terminates, otherwise choose x_{i+1} = $\mathbf{x_i} + \boldsymbol{\lambda_i}\mathbf{p_i}$ ∈ X according to Anyrule such that if $(\mathbf{x_{i_j}}, \mathbf{p_{i_j}})$ converges to ($\hat{x}$, $\hat{\mathbf{p}}$)then $\nabla f(\hat{x})\hat{\mathbf{p}}$ =0. Thestep size methods defined in Section 2.1.2have this property. See the theorem in Section 3.10 for further explanations.

Either the sequence {xi} generated by the Algorithm3.1.5 – 3.1.6terminates at stationary pointx_i, that is $\theta(x_i, Ix_i) = 0$, or each accumulation point $\bar{x}$is a stationary point,as expressed in statement 3.1.6a:

$\theta(x_j, I(\bar{x}) = 0$……………………………………………………..(3.1.6a)

Proof: If for some$\bar{\imath}$, $\nabla f(x_{\bar{\imath}})p_{\bar{\imath}} = 0$, then the algorithm terminates and, by 3.1.5b, $\theta(x_{\bar{\imath}}, I(x_{\bar{\imath}})) = 0$. Suppose now $-\nabla f(x_{\bar{\imath}})p_{\bar{\imath}} > 0$ for all i, and let $\bar{x}$ be any accumulation point of $\{x_i\}$. Take any subsequence $\{x_i\}_{L_1}$ of $\{x_i\}$ converging to$\bar{x}$. Then we extract a further subsequence $\{x_i\}_{L_2}$such that: $\{x_i, p_i\}_{L_2}$ converges to $(\bar{x}, \bar{p})$ by 3.1.6,and 3.1.5b, $\nabla f(\bar{x})\bar{p} = 0$.

3.1.7 $-\nabla f(x_{\bar{\imath}})p_{\bar{\imath}} \geq \sigma(-\theta(x_i, Ix_i))$ for $i \in L_2$.

Because there is a finite number of constraints $g_j(x) \leq 0$, j = 1, 2, , m, one subset J ⊂{1, 2, …, m} must occur and infinite number of times in the sequence of sets $\{I(x_i)\}_{L_3}$. Extract a further subsequence $L_3 \subset L_2$ such that only j occurs in $\{I(x_i)\}_{L_3}$.Thenthe statement in 3.1.7becomes $-\nabla f(x_{\bar{\imath}})p_{\bar{\imath}} \geq \sigma(-\theta(x_i, J))$ for $i \in L_3$ and in the limit $-\nabla f(x_{\bar{\imath}})p_{\bar{\imath}} \geq \lim_{i \in L_3} \sup \sigma(-\theta(x_i, J))$. By using 3.8 and the fact $\sigma(-\theta(x_i, J))$ is continuous in x, which follows from 3.1.3 and 3.1.5b, this last inequality becomes $0 \geq \sigma(-\theta(\bar{x}, J))$. This implies that by 3.1.5b, $\theta(\bar{x}, J) = 0$. Now if $I(\bar{x}) \subset J$ and $I(\bar{x}) \neq J$, then for $j \in J$, $j \in I(\bar{x})$ and $i \in L_3$. We then obtain $-\epsilon \leq g_j(x_i) \leq 0$ and in the limit $-\epsilon \leq g_j(\bar{x}) \leq 0$, and hence $j \in I(\bar{x})$, which is a contradiction. So $J \subset I(\bar{x})$and by3.1.1and 3.1.4we have: $0 = \theta(\bar{x}, J) \leq \theta(\bar{x}, I(\bar{x})) \leq 0$. Hence $\theta(\bar{x}, I(\bar{x}))$ and $\bar{x}$ is stationary.

We note here that 3.1.5although not used explicitly in the proof, is needed in all the step size methods suggested in 2.1.2. Also, the condition $\epsilon > 0$ can be relaxed to $\epsilon \geq 0$ in the above theorem. However, in establishing 3.1.5bin practical cases such as in the proof of Algorithm 2.1.1 – 2.1.2, here $\epsilon > 0$ is needed to prove that 3.1.5bholds.

We also note that ϵ can be set to $+\infty$ in which case the above theorem holds with I(x) = {1,2,..,m} for all $x \in X$. this is essentially the anti-jamming procedure of Topkis- Veinoit(Topkis- Veinoit, 1967) which is less efficient than the one proposed here.

We show further that the three step sizes proposed in2.1.1a, 2.1.1b, and 2.1.1chave the stated property in 3.1.6. Note that, step size results in more general are given in (Daniel, 1970).

3.1.8 Steps Size Theorem:

These conditions imply that:$u_i(\bar{x},)\, g_I(\bar{x},) = 0$ and hence $-\theta(\bar{x}, I(\bar{x},)) = 0$. If any of the mentioned constraint qualifications are satisfied, then the Arrow-Hurwicz-Uzawa constraint qualification (Mangasarian, 1970, p. 102) is satisfied (Mangasarian, 1970, Lemma 6, p. 103) and hence (Mangasarian, 1970, theorem 7, p. 105) the Kuhn-Tucker conditions are satisfied at $\bar{x}$,. (sufficiency) if $\theta(\bar{x}, I(\bar{x},)) = 0$, then $\eta \nabla f(\bar{x},) + u \nabla \overline{g(\bar{x},)} = 0$, $\bar{u}_I(\bar{x})\, g_i(\bar{x}) = 0$, $g(\bar{x}) \leqq 0$,$\bar{u}_I(\bar{x}) \geqq 0$ if $\bar{\eta} > 0$ then $\bar{x}$ solves the minimization problem (Mangasarian, 1970, theorem 2, p. 153).

Proof of Theorem 2.1.5.We invoke Theorem 3.1 to prove the convergence of the algorithm 2.1.1 -2.1.2. Hence, we have to show that conditions 3.1.1–3.1.4are satisfied. For notational simplicity, we shall drop the indices $I(x_i)$ and I, and the arguments of the functions. All functions are evaluated at x_i, and all g are $gI(x_i)$, u are $uI_{(xi)}$ and all x are x_i.

By using the Kuhn-Tucker condition (Mangasarian, 1970, theorem 7, p. 105), we have that at the solution (η,u) and some real number ξ must satisfy the Kuhn-Tucker conditions, where a prime denotes a transpose of matrix thus:

$2\eta\|\nabla f\|_H^2 + 2u\nabla gH\nabla g - \xi \geqq 0$……………………………………(3.1.8a)

$2\eta\nabla gH\nabla f + 2\nabla gH\nabla g'u - g - e\,\xi \geqq 0$……………………………………(3.1.8b)

$2\eta^2\|\nabla f\|_H^2 + 2\eta u\nabla gH\nabla f - ug - \xi\eta = 0$……………………………………(3.1.8c)

$2\eta^u\nabla gH\nabla f + 2u\nabla gH\nabla g'u - ug - eu\,\xi = 0$……………………………………(3.1.8d)

$\eta + 3eu - 1 = 0$……………………………………(3.1.8e)

By using 3.1.8c, 3.1.8d, 3.1.8e and 2.1.1awe get that:

$0 \geqq \theta(x, I(x)) = \|\eta\ \nabla f + u\,\nabla g\|_H^2 - \xi$……………………………………(3.1.8f)

Hence,

$\xi \geqq \|\eta\ \nabla f + u\,\nabla g\|_H^2$……………………………………(3.1.8g)

As such, using 3.1.8fand3.1.8g gives:

$-\theta(x, I(x)) \leqq -(\eta\ \nabla f + u\,\nabla g)\,H(\eta\ \nabla f + \nabla g'u - 2\nabla f)$…………………………… (3.1.8h)

Now Let:

$Q = -2(\eta\nabla f + u\,\nabla g)$…………………………………… (3.1.8i)

Note that by 3.1.8hand 3.1.8i, if q = 0, then $0 \leqq -\theta\ (x, I(x)) \leqq 0$. Hence, x is stationary and the algorithm terminates. So for $q \neq 0$, set:

All $\beta(x) = \text{minimum}\ \{1, \frac{M_1\xi}{M_2^2\ K\|q\|_H^2}, \frac{-1+(1+\varepsilon k/a^2)^{\frac{1}{2}}}{2kM_2}\}$…………………………… (3.1.8j)

By 3.1.8gand 3.1.8i we now have:

$\frac{M_1\xi}{M_2^2\ K\|q\|_H^2}, \frac{M_1\xi}{4M_2^2\ K\|\ \eta\ \ \nabla f+u\ \nabla g\|_H^2} \geqq \frac{M_1}{4M_2^2\ K}$……………………………………(3.1.8k)

Hence, by all and 3.1.8k:

$\beta(x) \geqq y\text{: minimum}\ \{\frac{M_1}{4M_2^2\ K}, \frac{-1+(1+\varepsilon k/a^2)^{\frac{1}{2}}}{2kM_2}\} > 0$……………………………………(3.1.8i)

We now define:

$t: = \beta(x)Hq$……………………………………(3.1.8m)

At this point, we shall now show that: $x + \mu t \in X$ for $0 \leqq \mu \geqq 1$.

For $j \in I(x)$, and $0 \leqq \mu \geqq 1$: we have 3.1.8n, by dropping the argument of β:

$g_j(x + \mu t) = g_j\ (x + \mu\beta Hq) \leqq g_j + \mu\beta\nabla g_j Hq + \frac{k\mu^2\beta^2}{2}\|Hq\|^2)$……………………………………(3.1.8n)

(byLipschitz continuity of∇g_j):

$= \mu\beta\ (g_j + \nabla g_j\, Hq) + (1 - \mu\beta)g_j + \frac{k\mu^2\beta^2}{2}\|Hq\|^2$……………………………………(3.1.8o)

$\leqq \mu\beta(-\xi + \frac{k\mu\beta M_2^2}{2M1}\ \|\ q\ \|_H^2)$. By 3.1.8b, $gj \leqq 0$, and $0 < \beta \leqq 1$.

$\leqq \frac{-\mu\beta\xi}{2}$ (by all since $\frac{\mu\beta}{2} \leqq \frac{\beta}{2} \leqq \frac{M_1\xi}{2M_2^2\ K\|q\|_H^2} \leqq 0$……………………………………(3.1.8p)

For $j \in I(x)$, and $0 \leqq \mu \geqq 1$:we have that:

$G_j\ (x + \mu t) \leqq g_j\ \mu\beta\nabla g_j\, Hg + \frac{k\mu^2\beta^2}{2}\|Hq\|^2 < -\varepsilon + \mu\beta\nabla g_j Hq + \frac{k\mu^2\beta^2}{2}\|Hq\|^2$ (since$j \in I(x)$)

$\leqq -\varepsilon + \mu\beta\propto M_2\ 2\propto + \frac{k\mu^2\beta^2}{2} M_2^2\ 4 \propto^2$ (by 3.1.8j, $\|q\| \leqq 2\propto$)

$=[-\frac{\varepsilon}{2} + 2\propto^2 M_2 \mu\beta + 2 \propto^2 M_2^2 k\mu^2\beta^2] -\frac{\varepsilon}{2} \leqq -\frac{\varepsilon}{2} < 0$ ....(3.1.8q)

It then follows from the inequality $=[-\frac{\varepsilon}{2} + 2\propto^2 M_2 \mu\beta + 2 \propto^2 M_2^2 k\mu^2\beta^2] -\frac{\varepsilon}{2}$, that the strictly convex quadratic expression in μβ has a negative root and a positive root equal to: $\frac{-1+(1+\varepsilon k/a^2)^{\frac{1}{2}}}{2kM_2}$ and for all μβ, is less than or equal this positive root. As such, the square-bracketed quadratic expression is non-positive.Hence we have just shown that $x + \mu\beta(x)Hq \in X$, for $0 \leqq \mu \geqq 1$and by 3.1.8l:
$X + \mu \gamma Hq \in X$ for $0 \leqq \mu \geqq 1 (\gamma > 0)$..(3.1.8r)

But by 2.1.1fand 3.1.8r, we have that:
$X + \mu v Hq \in X, \quad v = \max \{1, \frac{1}{2} \frac{1}{4}, \ldots\}$, for $0 \leqq \mu \geqq 1$..(3.1.8s)

So we either have: v = 1 or 2v> γ, otherwise by 3.1.8r we would have had $x + 2v\, Hq \in X$). Hence
$v \geqq \min \{1, \frac{\gamma}{2}\} = \gamma' > 0$.. (3.1.8t)

Furthermore, since p = vHq, 3.1.8s implies that:
$X + \mu p \in X$ for $0 \leqq \mu \geqq 1$..(3.1.8u)

3.1.8t is condition 3.1.56a of Theorem 3.1. To verify condition 3.1.5b, wenote that by 3.1.8h and 3.1.8i we have:
$-\theta (x, I(x)) \leqq \frac{1}{4} \| q \|_H^2 - qH\nabla f - \nabla fp = -v\, \nabla fHq \geqq -v\, \theta (x, I(x)) + \frac{v}{4} \| q \|_H^2 \geqq -v\, \theta (x, I(x))$
$\geqq -\gamma' \theta (x, I(x))$ .... (3.1.8v)

And since, by 3.1.8t, γ' is a positive number, condition 3.1.5bis then established.
We only need to verify conditions 3.1.1 to 3.1.4. Condition 3.1.1 follows from the fact that for x inX, $g_{I(x)}(x) \leqq 0$ and hence $\theta (x, I(x)) \leqq 0$. Condition 3.1.2 follows from the fact that if $\overline{x}$is optimal, then by the fritz John theorem (Mangasarian, 1970, Theorem 2, p. 99)there exists $(\bar{\eta}, \bar{u})$ such that:
$\bar{\eta}\nabla f(\overline{x}) + \bar{u}\nabla g(\overline{x}) = 0, \ \bar{u}g(\overline{x}) = 0, \ (\bar{n}, \bar{u}) \geqq 0, \ g(\overline{x}) \leqq 0, \bar{n} = e\bar{u} = 1$..............................(3.1.8w)

Hence,$\theta(\overline{x}, I(\overline{x})) = 0$. Condition 3.1.3follows from (Berge, 1963, Maximum Theorem, p. 116) since (η, u) lie in the fixed compact set $\{(\eta, u) \geqq 0\}$, and the objective function of 2.1.1ais continuous in (x, η, u). Also, condition 3.1.4follows from the fact that if $J \subset L$, then for $(\eta, u_J) \geqq 0$, $\eta + eu_J = 1$,(η, u_L) defined by $(\eta, u_J, 0_{Jd\,J})$ and also satisfies $(\eta, u) \geqq 0$ and $\eta + eu_L = 1$. This gives the same value of the objective function of 2.1.1a,hence $-\theta(\times, L) \leqq -\theta(\times, J)$.

Proof of Theorem 2.1.7: By direct application of the duality theorems of non-linear programming (Mangasarian, 1970, section 8.1, Theorems 4 and 6), Theorem 2.1.7holds for 2.1.1a as a primal problem andits dual thus:
$\text{Max} -\frac{1}{4} sH(X_i)^{-1} s\, S + u_{I(x_i)}\, g_{I(x_i)}(x_i) S, \eta, u$
Subject to :
$\eta\nabla f(X_i) + u_{I(x_i)} \nabla g_{I(x_i)}(x_i) + \frac{1}{2} H(x_i)^{-1} s = o, \eta = eu_{I(x_i)} = 1, \eta, u_{I(x_i)}) \geqq 0$..................(3.1.8x)
Substituting the first constraint in the objectives function and changing it to a minimization problem gives problems 2.1.1a. Note that: s=s=H (x_i)qas giving by 2.1.6b.

4. CONCLUSION

An algorithm for solving nonlinear programming problems with nonlinear constraints was presented. At each step, a convex quadratic problem with linear constraints as seen in equations2.1.1a-c, were solved by principal pivoting algorithm. Computational results indicate that by appropriate choice of $H(x_1)$ convergence increased appreciably.

The new anti-jamming procedure which is a modification and improvement of the Topkis-Veinott (Topkis and Veinott, 1967) procedure developed proved to be simpler than the Zoutendijk (Zoutendijk, 1960) and Zangwill (Zangwill, 1969) procedure in that one fixed ε tolerance is used throughout the algorithm.

The quadratic sub-problems are transformed dual problems to the feasible direction sub-problems of Topkis-Veinott (Topkis and Veinott, 1967) with the modified anti-jamming procedure.The sub problem presented in this paper has a quadratic objective function and linear constraints, and can be solved by fast quadratic programming algorithms.

The convergence of the algorithm is established by establishing a general theorem for the convergence of algorithms as seen in Section 3. This is achieved by using the concept of a general optimality function. The anti-jamming procedure is included in this general theorem together with various step size selection methods.

REFERENCES

Armijo L.L. (1966). Minimization of functions having Lipschitz continuous first partial derivatives. Pacific J. Math, 16, 1966, 1-3.

Bartels, R.H.,Golub, G. H and Saunders, M. A. (1970).Numerical techniques in mathematical programming.In Nonlinear Programming, ed. J. B. Rosen, O. L.Mangasarian and K. Ritter. Academic Press, 1970, 123-176.

Berge, C. (1963). Topological Spaces. MacMillan, New York, 1963.

Broyden, C. G. (1969). A new method of solving nonlinear simultaneous equations. Comp. J. 12, 1969, 94-99.

Colville, A. R. (1968). A comparative study of nonlinear programming codes, IBM New York Scientific Center Technical Report, No. 320-2949, June 1968.

Cottle, R. W. (1968).The Principal pivoting method of quadratic programming. In Mathematics of the decision sciences,Part 1, ed. G.b. Dantzig and A. F. Veinott, Amer. Math. Soc., Providence, R.I. , 1968, 144-162.

Cottle, R.W and Dantzig, G. B. (1968). Complementary pivot theory of mathematical

Daniel, J.W. (1970). Convergent step-sizes for gradient-like feasible direction algorithms for constrained optimization, in nonlinear programming. ed. J. B. Rosen, O. L. Mangasarian and K. Ritter, Academic Press, 1970, 245-274.

Fletcher, R and Powell, M.J.D. (1963).A rapidly convergent descent method for minimization.Comp. J. 6, 1963, 1963- 168.

Goldstein, A. A. (1967).Constructive Real Analysis.Harper and row, New York, 1967.

Mangasarian, O. L. (1970). Nonlinear programming. McGraw Hill, New York, 1970.

Ortega, J. M. and Rheinbolt, W.C. (1970).Iterative solution of nonlinear equations of several variables.Academic press, New York, 1970.

Polak, E. (1971).Computational Methods in Optimization.Academic Press, New York, 1971.

programming. *Linear Algebra and Appl*. 1, 1968, 103-125.

Stoer, J. (1971).On the numerical solution of constrained least-squares problems.SIAM J. Num. Anal.8, 1971, 382-411.

Topkis,D. M and Veinott, A. F. (1967).On the convergence of some feasible direction algorithms for nonlinear programming. SIAM J. Control 5, 1967, 268-279.

Zangwill, W. I. (1969). Nonlinear Programming, Prentice Hall.Englewood Cliffs, N. J., 1969.

Zoutendijk, G. (1960).Methods of Feasible Directions, Elsevier, Amsterdam, 1960.