J. Operations Research Soc. of Japan
Vol. 17, No. 4, December 1974

AN ALGORITHM FOR THE CONSTRAINED MAXIMIZATION IN NONLINEAR PROGRAMMING[1]

KUNIO TANABE

The Institute of Statistical Mathematics

(Received April 11; revised October 11, 1973)

Abstract

An algorithm is described for solving nonlinear programming problems. A continuous version of Rosen's gradient projection method is fully investigated and an algorithm is derived by discretizing the continuous model for numerical computation. A new iterative procedure is introduced for computing gradient project. The algorithm does not cause hunting in the neighbourhood of a maximal point and works well for both equality and inequality constrained maximization problems.

1. Introduction

A new algorithm is presented for solving nonlinear programming problem: To maximize a differentiable objective function $f(x_1, \cdots, x_n)$ of n variables constrained by differentiable equalities $g_j(x_1, \cdots, x_n)=0$, $j=1, \cdots, r$, and inequalities $g'_j(x_1, \cdots, x_n)\geqq 0$, $j=r+1, \cdots, r+s$.

Instead of the Lagrangean multiplier method, we take a geometric approach to the problem. We introduce a nonlinear autonomous system whose integral curve converges to a maximal point of the problem and then we derive an algorithm which is essentially an

[1] This is a revised version of the author's manuscript [9].

iterative procedure of numerical solution of the system by the Runge-Kutta-Merson method.

The algorithm is continuous version of Rosen's gradient projection method [6]. Our approach seems preferable to his direct analysis of discretized algorithm, because our algorithm can more faithfully trace the curved surface of the feasible region to a maximal point.

Our method can be successfully applied to a wide variety of nonlinear optimization problems and illustrative examples are shown in § 5.

2. The Statement of the Problem

Introducing squared slack variables, the inequality constraints are replaced by equality ones

$$g_{r+j}(x_1, \cdots, x_{n+s}) \equiv g'_{r+j}(x_1, \cdots, x_n) - x_{n+j}^2 = 0, \quad j=1, \cdots, s.$$

Then f and $g's$ can be considered as functions of $n+s$ variables $X \equiv (x_1, \cdots, x_n, x_{n+1}, \cdots, x_{n+s})$.

Let G be a mapping from $n+s$ dimensional Euclidian space R^{n+s} to $r+s$ dimensional one R^{r+s} described by an $(r+s)$-tuple

$$G(X) \equiv (g_1(X), \cdots, g_r(X), g_{r+1}(X), \cdots, g_{r+s}(X)).$$

Let $V(G)$ be the inverse image $G^{-1}(0)$ of $0 \in R^{r+s}$, *i.e.*

$$V(G) \equiv \{X \equiv (x_1, \cdots, x_{n+s});\ g_j(x_1, \cdots, x_n)=0,\ j=1, \cdots, r$$
$$\text{and } g_{r+j}(X) \equiv g'_{r+j}(x_1, \cdots, x_n) - x_{n+j}^2 = 0,\ j=1, \cdots, s\}$$

Assumption. We assume that

1) f and $g's$ belong to class C^2, the family of twice continuously differentiable functions of X.
2) the Jacobian matrix $J_G(X) \equiv (\partial g_j/\partial x_i)$ of the mapping G has rank $r+s$ for all X in $V(G)$. ($V(G)$ is nonempty.)

By the implicit function theorem, it is easily seen that $V(G)$ is an $(n-r)$-dimensional differentiable manifold in R^{n+s} and that f is a differentiable function on the manifold $V(G)$.

Then the problem is to find a maximum point of the function f on the manifold $V(G)$.

3. Theoretical Bases

Let $(df)_p \in T_p^*(V)$ be the differential of f at a point $p \in V(G)$, where $T_p^*(V)$ is the dual of the tangent space $T_p(V)$ of the manifold $V(G)$ at a point p. A point $p \in V(G)$ is called a critical point of f on the manifold $V(G)$, if $(df)_p=0$. Clearly a maximal point is a critical point.

The following is easily deduced by the implicit function theorem. For example, see Hancock [4] or Tanabe [9].

Proposition 1. A point $p \in V(G)$ is a critical point of f on $V(G)$, iff $\nabla f((X)_p)$ is in the linear subspace of R^{n+s} spanned by $\nabla g_1((X)_p)$. $\cdots$, $\nabla g_{r+s}((X)_p)$, where ∇ is the gradient operator $(\partial/\partial x_1, \cdots, \partial/\partial x_{n+s})^t$ and $(X)_p \in R^{n+s}$ is the value of X at a point $p \in V(G)$.

In the following, $(X)_p$ will be identified with p.

Next consider an autonomous system

$$dX/dt=[I-J_G^t(X)(J_G(X)J_G^t(X))^{-1}J_G(X)]\nabla f(X) \tag{1}$$

where I is the identity matrix and $J_G^t(X)$ is the transposed of the Jacobian matrix $J_G(X)$ of the mapping G, *i.e.*

$$\begin{aligned} J_G^t(X) &= (\partial g_j/\partial x_i)(X) \\ &= (\nabla g_1(X), \cdots, \Delta g_{r+s}(X)) \\ &= \left(\begin{array}{c|ccc} \tilde{\nabla} g_1, \cdots, \tilde{\nabla} g_r, & & \tilde{\nabla} g_{r+1}, \cdots, \tilde{\nabla} g_{r+s} & \\ \hline & -2x_{n+1} & & 0 \\ 0 & & \ddots & \\ & 0 & & -2x_{n+s} \end{array}\right) \end{aligned} \tag{2}$$

where $\tilde{\nabla} \equiv (\partial/\partial x_1, \cdots, \partial/\partial x_n)^t$. The right-hand-side of Eq. (1) will be denoted by $\Psi(X) \in R^{n+s}$.

The following theorem is easily shown and very useful for the calculation of $\Psi(X)$.

Theorem 2. There exists a unique differentiable mapping $\Lambda(X) \in R^{r+s}$ defined on $V(G)$ which satisfies

$$(3)\qquad \left[\begin{array}{c:c} I & J_G^t(X) \\ \hdashline J_G(X) & 0 \end{array}\right]\left[\begin{array}{c} \Psi(X) \\ \hdashline \Lambda(X) \end{array}\right]=\left[\begin{array}{c} \nabla f(X) \\ \hdashline 0 \end{array}\right]$$

or equivalently

$$(4)\qquad \Psi(X)=\nabla f(X)-J_G^t(X)\Lambda(X)\,, \quad \text{and}$$
$$J_G^t(X)\nabla f(X)=J_G(X)J_G^t(X)\Lambda(X)\,.$$

It should be noted that the matrix $J_G(X)J_G^t(X)$ is nonsingular.

Proposition 3. A point $p\in V(G)$ is a critical point of the function f on the manifold $V(G)$, iff

$$(5)\qquad \Psi((X)_p)=0\in R^{n+s}\,.$$

Proof. In this proof, $(X)_p$ is identified with X. If $\Psi(X)=0$, it follows from Eq. (4) that $\nabla f(X)=J^t(X)\Lambda(X)$. This means that $\nabla f(X)$ is a linear combination of $\nabla g_1(X)$, $\cdots$, $\nabla g_{r+s}(X)$.

Conversely, if $\nabla f(X)=\sum_{i=1}^{r+s}\lambda_i\nabla g_i(X)$, then Eq. (3) is satisfied by $\Psi(X)=0\in R^{n+s}$ and $\Lambda(X)=(\lambda_1, \cdots, \lambda_{r+s})^t$. Since Eq. (3) has a unique solution, then it follows that $\Psi(X)=0$. Hence Proposition 1 applies to give the desired result.

From this proposition, it follows that a critical point of the system (1) is just a critical point of the function f on the manifold $V(G)$. The set of critical points will be denoted by C_f, *i.e.*

$$C_f\equiv\{p\in V(G);\ (df)_p=0\}=\{X\in V(G);\ \Psi(X)=0\}\,.$$

It should be noted that the set C_f is closed in the manifold $V(G)$.

Theorem 4. For any feasible point $X^0\in V(G)$, there exists a unique solution $X(t, X^0)$ of the system (1) defined in some interval $[0, M)$ with the initial point $X_0=X(0, X^0)$. The solution has the following properties:

1) The trajectory of the solution is feasible, *i.e.*

$$X(t, X^0)\in V(G) \quad \text{for all} \quad t\in[0, M)\,.$$

That is, the manifold $V(G)$ is an invariant set of the system (1).

2)

$$(6)\qquad df(X(t, X^0))/dt=||\Psi(X(t, X^0))||^2\geqq 0$$

for $t\in[0, M)$. Further, if the initial point $X^0\in V(G)$ is not a critical point, then $df(X(t, X^0))/dt>0$ for $t\in[0, M)$.

Proof. By the assumption, $J_G(X)$ is continuous in the neighbourhood of $V(G)$ in R^{n+s}, and $J_G(X)$ is of full rank on $V(G)$. Hence, there exists an open neighbourhood $U_V\subset R^{n+s}$ of $V(G)$ such that $J_G(X)J_G^t(X)$ is nonsingular in U_V. From the elementary theory of differential equations, we have the first half of the theorem.

Next, for every $i=1, \cdots, r+s$, we have

$$dg_i(X(t, X^0))/dt=\langle\nabla g_i(X), dX/dt\rangle =\langle\nabla g_i(X), \Psi(X)\rangle . \tag{7}$$

It follows from Eq. (3) that

$$J_G(X)\Psi(X)=0\in R^{r+s} . \tag{8}$$

This means that

$$\langle\nabla g_i(X), \Psi(X)\rangle=0 \quad \text{for} \quad i=1, \cdots, r+s . \tag{9}$$

Then it follows from Eqs. (7) and (9) that $g_i(X(t, X^0))$ is constant in $t\in[0, M)$, so that

$$g_i(X(t, X^0))=g_i(X(0, X^0)) =g_i(X^0) =0 \quad \text{for} \quad t=1, \cdots, r+s . \tag{10}$$

Thus the trajectory $X(t, X^0)$, $t\in[0, M)$, lies in the manifold $V(G)$.
Finally we have

$$\begin{aligned} df(X(t, X^0))/dt&=\langle\nabla f(X), \Psi(X)\rangle \\ &=\langle\nabla f(X), P(X)\nabla f(X)\rangle \\ &=\langle\nabla f(X), P^2(X)\nabla f(X)\rangle \\ &=\langle P(X)\nabla f(X), P(X)\nabla f(X)\rangle \\ &=\|\Psi(X(t, X^0))\|^2 \end{aligned} \tag{11}$$

where $P(X)$ denotes the matrix $[I-J_G^t(X)(J_G(X)J_G^t(X))^{-1}J_G(X)]$, which is symmetric and idempotent.

If $\|\Psi(X(t, X^0))\|$ vanishes at $\tilde{t}\in[0, M)$, then it follows that

$$\Psi(X(\tilde{t}, X^0))=0\in R^{n+s} , \tag{12}$$

hence $X(\tilde{t}, X^0)$ is a critical point. On the other hand, $X(t)\equiv X(\tilde{t}, X^0)$, $t\in[0, \infty)$, is obviously a solution of the system (1). From the uniqueness of the solution of (1), it follows then that

$$X(t, X^0) \equiv X(\tilde{t}, X^0), \ t \in [0, \infty).$$

This contradicts the assumption that X^0 is not a critical point. Thus the proof is completed.

It is seen from this theorem that there is no periodic solution of (1).

Theorem 5. Let $X^0 \in V(G)$ be a non-critical point. If the solution $X(t, X^0)$ of (1) defined in $[0, M)$ with the initial point X^0 is bounded, then $M=\infty$ and the positive limit set Γ is a compact connected set contained in the boundary ∂C_f of the set C_f of critical points.

Proof. Since the solution lies in the manifold $V(G)$, it follows from the assumption that the positive limit set Γ of the solution is non-empty, invariant and compact in the manifold $V(G)$, hence $M=\infty$. Let K be a compact set in the manifold $V(G)$ which contains the trajectory $X(t, X^0)$, $t \in [0, \infty)$. From the continuity of f on $V(G)$, it follows that f is bounded on K. As is seen in Theorem 4, $f(X(t, X^0))$ is a strictly monotone increasing function of $t \in [0, \infty)$, then, it follows that $f(X(t, X^0))$ converges to $\tilde{l}<\infty$ as $t\to\infty$. Again from the continuity of f, it follows that

$$f(\tilde{X}) \equiv \tilde{l} \quad \text{for} \quad \tilde{X} \in \Gamma.$$

Let $\tilde{X}$ be a point in the positive limit set Γ. Since $\Gamma \subset V(G)$ is an invariant set of the system (1), there exists a solution $X(t, \tilde{X})$ defined in some interval $[0, N)$ with the initial point $\tilde{X}$ such that

$$X(t, \tilde{X}) \in \Gamma \quad \text{for} \quad t \in [0, N).$$

If $\tilde{X}$ is not a critical point, then it follows from Theorem 4 that $f(X(t, \tilde{X}))$ is a strictly monotone increasing function of $t \in [0, N)$. But this contradicts the fact that $f(X)$ is constant on Γ. Thus Γ is contained in C_f. It is easily seen that Γ is connected. (See [1].)

It is easily seen that $\Gamma \subset \partial C_f$, so the proof is omitted.

It follows from the theorem that if the solution converges to a point $\hat{X}$, then $\hat{X}$ is a critical point of the function f on the manifold $V(G)$.

Corollary 6. In the case of Theorem 5, we have

1)

(13) $$l(t)\equiv\int_0^t\left\|\frac{dX}{dt}\right\|dt<\kappa\sqrt{t},$$

where $\kappa\equiv\sqrt{f(\Gamma)-f(X^0)}$ and $f(\Gamma)=\tilde{l}$, and

2)

(14) $$dl/dt=df/dl=||\Psi(X(t,\ X^0))||.$$

Proof. We have by Theorem 4,

$$f(X(t,\ X^0))-f(X^0)=\int_0^t||\Psi(X(t,\ X^0))||^2dt\ .$$

By the proof of Theorem 5, we have $\lim_{t\to\infty}f(X(t,\ X^0))=\tilde{l}$, hence $\int_0^\infty||\Psi(X(t,\ X^0))||^2dt=\tilde{l}-f(X^0)$. By Schwartz's inequality, we have

$$l(t)=\int_0^t||\Psi(X(t,\ X^0))||dt\leqq\left(\int_0^t||\Psi(X(t,\ X^0))||^2dt\right)^{1/2}\left(\int_0^t dt\right)^{1/2}$$
$$<\sqrt{\tilde{l}-f(X^0)}\sqrt{t}\ .$$

Eq. (14) is easily seen from Theorem 4.

Let $L(X^0)$ be a connected component of the set $\{X\in V(G);\ f(X)\geqq f(X^0)\}$, that contains $X^0\in V(G)$.

Corollary 7. If $L(X^0)$ is bounded, the conclusion of the previous theorem is valid.

Proof. From Theorem 4, it follows that the solution $X(t,\ X^0)$ lies in $L(X^0)$ for $t\in[0,\ M)$. Then Theorem 5 applies to give the desired result.

It should be noted that the assumption is satisfied when $V(G)$ is bounded in R^{n+s}.

Corollary 8. If $C_f\cap K$ consists of isolated critical points of f, besides the assumption of Theorem 5 being satisfied, then the solution $X(t,\ X^0)$ converges to a critical point of f as t tends to infinity, *i.e.*

$$\lim_{t\to\infty}X(t,\ X^0)=\tilde{X}\in C_f\ .$$

Proof. The result is easy to establish, so the proof is omitted.

It should be noted here that a regular critical point is an isolated point, where a critical point of f is 'regular,' if its Hessian $\partial^2 f$ is regular at this point. (See Matsushima [5].)

Let the distance of two subset A, B of R^{n+s} be denoted by $\rho(A, B)$, and let a neighbourhood $U_{A,\delta}$ of A be defined by $U_{A,\delta}\equiv\{X;\ \rho(X, A)<\delta\}$.

A compact set $R\subset C_f$ of critical points of the dynamical system (1) is 'stable,' if for each $\lambda>0$ there is $\mu\leqq\lambda$ such that any trajectory which starts from a point in $V(G)\cap U_{R,\mu}$ remains in $V(G)\cap U_{R,\lambda}$; 'asymptotically stable,' if it is stable and in addition for each $\varepsilon>0$ there are τ and δ such that

$$X(t, X^0)\in V(G)\cap U_{R,\varepsilon} \text{ for any } t>\tau$$
$$\text{and any } X^0\in V(G)\cap U_{R,\delta}.$$

A solution $X(t, X^0)$ of the dynamical system (1) is 'quasi-asymptotically stable,' if for each ε, there are τ and δ such that

$$||X(t, X^0)-X(t, X))||<\varepsilon \text{ for any } t>\tau, \text{ and any } X\in U_{X^0,\delta}.$$

Theorem 9. Let $\hat{C}$ be a compact connected component of C_f such that

$$\hat{C}\cap\overline{(C_f-\hat{C})}=\phi.$$

If the function f is maximal on the set $\hat{C}$, then $\hat{C}$ is asymptotically stable set of the system (1).

Proof. Let $f(X)=m$ on $\hat{C}$. There exists a relatively compact neighbourhood $U_{\hat{C}}\subset V(G)$ of $\hat{C}$ such that

$$\bar{U}_{\hat{C}}\cap\overline{(C_f-\hat{C})}=\phi, \quad \text{and}$$
$$m>f(X) \quad \text{for} \quad X\in\bar{U}_{\hat{C}}-\hat{C}.$$

Let m' be a maximal value of $f(X)$ on the boundary of $U_{\hat{C}}$. Of course $m>m'$. Since $f(X)$ is continuous on $V(G)$, we can choose a neighbourhood $V_{\hat{C}}\subset U_{\hat{C}}$ of $\hat{C}$ such that

$$f(X)>(m+m')/2 \quad \text{for} \quad X\in V_{\hat{C}}.$$

Then for any solution $X(t, X^0)$ of (1) defined in $[0, M)$ with the initial point $X^0\in V_{\hat{C}}$, we have

$$f(X(t, X^0))>(m+m')/2 \quad \text{for} \quad t\in[0, M).$$

Thus the solution of (1) which starts from a point in the neighbourhood $V_{\hat{C}}$ remains in the relatively compact set $U_{\hat{C}}$. Next from Theorem 5, it follows that the solution approaches to a critical set in $\bar{U}_{\hat{C}}$ as $t\rightarrow\infty$. On the other hand,

$$C_f \cap \bar{U}_{\hat{C}} = \hat{C},$$

so that the solution approaches to $\hat{C}$. This completes the proof.

The next corollary follows immediately from the theorem.

Corollary 10. If an isolated critical point p is a maximal point of f on $V(G)$, then the system (1) is asymptotically stable at the point p in the manifold $V(G)$.

Theorem 11. If a point X' is not a maximal point of f on $V(G)$, then the system is not asymptotically stable at the point X'. If we assume further that

$$X' \notin (\overline{C_f - X'}),$$

then the system (1) is unstable at $X' \in V(G)$ in the manifold $V(G)$.

Proof. For any neighbourhood $U_{X'} \subset V(G)$ of X', there exists a point $X'' \in U_{X'}$ such that

$$f(X'') > f(X').$$

By the continuity of f, there exists a relatively compact neighbourhood $W_{X'} \subset U_{X'}$ of X' such that

$$(f(X') + f(X''))/2 > f(X)$$

for $X \in \overline{W}_{X'}$. The solution $X(t, X'')$ with the initial point X'' and defined in $[0, M)$ remains outside of $W_{X'}$, since $f(X(t, X'')$ is a non-decreasing function of t. Thus we have the first part of the theorem. Next by the assumption there exists a relatively compact neighbourhood $V_{X'} \subset V(G)$ of X' such that

$$\bar{V}_{X'} \cap (\overline{C_f - X'}) = \phi.$$

Suppose the point X' is stable, there exists a neighbourhood $U'_{X'} \subset V_{X'}$ of X' such that every solution which starts from a point in $U'_{X'}$ remains in $V_{X'}$. As is seen in the above proof, there exists a point $X'' \in U'_{X'}$ and an open neighbourhood $W'_{X'} \subset U'_{X'}$ such that the solution $X(t, X'')$ is outside of $W'_{X'}$. Since the trajectory $X(t, X'')$, $t \in [0, M)$ lies in the bounded set $V_{X'} - W'_{X'}$, it follows from Theorem 5 that the solution $X(t, X'')$ approaches to a critical set in $\bar{V}_{X'}-W'_{X'}$. But this contradicts the fact that there are no critical points in $\bar{V}_{X'}-W'_{X'}$. Then X' cannot be a stable point.

Proposition 12. If the solution $X(t, X^0)$ which starts from $X^0 \in V(G)$ converges to an isolated maximal point $\hat{X}$ of f on $V(G)$, then the solution $X(t, X^0)$ is a quasi-asymptotically stable solution of (1).

Proof. For any neighbourhood $U_{\hat{X}}$ of $\hat{X}$, by Corollary 10, there exists an open neighbourhood $V_{\hat{X}} \subset U_{\hat{X}}$ such that every solution which starts from a point in $V_{\hat{X}}$ converges to $\hat{X}$ as $t \to \infty$. By the assumption, there exists $\tau \in [0, \infty)$ such that $X(\tau, X^0) \in V_{\hat{X}}$. Since $X(t, X^0)$ is continuous in the argument of $X^0 \in V(G)$, there exists an open neighbourhood $W_{X^0} \subset V(G)$ of X^0 such that

$$X(\tau, X') \in V_{\hat{X}} \quad \text{for} \quad X' \in W_{X^0}.$$

Thus every solution which starts from a point in W_{X^0} converges to $\hat{X}$. This completes the proof.

To illustrate the meaning of Theorems 5–12, we give Fig. 1 which shows schematically the behaviour of trajectories of system (1) on a manifold $V(G)$.

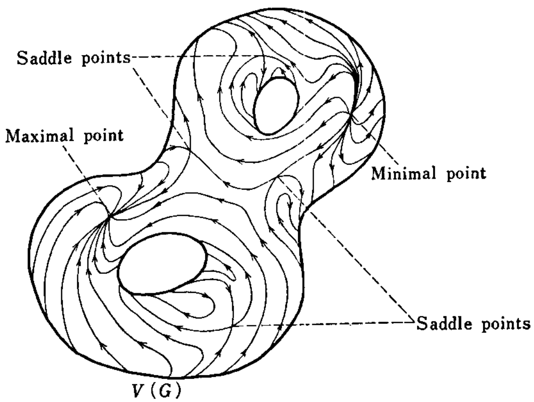

Fig. 1. Flows on $V(G)$.

4. Computational Algorithm

From the foregoing discussion, we can derive an algorithm for nonlinear programming problems. We assume that an initial feasible point X^0 is known.

An algorithm introduced here is an iterative procedure of numerical solution of the system (1) with the initial point X^0 by the Runge-Kutta-Merson method. The reason why this method is adopted is that it is self-starting, hence a change in stepsize is easily effected at any intermediate stage of the calculation and that the evaluation of the derivative of the right-hand-side $\Psi(X)$ of Eq. (1) is not required. It requires only several evaluations of $\Psi(X)$ at each iteration. Theorem 2 shows that $\Psi(X)$ is calculated by solving the linear simultaneous equations (4): First $\Lambda(X)$ is calculated by solving the linear equations of order $r+s$,

$$J_G(X)\nabla f(X)=(J_G(X)J_G^t(X))\Lambda(X)\,. \tag{4.1}$$

Then we obtain $\Psi(X)$, substituting $\Lambda(X)$ into

$$\Psi(X)=\nabla f(X)-J_G^t(X)\Lambda(X)\,. \tag{4.2}$$

But when Eq. (4.1) is of large order, it is difficult to solve it satisfactorily [2]. Besides, even when $J_G(X)$ is sparse $(J_G(X)J_G^t(X))$ is often dense. In this case, therefore, it is unrealistic to evaluate $\Psi(X)$ by solving Eq. (4). Fortunately, we have another useful method for calculating $\Psi(X)$ [7].

Theorem 13. Let mappings $\zeta(\phi)$ from R^{n+s} to R^{n+s} be defined as

$$\zeta_i(\phi)=\phi-\frac{\langle\phi,\nabla g_i(X)\rangle}{\|\nabla g_i(X)\|^2}\nabla g_i(X)\,,\quad i=1,\cdots,r+s\,.$$

Let $Z(\phi)=\zeta_1\circ\zeta_2\circ\cdots\circ\zeta_{r+s}(\phi)$. Then the sequence of vectors $\{\phi^i\}\in R^{n+s}$ generated by the recurrence relation

$$\phi^0=\nabla f(X)\,,\quad \phi^{i+1}=Z(\phi^i)\,,\quad i=0,1,2,\cdots$$

always converges to $\Psi(X)$, *i.e.*

$$\lim_{i\to\infty}\phi^i=\Psi(X)\,.$$

Note that the theorem is valid even when $\nabla g_i(X)$ are linearly dependent. This iterative procedure can make full use of sparseness of $J_G(X)$ to produce $\Psi(X)$ and is effective especially when it is applied to problems of large scale.

It is seen from the form of $J_G(X)$ and Eq. (4) that the introduction of slack variables does not increase the computational work for $\Psi(X)$ seriously.

When the stepsize Δt is small enough, the Runge-Kutta method gives fourth order accuracy in the increment of each component of X at every step of iteration. Generally, numerical solution diverges from the solution $X(t, X^0)$ as iteration proceeds. But this does not matter much, for we need not necessarily give an approximation to the solution $X(t, X^0)$ in the interval $[0, \infty)$ but to produce from one feasible point X^{i-1} another X^i successively by applying one step of the iteration process of the Runge-Kutta method to the solution $X(t, X^{i-1})$ which starts from X^{i-1}. Calculating the values of $g's$ at X^i, we can check the feasibility of X^i at each step of iteration. Hence the stepsize can be chosen larger than usually so long as X^i is feasible.

When X^i becomes infeasible, we can either iterate again from X^{i-1} with a smaller stepsize or correct X^i back into $V(G)$ by minimizing the following function

$$v(X) \equiv \sum_{j=1}^{r+s} g_j^2(X)$$

subject to

$$f(X) = f(X^i) .$$

The correction procedure is to solve numerically the following system which is similar to (1)

$$dX/dt = \nabla v - \langle \nabla v, \Delta f \rangle \Delta f / \|\Delta f\|^2 . \tag{15}$$

If we only keep X^i's in $V(G)$ in this way, Theorems 4–12 suggest that

1) there is little danger for numerical solution X^i to converge to a point which is not a maximal point. (The algorithm never con-

verges to a minimal point except for the case that the initial feasible point is a minimal point.)

2) if $f(X)$ is unimodal with respect to $V(G)$, the algorithm is expected to converge to the global maximum. Then if the original problem is a convex programming, the maximum point will always be found.

Introduction of the squared slack variables, of course, increases the number of critical points of the problem. But this has little effect on the algorithm, because a critical point of the transformed problem which is not a critical point of the original problem cannot be a maximal point of the transformed problem and because a maximal point of one problem exactly coincides with a maximal point of the other.

The algorithm has been coded in FORTRAN IV for HITAC 5020

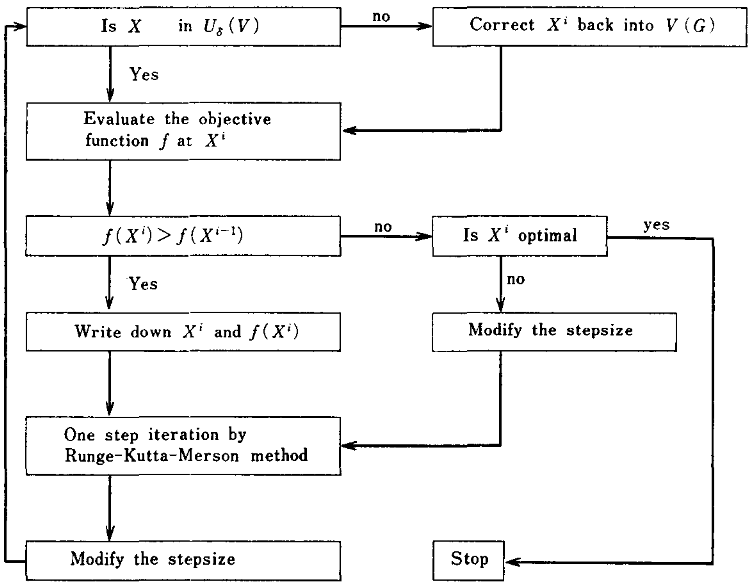

Fig. 2. Flow chart.

computer. The stepsizes are controlled automatically from the information available in the calculation of $\Psi(X)$ at each step of iteration. Only active constraints are involved in the calculation of $\Psi(X)$ at each iteration. Then slack variables are used in the neighbourhood of the edges of inequality constraints to produce a feasible direction $\Psi(X)$.

We give an outline of the algorithm by means of a simplified flow chart (Fig. 2). Let U_δ be a neighbourhood of $V(G)$ in R^{n+s} such that

$$U_\delta \equiv \{X \in R^{n+s};\ ||G(X)||_\infty \leqq \delta\}\ ,$$

where $||*||_\infty$ denotes the maximum among the absolute values of components of *.

5. Computational Experiences

In this section, several numerical examples will be given. All the problems were computed on HITAC 5020 computer.

Example 1. To maximize

$$f=\frac{1}{(x_1+1)^2+x_2^2}$$

subject to

$$g_1=x_1^2+x_2^2-4\geqq 0\ ,$$
$$g_2=16-x_1^2-x_2^2\geqq 0\ .$$

The optimum is at (−2.0, 0.0). The feasible region is not convex. The transformed problem is stationary at (4.0, 0.0), (2.0, 0.0) and (−4.0, 0.0). Starting from the initial feasible point (3.0, 2.5), we obtained the convergent sequence of iterates given in Table 1.

Example 2. To maximize

$$f=x_1^2+x_2^2+x_3^2$$

subject to

$$g_1=x_1^2x_2^2+x_2^2x_3^2+x_3^2x_1^2-x_1x_2x_3=0\ ,\quad x_1x_2x_3\neq 0\ .$$

The optimum is at (1/3, 1/3, 1/3), (−1/3, −1/3, 1/3), (1/3, −1/3, −1/3)

Table 1.

Iteration number	x_1	x_2	$f(X)$	$\|\|G(X)\|\|_\infty$	Δt
1	0.300000 E 1	0.250000 E 1	0.4494382 E−1	0.0000 E−40	0.1000
2	0.299977 1	0.250015 1	0.4494606 −1	0.0000 −40	0.2000
3	0.299931 1	0.250044 1	0.4495054 −1	0.0000 −40	0.4000
4	0.299839 1	0.250102 1	0.4495953 −1	0.0000 −40	0.8000
5	0.299655 1	0.250218 1	0.4497761 −1	0.0000 −40	0.1600 E 1
6	0.299284 1	0.250450 1	0.4501409 −1	0.0000 −40	0.3200 1
7	0.298535 1	0.250911 1	0.4508846 −1	0.0000 −40	0.6400 1
8	0.297006 1	0.251824 1	0.4524297 −1	0.0000 −40	0.1280 2
9	0.293816 1	0.253610 1	0.4557694 −1	0.0000 −40	0.2560 2
10	0.286848 1	0.256990 1	0.4636159 −1	0.0000 −40	0.5120 2
11	0.269927 1	0.262643 1	0.4858430 −1	0.0000 −40	0.1024 3
12	0.215156 1	0.264909 1	0.5899710 −1	0.0000 −40	0.1024 3
13	0.164993 1	0.255978 1	0.7366714 −1	0.0000 −40	0.5120 2
14	0.125709 1	0.246020 1	0.8970995 −1	0.0000 −40	0.2560 2
15	0.986430	0.238781 1	0.1036535	0.0000 −40	0.1280 2
16	0.619730	0.229423 1	0.1267909	0.0000 −40	0.1280 2
17	0.429111	0.216009 1	0.1738021	0.0000 −40	0.1280 2
18	−0.442092	0.202338 1	0.2269977	0.0000 −40	0.6400 1
19	−0.799165	0.187428 1	0.2814320	0.0000 −40	0.3200 1
20	−0.130309 1	0.153464 1	0.4086647	0.0000 −40	0.3200 1
21	−0.163219 1	0.116483 1	0.5693191	0.0000 −40	0.1600 1
22	−0.179822 1	0.881128	0.7074401	0.0000 −40	0.8000
23	−0.192287 1	0.533179	0.8637817	0.0000 −40	0.8000
24	−0.197978 1	0.285082	0.9604012	0.0000 −40	0.8000
25	−0.199563 1	0.132871	0.9911475	0.0000 −40	0.8000
26	−0.199911 1	0.601932 −1	0.9981647	0.0000 −40	0.8000
27	−0.199982 1	0.270983 −1	0.9996253	0.0000 −40	0.1600 1
28	−0.200000 1	0.535574 −2	0.9999803	0.0000 −40	0.1600 1
29	−0.200000 1	0.105821 −2	0.9999989	0.0000 −40	0.1600 1
30	−0.200000 1	0.209085 −3	0.9999990	0.0000 −40	0.3200 1
31	−0.200000 1	0.105051 −3	0.1000004	0.8583 − 5	0.3200 1
32	−0.200000 1	0.527862 −4	0.1000006	0.1144 − 4	0.3200 1

and (−1/3, 1/3, −1/3). Starting from the initial feasible point (0.1, 0.02, −0.00202124), we obtained the sequence given in Table 2. The result shows that the method can well trace the sharply curved surface. It should be noted that although the contour surfaces of the objective function and the constraint are closely attached at the optimum points in the problems of Examples 1 and 2, the method does not suffer zigzaging in the neighbourhood of these points.

Example 3. To maximize

$$f=(x_1-3)^2(4-x_2)$$

subject to

Table 2.

I.N.	x_1	x_2	x_3	$f(X)$	$\|\|G(X)\|\|_\infty$	Δt
1	0.100000	0.200000 E-1	-0.202124 E-2	0.1040408 E-1	0.1065 E-10	0.1000
2	0.110512	0.220781 -1	0.247169 -2	0.1270643 -1	0.6151 -11	0.2000
3	0.134962	0.268812 -1	0.369943 -2	0.1895102 -1	0.1913 - 9	0.4000
4	0.201233	0.396093 -1	0.834129 -2	0.4213332 -1	0.2719 - 7	0.8000
5	0.299786	0.572476 -1	0.191715 -1	0.9351622 -1	0.2500 - 6	0.4000
6	0.365550	0.675110 -1	0.295953 -1	0.1390605	0.3031 - 6	0.2000
7	0.444539	0.763349 -1	0.474607 -1	0.2056945	0.7613 - 6	0.2000
8	0.466053	0.772197 -1	0.542811 -1	0.2261151	0.7698 - 6	0.1000
9	0.486765	0.760000 -1	0.633327 -1	0.2467272	0.8310 - 6	0.5000 E-1
10	0.493676	0.736014 -1	0.690249 -1	0.2538979	0.8059 - 6	0.2500 -1
11	0.494493	0.726753 -1	0.710456 -1	0.2548524	0.8124 - 6	0.1250 -1
12	0.494542	0.725821 -1	0.720728 -1	0.2550346	0.8107 - 6	0.2500 -1
13	0.494424	0.732534 -1	0.732372 -1	0.2551847	0.8000 - 6	0.2500 -1
14	0.494288	0.741717 -1	0.741709 -1	0.2553234	0.8000 - 6	0.5000 -1
15	0.494006	0.760517 -1	0.760545 -1	0.2556099	0.7999 - 6	0.1000
16	0.493410	0.801196 -1	0.797619 -1	0.2562344	0.6895 - 6	0.1000
17	0.493253	0.809608 -1	0.809125 -1	0.2564000	0.6883 - 6	0.2500 -1
18	0.493087	0.819463 -1	0.819393 -1	0.2565643	0.6880 - 6	0.5000 -1
19	0.492743	0.839846 -1	0.839919 -1	0.2569038	0.6886 - 6	0.1000
20	0.492011	0.883638 -1	0.880673 -1	0.2576392	0.5936 - 6	0.1000
21	0.491622	0.903423 -1	0.904495 -1	0.2580350	0.5387 - 6	0.5000 -1
22	0.491204	0.926344 -1	0.926084 -1	0.2584391	0.5339 - 6	0.5000 -1
23	0.490765	0.948935 -1	0.948975 -1	0.2588606	0.5342 - 6	0.1000
24	0.489820	0.996127 -1	0.995740 -1	0.2597616	0.5315 - 6	0.1000
25	0.488786	0.104392	0.104601	0.2607510	0.4571 - 6	0.1000
26	0.488229	0.107017	0.107033	0.2612760	0.4562 - 6	0.5000 -1
27	0.487643	0.109606	0.109608	0.2618231	0.4562 - 6	0.1000
28	0.486385	0.114935	0.114933	0.2629899	0.4558 - 6	0.2000
29	0.483486	0.126201	0.126310	0.2656398	0.4426 - 6	0.2000
30	0.481822	0.132267	0.132246	0.2671362	0.4394 - 6	0.1000
31	0.479998	0.138486	0.138487	0.2687556	0.4405 - 6	0.2000
32	0.475819	0.151636	0.151631	0.2723888	0.4407 - 6	0.4000
33	0.464935	0.180417	0.180659	0.2813523	0.3806 - 6	0.4000
34	0.458003	0.196115	0.196110	0.2866866	0.3740 - 6	0.2000
35	0.449948	0.212211	0.212211	0.2925199	0.3730 - 6	0.4000
36	0.430358	0.244818	0.244818	0.3050799	0.3565 - 6	0.4000
37	0.406940	0.275201	0.275201	0.3070714	0.3118 - 6	0.4000
38	0.382725	0.299608	0.299608	0.3260084	0.2893 - 6	0.4000
39	0.362291	0.315877	0.315877	0.3308112	0.3197 - 6	0.4000
40	0.348481	0.324962	0.324962	0.3326389	0.3272 - 6	0.4000
41	0.340687	0.329467	0.329467	0.3331653	0.3129 - 6	0.4000
42	0.336761	0.331575	0.331575	0.3332925	0.3065 - 6	0.4000
43	0.334899	0.332538	0.332538	0.3333205	0.3029 - 6	0.8000
44	0.333644	0.333174	0.333174	0.3333277	0.2912 - 6	0.8000
45	0.333393	0.333300	0.333300	0.3333280	0.2975 - 6	0.8000
46	0.333359	0.333317	0.333317	0.3333281	0.2955 - 6	0.8000

$$g_1=5-x_1^2-x_2^2\geqq 0$$
$$g_2=5-(x_1-1)^2-x_2^2\geqq 0$$
$$g_3=5-x_1^2-(x_2-1)^2\geqq 0$$
$$g_4=x_1\geqq 0$$
$$g_5=x_2\geqq 0\,.$$

The optimum is at $(\sqrt{19}/2,\ 1/2)\doteqdot(2.179447,\ 0.5)$. Starting from the initial feasible point (0.5, 2.0), we obtained the sequence given in Table 3.

Table 3.

Iteration number	x_1	x_2	$f(X)$	$\|\|G(X)\|\|_\infty$	Δt
1	0.500000	0.200000 E 1	−0.1250000 E 2	0.0000 E−40	0.1000
2	0.879401	0.197661 1	−0.9099079 1	0.0000 −40	0.1000
3	0.116447 E 1	0.187023 1	−0.7175546 1	0.0000 −40	0.1000
4	0.139033 1	0.172993 1	−0.5881814 1	0.0000 −40	0.1000
5	0.156826 1	0.158094 1	−0.4958768 1	0.0000 −40	0.2000
6	0.180982 1	0.130754 1	−0.3813944 1	0.0000 −40	0.2000
7	0.194848 1	0.109424 1	−0.3212893 1	0.0000 −40	0.2000
8	0.202925 1	0.937836	−0.2885658 1	0.0000 −40	0.4000
9	0.210904 1	0.742651	−0.2585691 1	0.0000 −40	0.4000
10	0.214307 1	0.638102	−0.2468716 1	0.0000 −40	0.4000
11	0.215951 1	0.580086	−0.2415890 1	0.0000 −40	0.8000
12	0.217290 1	0.527769	−0.2375338 1	0.3147 − 4	0.8000
13	0.217718 1	0.509812	−0.2362953 1	0.3529 − 4	0.1600 E 1
14	0.217917 1	0.501233	−0.2357307 1	0.3815 − 4	0.1600 1
15	0.217942 1	0.500172	−0.2356623 1	0.3719 − 4	0.1600 1
16	0.217945 1	0.500037	−0.2356541 1	0.3338 − 4	0.3200 1
17	0.217945 1	0.500012	−0.2356525 1	0.3338 − 4	0.3200 1

6. Concluding Remark

The algorithm works well for both equality and inequality constrained maximization problem. Optimality will be obtained from a very poor approximation to the desired point. Hunting does not occur in the neighbourhood of optimal point.

Acknowledgement

The author would like to thank Dr. M. Shibuya for helpful suggestions for making computer programs.

References

[1] Bhatia, N. P. and G. P. Szegö, *Dynamical Systems: Stability Theory and Applications*, Springer Verlag, 1967.

[2] Forsythe, G. E. and C. B. Moler, *Computer Solution of Linear Algebraic Systems*, Prentice Hall, 1967.

[3] Fox, L., *Numerical Solution of Ordinary and Partial Differential Equations*, Pergamon Press, 1962.

[4] Hancock, H., *Theory of Maxima and Minima*, Gin and Company, 1917, Dover Publications, 1960.

[5] Matsushima, Y., *Introduction to Manifold* (in Japanese), Shyokabo, 1965.

[6] Rosen, J. B., "The Gradient Projection Method for Nonlinear Programming. Part II, Nonlinear Constraints," *J. Soc. Indust. Appl. Math.*, **9**, 4 (1961).

[7] Tanabe, K., "Projection Method for Solving a Singular System of Linear Equations and its Applications," *Numerische Mathematik*, **17** (1971), 203-214.

[8] Tanabe, K., "Algorithm for the Constrained Maximization Technique for Nonlinear Programming," *Proceedings of the Second Hawaii International Conference on System Science*, 1969, 487-490.

[9] Tanabe, K., "An Algorithm for the Constrained Maximization in Nonlinear Programming," *Research Memorandum*, No. 31, Institute of Statistical Mathematics, 1969.