Journal of the Operations Research
Society of Japan
Vol. 20, No. 2, June, 1977

AN ITERATION METHOD FOR NONLINEAR PROGRAMMING PROBLEMS: II

HISASHI MINE and KATSUHISA OHNO, *Kyoto University*
TATSUO NODA, *Toyama College of Technology*

(Received July 26, 1976; Revised December 20, 1976)

Abstract This note proposes a simple and practical iteration method for finding a local minimum of a nonlinear programming problem with inequality and equality constraints. The iteration method seeks a point which satisfies the Kuhn-Tucker conditions. It can be shown that the sequence of points generated by the iteration method converges to the local optimal solution.

1. Introduction

Let R^n be the n-dimensional Euclidean space, and let $f(x)$, $h_i(x)$ $(i=1,2,\ldots,m)$ and $g_j(x)$ $(j=1,2,\ldots,\ell)$ be real-valued functions defined on R^n.

Consider the following nonlinear programming problem:

(P)
$$\begin{aligned} &\text{Minimize } f(x), \\ &\text{subject to} \\ &h_i(x) \leqq 0 \quad (i=1,2,\ldots,m) \\ &\text{and} \\ &g_j(x) = 0 \quad (j=1,2,\ldots,\ell). \end{aligned}$$

It is assumed that f, h_i and g_j are three times continuously differentiable on R^n.

Mine, Ohno and Noda [3] discuss Problem (P) and devise a simple and practical iteration method for solving (P). Note that in [3], inequality constraints are reduced to equality constraints by introducing slack variables.

Luenberger [2] deals with a nonlinear programming problem with inequality constraints and converts it to an unconstrained minimization problem. In [2], a computational procedure based on the conjugate residual scheme is

applied to the unconstrained minimization problem.

In this note, Problem (P) is transformed into an unconstrained minimization problem which has more tractable form than the unconstrained problem in [2]. The present unconstrained minimization problem contains no slack variables for inequality constraints and the present iteration method is different from the one in [3].

§2 proposes the iteration method and shows its local convergence. §3 notes some remarks on the method and §4 gives a numerical example.

2. Algorithm

Let

$$x = (x_1, x_2, \ldots, x_n),$$

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m),$$

$$\mu = (\mu_1, \mu_2, \ldots, \mu_\ell),$$

$$h(x) = (h_1(x), h_2(x), \ldots, h_m(x))$$

and

$$g(x) = (g_1(x), g_2(x), \ldots, g_\ell(x)).$$

Superscript * is used to denote transposition. Then, the Lagrangian function $\phi(x, \lambda, \mu)$ associated with Problem (P) is:

$$\phi(x, \lambda, \mu) = f(x) + \lambda\, h(x)^* + \mu\, g(x)^*.$$

Denote by $\partial h(x)/\partial x$ and $\partial g(x)/\partial x$ the $m\times n$ and $\ell\times n$ Jacobian matrices with (i, j) components $\partial h_i(x)/\partial x_j$ and $\partial g_i(x)/\partial x_j$, respectively. Let ϕ_x and ϕ_{xx} be the gradient row vector with component $\partial\phi/\partial x_i$ and the Hessian matrix with (i, j) component $\partial^2\phi/\partial x_i \partial x_j$, respectively. Define an $(n+m+\ell)$-dimensional vector $y(x, \lambda, \mu)$ and an $(n+m+\ell)\times(n+m+\ell)$ matrix $A(x, \lambda, \mu)$ as follows:

$$y(x, \lambda, \mu) = (\phi_x(x, \lambda, \mu),\ h(x)(\mathrm{diag}(\lambda)),\ g(x))$$

and

$$A(x, \lambda, \mu) = \begin{pmatrix} \phi_{xx}(x, \lambda, \mu) & (\partial h(x)/\partial x)^* & (\partial g(x)/\partial x)^* \\ \mathrm{diag}(\lambda)(\partial h(x)/\partial x) & \mathrm{diag}\ h(x) & 0 \\ \partial g(x)/\partial x & 0 & 0 \end{pmatrix},$$

where $\mathrm{diag}(\lambda)$ is the diagonal matrix with the i th diagonal component λ_i. Define the Euclidean norms for an n-dimensional vector x and an $m\times n$ matrix

$B = (b_{ij})$ as follows:

$$||x|| = \left(\sum_{j=1}^{n} x_j^2\right)^{1/2}$$

and

$$||B|| = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} b_{ij}^2\right)^{1/2}.$$

In the following, we suppose the second-order sufficient conditions under which a point $\bar{x}$ is an isolated local minimum of Problem (P).

Sufficient conditions (Fiacco-McCormick [1, page 30]):

(1) $h(\bar{x}) \leqq 0$,

(2) $g(\bar{x}) = 0$,

(3) $h(\bar{x})(\mathrm{diag}(\bar{\lambda})) = 0$,

(4) $\bar{\lambda}_i > 0$ for all $i \in \bar{B} = \{i;\ h_i(\bar{x}) = 0\}$,

(5) $\phi_x(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$,

(6) $u\ \phi_{xx}(\bar{x}, \bar{\lambda}, \bar{\mu})\ u^* > 0$ for every nonzero vector u

satisfying

$$u\ (h_i(\bar{x}))_x^* = 0 \quad \text{for } i \in \bar{B}$$

and

$$u\ (g_j(\bar{x}))_x^* = 0 \quad \text{for } j = 1, 2, \ldots, \ell.$$

In addition, if

(7) the vectors

$$\{(h_i(\bar{x}))_x;\ i \in \bar{B}\},\ \{(g_j(\bar{x}))_x;\ j = 1, 2, \ldots, \ell\}$$

are linearly independent,

then $A(\bar{x}, \bar{\lambda}, \bar{\mu})$ is nonsingular (Fiacco-McCormick [1, page 80]).

In order to simplify the notations, denote by z an $(n+m+\ell)$-dimensional vector (x, λ, μ) and by $\bar{z}$ the triple $(\bar{x}, \bar{\lambda}, \bar{\mu})$ which satisfies the above sufficient conditions.

Further define $E(z)$ as

$$E(z) = ||\phi_x(z)||^2 + \sum_{i=1}^{m} (\lambda_i h_i(x))^2 + \sum_{j=1}^{\ell} (g_j(x))^2.$$

In order to find $\bar{z}$, we minimize $E(z)$ by the iteration method similar to the one in [3].

The proposed algorithm is:

Step 1: Set $k = 0$ and choose an initial point $z^{(0)}$, a certain value $\varepsilon > 0$, and a constant α such that $0 < \alpha < 2$.

Step 2: Calculate $z^{(k+1)}$ by the formula

$$z^{(k+1)} = z^{(k)} - \alpha \, ||A(z^{(k)})||^{-2} \, y(z^{(k)}) \, A(z^{(k)}).$$

Step 3. Stop if $|z_j^{(k+1)} - z_j^{(k)}| < \varepsilon$ for $j = 1, 2, \ldots, n+m+\ell$.

Otherwise, set $k = k+1$ and return to Step 2.

The following theorem shows the local convergence of the algorithm. The proof is the same as in [3] and omitted.

Theorem. If $\bar{z}$ satisfies conditions (1)-(7), then there exists a neighbourhood $U(\bar{z})$ such that for any starting point $z^{(0)} \in U(\bar{z})$ the sequence $z^{(k)}$ remains in $U(\bar{z})$ and converges to $\bar{z}$.

This theorem means that $\bar{z}$ is a point of attraction of the proposed algorithm (see, Ortega-Rheinboldt [4, page 299]).

3. Remarks

Remark 1. From an analogous argument to the one in [3, page 142-143], it follows that

$$||z^{(k+1)} - \bar{z}|| < K \, ||z^{(k)} - \bar{z}||,$$

where

$$K = \max (|1 - L|, \, |1 - M|),$$

$$L = \frac{\alpha}{||A(\bar{z})||^2 + \varepsilon} \left(\min_{||\rho||=1} ||\rho(A(\bar{z}))^*||^2 - \varepsilon \right),$$

and

$$M = \frac{\alpha(||A(\bar{z})||^2 + 2\varepsilon)}{||A(\bar{z})||^2 - \varepsilon}.$$

Since $0 < L < M < 2$, we have $0 < K < 1$. Note that as the value of K becomes small, the sequence $z^{(k)}$ may converge more rapidly to $\bar{z}$. Therefore, the parameter α may be chosen so that K becomes small. In general, ε can be neglected when $z^{(k)}$ is sufficiently close to $\bar{z}$.

Let

$$r = \min_{||\rho||=1} ||\rho(A(\bar{z}))^*||^2 / ||A(\bar{z})||^2.$$

Then $0 < r \leq 1$ and we have the following approximations to L, M and K:

$$L \sim \alpha \, r,$$

$$M \sim \alpha$$

and

$$K \sim R(\alpha) \equiv \max\ (|\alpha\, r - 1|,\ |\alpha - 1|).$$

Since $0 < r \leq 1$, $R(\alpha)$ attains the minimum at $\alpha = \dfrac{2}{r+1}$ $\left(1 \leq \dfrac{2}{r+1} < 2\right)$.

Remark 2. Let us discuss the operation count, that is, the number of multiplications and divisions involved within one iteration of the algorithm. The operation count for computing $||A(z^{(k)})||^2$ is

$$\frac{n(n+1)}{2} + n\,m + 3\,m + \ell\,n,$$

and that for computing $y(z^{(k)})A(z^{(k)})$ is

$$n(n + m + \ell) + m(n + 1) + \ell\,n.$$

Therefore the operation count of one iteration is

$$\frac{3}{2}\,n^2 + \frac{3}{2}\,n + 3\,n\,m + 3\,n\,\ell + 5\,m + \ell.$$

Similarly, the operation count of one iteration of the algorithm in [3] is

$$\frac{3}{2}\,(n+m)^2 + \frac{3}{2}\,(n+m) + 3(n+m)m + 3(n+m)\ell + 5\,m + \ell.$$

It is clear that the operation count of the present algorithm is much less than that of the previous algorithm [3].

4. Numerical Example

As in [3], the Rosen-Suzuki Problem was solved as a numerical example.

The Rosen-Suzuki Test Problem [5]: Minimize

$$f(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4,$$

subject to

$$h_1(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8 \leq 0,$$

$$h_2(x) = x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10 \leq 0,$$

$$h_3(x) = 2x_1^2 + x_2^2 + x_3^2 + 2x_1 - x_2 - x_4 - 5 \leq 0.$$

Minimum point $\bar{x}$ and the minimum value $f(\bar{x})$ are given by

$$\bar{x} = (0,\ 1,\ 2,\ -1)$$

and

$$f(\bar{x}) = -44.$$

Computations with $\varepsilon = 10^{-4}$ were carried out on FACOM 230-75 computer of Kyoto University Computation Center. These results are shown in the following tables.

Tables. Computation Results for the Rosen-Suzuki Problem.

$z^{(0)} = (0.0, 0.0, \ldots, 0.0)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec.)
0.9	0.01127	1.01824	1.99161	-0.99047	-43.9573	1.1
1.3	0.00389	1.01884	1.99253	-0.99652	-43.9609	1.0
1.9	-0.00438	1.01958	1.99793	-1.00180	-44.0185	0.9

$z^{(0)} = (1.1, 1.1, \ldots, 1.1)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec.)
0.9	0.02265	1.01639	2.00811	-0.95063	-44.0189	3.0
1.3	0.01723	1.01737	2.00514	-0.96167	-44.0114	2.8
1.9	0.01171	1.01607	2.00263	-0.97292	-44.0045	2.2

$z^{(0)} = (1.2, 1.2, \ldots, 1.2)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec.)
0.9	0.01398	0.99994	2.00692	-0.96663	-43.9912	0.7
1.3	0.00030	1.01096	1.99807	-0.99659	-43.9917	0.5
1.9	-0.00912	1.01772	1.99197	-1.00250	-44.0013	0.4

Acknowledgement

The authors are indebted to Mr. M. Fukushima of Kyoto University for helpful discussions and suggestions. They also would like to thank the referees for their valuable comments.

References

[1] Fiacco, A.V., and McCormick, G.P.: *Nonlinear Programming:Sequential Unconstrained Minimization Techniques*, John Wiley, New York, 1968.

[2] Luenberger, D.G.: An Approach to Nonlinear Programming. *Journal of Optimization Theory and Applications*, Vol. 11 (1973), pp. 219-227.

[3] Mine, H., Ohno, K., and Noda, T.: An Iteration Method for Nonlinear Programming Problems. *Journal of the Operations Research Society of Japan*, Vol. 19 (1976), pp. 137-146.

[4] Ortega, J.M., and Rheinboldt, W.C.: *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[5] Rosen, J.B., and Suzuki, S.: Construction of Nonlinear Programming Test Problems. *Communications of the ACM*, Vol. 8 (1965), p. 113.

Hisashi MINE and Katsuhisa OHNO:
Department of Applied Mathematics
and Physics, Faculty of Engineering
Kyoto University, Kyoto, Japan

Tatsuo NODA:
Department of Applied Mathematics
Toyama College of Technology
Kurokawa, Kosugi-machi, Toyama
Japan