Iteration Methods for Solving Nonlinear Programming Problems

AUTHOR(S):

MINE, Hisashi; OHNO, Katsuhisa; NODA, Tatsuo

CITATION:

MINE, Hisashi ...[et al]. Iteration Methods for Solving Nonlinear Programming Problems. Memoirs of the Faculty of Engineering, Kyoto University 1978, 39(4): 535-547

ISSUE DATE:

1978-01-31

URL:

http://hdl.handle.net/2433/281054

Iteration Methods for Solving Nonlinear Programming Problems

By

Hisashi MINE,* Katsuhisa OHNO* and Tatsuo NODA**

(Received June 29, 1977)

Abstract

This paper proposes simple and practical iteration methods for finding an optimal solution of a nonlinear programming problem with inequality and equality constraints. The iteration methods seek a point which satisfies the Kuhn-Tucker conditions. It is shown that the sequence of points generated by the iteraion methods converges to the optimal solution. Numerical results show the efficency of the proposed methods.

1. Introduction

Let R^n be the n-dimensional Euclidean space, and let $f(x)$, $h_i(x)$ $(i=1, 2, \dots, m)$ and $g_j(x)$ $(j=1, 2, \dots, l)$ be real-valued functions defined on R^n. Let us consider the following nonlinear programming problem:

(P) Minimize $f(x)$,
subject to
$h_i(x) \leqq 0 \quad (i=1, 2, \dots, m)$
and
$g_j(x)=0 \quad (j=1, 2, \dots, l)$.

This paper improves the iteration method for finding the optimal solution of (P) proposed in the previous paper.[4,5] Throughout this paper, it is assumed that the functions f, $h_i(i=1, 2, \dots, m)$ and $g_j(j=1, 2, \dots, l)$ are three times continuously differentiable on R^n.

Section 2 shows the Kuhn-Tucker conditions for (P) and devises an iteration method for finding an optimal solution of (P), so that the Kuhn-Tucker conditions are satisfied. In Section 3 are given several lemmas which are used in proving the

* Department of Applied Mathematics and Physics.
** Department of Applied Mathematics, Toyama College of Technology, Toyama.

local convergence of the proposed method in Section 4. A modified version of the method is also given in Section 5. As a numerical example, the Rosen-Suzuki Problem[7] is solved by the proposed and modified methods presented in Section 6.

2. Iteration method

Let

$x=(x_1, x_2, \ldots, x_n)$ be an n-dimensional vector,
$\lambda=(\lambda_1, \lambda_2, \ldots, \lambda_m)$ an m-dimensional vector,

and

$\mu=(\mu_1, \mu_2, \ldots, \mu_l)$ an l-dimensional vector.

Define m-dimensional and l-dimensional vector-valued functions h and g as follows:

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

and

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

Then, the Lagrangian function $\phi(x, \lambda, \mu)$ associated with Problem (P) is

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

where superscript * denotes transposition. Denote by $\partial h(x)/\partial x$ and $\partial g(x)/\partial x$ the $m\times n$ and $l\times n$ 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 components $\partial\phi/\partial x_i$ and the Hessian matrix with (i, j) component $\partial^2\phi/\partial x_i\partial x_j$, respectively.

In the following, the Kuhn-Tucker conditions[1] are introduced, under which point $\bar{x}$ is an optimal solution of Problem (P).

The Kuhn-Tucker conditions[1]:

$$h(\bar{x})\leq 0, \qquad \cdots\cdots\cdots\cdots(1)$$
$$g(\bar{x})=0, \qquad \cdots\cdots\cdots\cdots(2)$$
$$h(\bar{x})(\mathrm{diag}(\bar{\lambda}))=0, \qquad \cdots\cdots\cdots\cdots(3)$$
$$\bar{\lambda}_i>0 \quad \text{for all} \quad i\in\bar{B}=\{i;\ h_i(\bar{x})=0\}, \qquad \cdots\cdots\cdots\cdots(4)$$
$$\phi_x(\bar{x}, \bar{\lambda}, \bar{\mu})=0, \qquad \cdots\cdots\cdots\cdots(5)$$

and

$$v\phi_{xx}(\bar{x}, \bar{\lambda}, \bar{\mu})v^*>0, \qquad \cdots\cdots\cdots\cdots(6)$$

for every non-zero vector v satisfying $v(h_i(\bar{x}))_x^*=0$ for $i\in\bar{B}$ and $v(g_j(\bar{x}))_x^*=0$ for $j=1, 2, \ldots, l$.

In addition, suppose that the vectors

$\{(h_i(\bar{x}))_x ; i \in \bar{B}\}$, $\{(g_j(\bar{x}))_x ;\ j=1, 2, \ldots, l\}$ are linearly independent. $\cdots\cdots\cdots\cdots(7)$

In order to simplify the notation, denote by z the $(n+m+l)$-dimensional vector (x, λ, μ) and by $\bar{z}$ the triple $(\bar{x}, \bar{\lambda}, \bar{\mu})$ which satisfies the above conditions. Define the $(n+m+l)$-dimensional vector $y(z)$ and the $(n+m+l)\times(n+m+l)$ matrix $A(z)$ as follows:

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

and

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

where $\mathrm{diag}(\lambda)$ is the diagonal matrix with the i th diagonal component λ_i. Put

$$A(z)=(A_1(z)\ \ A_2(x)),$$

where the $(n+m+l)\times n$ matrix $A_1(z)$ and the $(n+m+l)\times(m+l)$ matrix $A_2(x)$ are as follows:

$$A_1(z)=\begin{pmatrix} \phi_{xx}(z) \\ \mathrm{diag}(\lambda)(\partial h(x)/\partial x) \\ \partial g(x)/\partial x \end{pmatrix},$$

$$A_2(x)=\begin{pmatrix} (\partial h(x)/\partial x)^* & (\partial g(x)/\partial x)^* \\ \mathrm{diag}\ h(x) & 0 \\ 0 & 0 \end{pmatrix}.$$

The proposed method in the previous paper[5] is based on a method for minimizing $E(z)$ given by

$$E(z)=\|\phi_x(z)\|^2+\sum_{i=1}^{m}(\lambda_i h_i(x))^2+\sum_{j=1}^{l}(g_j(x))^2. \qquad \cdots\cdots\cdots\cdots(8)$$

The Kuhn-Tucker conditions imply that if $\bar{z}$ satisfies $E(\bar{z})=0$, (1) and (4), then $\bar{z}$ is an optimal solution of (P). The previous iteration method is given by

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

where

$$\|A(z)\|_E^2=\sum_{i,j=1}^{n+m+l}(a_{ij}(z))^2$$

and α is a constant satisfying $0<\alpha<2$.

In this paper, we consider the submatrices $A_1(z)$ and $A(x_2)$ instead of $A(z)$. Given an initial point $x^{(0)}$, we can find the Lagrange multiplier $w^{(1)}=(\lambda^{(1)}, \mu^{(1)})$ corresponding to $x^{(0)}$ by minimizing $E(x^{(0)}, w)$. This $w^{(1)}$ can be obtained by solving the system of $(m+l)$ linear equations

$$A_2(x^{(0)})^*A_2(x^{(0)})w^*=-A_2(x^{(0)})^*(f_x(x^{(0)}), 0, g(x^{(0)}))^*.$$

With $w^{(1)}$, an improved point $x^{(1)}$ is then determined by minimizing $E(x, w^{(1)})$. Summarizing this procedure, we obtain the following iteration method for solving (P) :

Step 1: Let $x^{(0)}$ be given. Set $k=0$ and choose a positive number $\varepsilon>0$.

Step 2: Solve the system of $(m+l)$ linear equations
$A_2(x^{(k)})^*A_2(x^{(k)})w^*=-A_2(x^{(k)})^*(f_x(x^{(k)}), 0, g(x^{(k)}))^*$
and set $w^{(k+1)}=(\lambda^{(k+1)}, \mu^{(k+1)})$ as the solution.

Step 3: Find $x^{(k+1)}$ which minimizes $E(x, w^{(k+1)})$. Stop if $\max_j|x_j^{(k+1)}-x_j^{(k)}|<\varepsilon$.
Otherwise, set $k=k+1$ and return to Step 2.

Remark. Since many computational methods for solving unconstrained minimization problems are available[2,3,4,5], any of these methods can be applied for finding $x^{(k+1)}$ which minimizes $E(x, w^{(k+1)})$ in Step 3. For example, the previous iteration method leads to the following algorithm:

Set $\hat{x}^{(0)}=x^{(k)}$ and $p=0$. Calculate $\hat{x}^{(p+1)}$ by

$$\hat{x}^{(p+1)}=\hat{x}^{(p)}-\frac{\alpha}{\|A_1(\hat{x}^{(p)}, w^{(k+1)})\|_E^2}\,y(\hat{x}^{(p)}, w^{(k+1)})A_1(\hat{x}^{(p)}, w^{(k+1)})$$

with the stopping criterion $\max_j|\hat{x}_j^{(p+1)}-\hat{x}_j^{(p)}|<\varepsilon_1$, where ε_1 is suitably chosen for the initial point $x^{(k)}$ and α is a constant such that $0<\alpha<2$.

And set $x^{(k+1)}=\hat{x}^{(p+1)}$.

3. Preliminaries

Denote by $\|x\|$ and $\|A\|$ the Euclidean norm and the corresponding matrix norm, i. e.,

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

and

$$\|A\|=\rho^{1/2},$$

where ρ is the maximum eigenvalue of A^*A.

First, the following lemma holds.

Lemma 1. If conditions (1)-(7) are satisfied, then there exist neighbourhoods $V_1(\bar{z})$ and $V_2(\bar{x})$ such that

$$\text{rank } A(z) = n+m+l, \quad z \in V_1(\bar{z}),$$
$$\text{rank } A_1(z) = n, \quad z \in V_1(\bar{z}),$$

and

$$\text{rank } A_2(x) = m+l, \quad x \in V_2(\bar{x}).$$

Proof. From Fiacco-McCormick[1], it follows that rank $A(\bar{z}) = n+m+l$. It is clear by (7) that

$$\text{rank } A_2(\bar{x}) = m+l.$$

Therefore

$$\text{rank } A_1(\bar{z}) = n.$$

The continuity of f, h_i $(i=1, 2, \ldots, m)$ and g_j $(j=1, 2, \ldots, l)$ implies the desired result.

Now define an $n \times n$ matrix $C(z)$ by

$$C(z) = \sum_{k=1}^{n} \frac{\partial \phi}{\partial x_k}\left(\frac{\partial \phi(z)}{\partial x_k}\right)_{xx} + \sum_{i=1}^{m} \lambda_i h_i(x) \lambda_i (h_i(x))_{xx}$$
$$+ \sum_{j=1}^{l} g_j(x) (g_j(x))_{xx}.$$

Then we have the following lemma.

Lemma 2. Under the same conditions as in Lemma 1,

$$\det E_{xx}(z) \neq 0, \quad z \in V_1(\bar{z})$$

holds.

Proof. From conditions (2), (3) and (5), it follows that

$$E_{xx}(\bar{z}) = 2[A_1(\bar{z})^* A_1(\bar{z}) + C(\bar{z})]$$
$$= 2A_1(\bar{z})^* A_1(\bar{z}).$$

Since $A_1(\bar{z})$ is an $(n+m+l) \times n$ matrix, Lemma 2 follows from Lemma 1.

Conditions (2), (3) and (5) show

$$E_x(\bar{z}) = 2[\phi_x(\bar{z}) \phi_{xx}(\bar{z}) + \sum_{i=1}^{m} (\bar{\lambda}_i h_i(\bar{x})) \bar{\lambda}_i (h_i(\bar{x}))_x$$
$$+ \sum_{j=1}^{l} g_j(\bar{x}) (g_j(\bar{x}))_x]$$
$$= 0. \qquad \cdots\cdots\cdots\cdots (9)$$

Consequently Lemma 2, (9) and the implicit function theorem[6] imply that the equation

$$E_x(z) = E_x(x, w) = 0$$

has a unique solution

$$x=\psi(w),\ (x,\ w)\in V(\bar{x})\times W(\bar{w}),$$

where $V(\bar{x})$ and $W(\bar{w})$ are neighbourhoods of $\bar{x}$ and $\bar{w}$, respectively.

Since (8) can be rewritten as

$$E(x,\ w)=\| A_2(x)w^*+(f_x(x),\ 0,\ g(x))^*\|^2, \qquad \cdots\cdots\cdots\cdots(10)$$

Lemma 1 implies that

$$w=-(f_x(x),\ 0,\ g(x))A_2(x)(A_2(x)^*A_2(x))^{-1}$$

minimizes the value of (10) for an arbitrarily fixed $x\in V_2(\bar{x})$. Define the $(m+l)$-dimensional vector $u(x)$ and the $(m+l)\times(m+l)$ matrix $A_3(x)$ by

$$u(x)=(f_x(x),\ 0,\ g(x))A_2(x)$$

and

$$A_3(x)=(A_2(x)^*A_2(x))^{-1}.$$

Put

$$L_1=\sup_{x\in V_2(\bar{x})}\| u(x)\| \qquad \cdots\cdots\cdots\cdots(11)$$

and

$$L_2=\sup_{x\in V_2(\bar{x})}\| A_3(x)\|. \qquad \cdots\cdots\cdots\cdots(12)$$

Then the following lemma holds.

Lemma 3. For any $x',\ x''\in V_2(\bar{x})$, there exist $M_1>0$ and $M_2>0$ such that

$$\| u(x'')-u(x')\|\leqq M_1\|x''-x'\| \qquad \cdots\cdots\cdots\cdots(13)$$

and

$$\| A_3(x'')-A_3(x')\|\leqq M_2\|x''-x'\|. \qquad \cdots\cdots\cdots\cdots(14)$$

Proof. First, we shall show (13). By the definition, we have

$$u(x)=(f_x(x)(\partial h(x)/\partial x)^*,\ f_x(x)(\partial g(x)/\partial x)^*).$$

Therefore,

$$\| u(x'')-u(x')\|^2=\| f_x(x'')(\partial h(x'')/\partial x)^*-f_x(x')(\partial h(x')/\partial x)^*\|^2 + \| f_x(x'')(\partial g(x'')/\partial x)^*-f_x(x')(\partial g(x')/\partial x)^*\|^2.$$

Define the m-dimensional vector $p(x)$ and the l-dimensional vector $q(x)$ as follows:

$$p(x)=f_x(x)(\partial h(x)/\partial x)^*,$$
$$q(x)=f_x(x)(\partial g(x)/\partial x)^*.$$

Put

$$M_3 = \sup_{x \in V_2(\bar{x})} \|\partial p(x)/\partial x\|$$

and

$$M_4 = \sup_{x \in V_2(\bar{x})} \|\partial q(x)/\partial x\|.$$

Then it follows that

$$\|u(x'') - u(x')\|^2 \leq (M_3^2 + M_4^2)\|x'' - x'\|^2.$$

Consequently, (13) holds with $M_1 = (M_3^2 + M_4^2)^{1/2}$.

We shall now show (14). From (12), it follows that for an arbitrary $x', x'' \in V_2(\bar{x})$,

$$\begin{aligned}\|A_3(x'') - A_3(x')\| &= \|A_3(x'')(A_3(x')^{-1} - A_3(x'')^{-1})A_3(x')\| \\ &\leq \|A_3(x'')\|\|A_3(x')\|\|A_3(x')^{-1} - A_3(x'')^{-1}\| \\ &\leq L_2^2\|A_3(x')^{-1} - A_3(x'')^{-1}\|.\end{aligned}$$

Denote by $b_{ij}(x', x'')$ the (i, j) component of the $(m+l)\times(m+l)$ matrix $(A_3(x')^{-1} - A_3(x'')^{-1})$.

Since

$$A_3(x')^{-1} - A_3(x'')^{-1} = A_2(x')^* A_2(x') - A_2(x'')^* A_2(x'')$$

is symmetric, the inequality

$$\|A_3(x')^{-1} - A_3(x'')^{-1}\| \leq (m+l) \max_{1 \leq i, j \leq m+l} |b_{ij}(x', x'')| \qquad \cdots\cdots\cdots\cdots (15)$$

holds. (See, for example, Ortega-Rheinboldt.[6])

Now define the $m\times m$ matrix $A_{11}(x)$, the $m\times l$ matrix $A_{12}(x)$ and the $l\times l$ matrix $A_{22}(x)$ as follows:

$$\begin{aligned}A_{11}(x) &= (\partial h(x)/\partial x)(\partial h(x)/\partial x)^* + (\text{diag } h(x))^2, \\ A_{12}(x) &= (\partial h(x)/\partial x)(\partial g(x)/\partial x)^*\end{aligned}$$

and

$$A_{22}(x) = (\partial g(x)/\partial x)(\partial g(x)/\partial x)^*.$$

Then we have

$$A_3(x)^{-1} = \begin{bmatrix} A_{11}(x) & A_{12}(x) \\ A_{12}(x)^* & A_{22}(x) \end{bmatrix}$$

and

$$\max_{1 \leq i, j \leq m+l} |b_{ij}(x', x'')| = \max\Big[\max_{\substack{i \neq j \\ 1 \leq i, j \leq m}} |h_i(x'))_x (h_j(x'))_x^* - (h_i(x''))_x (h_j(x''))_x^*|,$$

$$\max_{1\leq i\leq m} \; |\|(h_i(x'))_x\|^2-\|(h_i(x''))_x\|^2+(h_i(x'))^2 - (h_i(x''))^2|,$$

$$\max_{\substack{1\leq i\leq m\\ 1\leq j\leq l}} \; |(h_i(x'))_x(g_j(x'))_x^* -(h_i(x''))_x(g_j(x''))_x^*|,$$

$$\max_{1\leq i,\, j\leq l} \; |(g_i(x'))_x(g_j(x'))_x^* - (g_i(x''))_x(g_j(x''))_x^*|]. \quad \cdots\cdots\cdots\cdots (16)$$

Moreover, define the real-valued functions $p_{ij}(x)$, $p_i(x)$, $q_{ij}(x)$ and $r_{ij}(x)$ as follows:

$$p_{ij}(x)=(h_i(x))_x(h_j(x))_x^* \quad (i\neq j),$$
$$p_i(x)=\|(h_i(x))_x\|^2+(h_i(x))^2,$$
$$q_{ij}(x)=(h_i(x))_x(g_j(x))_x^*$$

and

$$r_{ij}(x)=(g_i(x))_x(g_j(x))_x^*.$$

Put

$$L_{ij}=\sup_{x\in V_2(\bar{x})}\|(p_{ij}(x))_x\|,$$
$$L_i=\sup_{x\in V_2(\bar{x})}\|(p_i(x))_x\|,$$
$$M_{ij}=\sup_{x\in V_2(\bar{x})}\|(q_{ij}(x))_x\|,$$
$$N_{ij}=\sup_{x\in V_2(\bar{x})}\|(r_{ij}(x))_x\|.$$

Then, from (15) and (16)

$$\|A_3(x')^{-1}-A_3(x'')^{-1}\|\leq(m+l)[\max(\bar{L}_{ij}, \bar{L}_i, \bar{M}_{ij}, \bar{N}_{ij})]\|x'-x''\|$$

holds, where

$$\bar{L}_{ij}=\max_{\substack{i\neq j\\ 1\leq i,\, j\leq m}} L_{ij},$$
$$\bar{L}_i=\max_{1\leq i\leq m} L_i,$$
$$\bar{M}_{ij}=\max_{\substack{1\leq i\leq m\\ 1\leq j\leq l}} M_{ij},$$
$$\bar{N}_{ij}=\max_{1\leq i,\, j\leq l} N_{ij}.$$

This shows that (14) holds with

$$M_2=L_2^2(m+l)[\max(\bar{L}_{ij}, \bar{L}_i, \bar{M}_{ij}, \bar{N}_{ij})].$$

Now put $\xi(x)$ and $U(\bar{x})$ as follows:

$$\xi(x)\equiv-u(x)A_3(x)=w,$$

and

$$U(\bar{x}) \equiv V(\bar{x}) \cap V_2(\bar{x}).$$

Then we have the following lemma.

Lemma 4. For arbitrary $x', x'' \in U(\bar{x})$ and arbitrary $w', w'' \in W(\bar{w})$, the following inequalities hold.

$$\|\xi(x'') - \xi(x')\| \leq (L_2M_1 + L_1M_2)\|x'' - x'\|, \qquad \cdots\cdots\cdots\cdots (17)$$

$$\|\phi(w'') - \phi(w')\| \leq K\|w'' - w'\|. \qquad \cdots\cdots\cdots\cdots (18)$$

Proof. Inequality (17) is shown as follows. For arbitrary $x', x'' \in U(\bar{x})$, it follows from (11), (12), (13) and (14) that

$$\begin{aligned}\|\xi(x'') - \xi(x')\| &= \|u(x'')A_3(x'') - u(x'')A_3(x') + u(x'')A_3(x') - u(x')A_3(x')\| \\ &\leq \|A_3(x'') - A_3(x')\|\|u(x'')\| + \|A_3(x')\|\|u(x'') - u(x')\| \\ &\leq (M_2L_1 + L_2M_1)\|x'' - x'\|.\end{aligned}$$

Moreover, observing that ϕ is differentiable, by the implicit function theorem, inequality (18) follows by setting

$$K = \sup_{(x,w)\in U(\bar{x})\times W(\bar{w})} \|-[E_{xx}(x,w)]^{-1}E_{xw}(x,w)\|.$$

4. Convergence Proof

The following theorem shows the local convergence of the iteration method proposed in Section 2.

Theorem 1. If $\bar{z}$ satisfies conditions (1)-(7), and the inequality

$$\tilde{K} \equiv K(L_1M_2 + L_2M_1) < 1$$

holds, then there exists a neighbourhood $U(\bar{x})$ such that for any starting point $x^{(0)} \in U(\bar{x})$, the sequence $x^{(k)}$ remains in $U(\bar{x})$ and converges to $\bar{x}$.

Proof. Note that

$$\bar{x} = \phi(\xi(\bar{x})).$$

For any $x^{(0)} \in U(\bar{x})$, (17) and (18) show

$$\begin{aligned}\|x^{(k+1)} - \bar{x}\| &= \|\phi(\xi(x^{(k)})) - \phi(\xi(\bar{x}))\| \\ &\leq K\|\xi(x^{(k)}) - \xi(\bar{x})\| \\ &\leq \tilde{K}\|x^{(k)} - \bar{x}\|.\end{aligned}$$

This completes the proof.

The following corollary follows immediately from Theorem 1.

Corollary 1. If the conditions in Theorem 1 are satisfied, then the sequence $\{w^{(k)}\}$

converges to $\bar{w}$.

Proof. The corollary immediately follows since

$$w^{(k+1)} = \xi(x^{(k)})$$

and

$$\bar{w} = \xi(\bar{x}).$$

Further we have the following corollary.

Corollary 2. Suppose that the same conditions as in Theorem 1 hold. Then $E(x^{(k+1)}, w^{(k+1)}) \leqq E(x^{(k)}, w^{(k)})$.

Proof. Since Step 3 implies that

$$E(x^{(k+1)}, w^{(k+1)}) \leqq E(x^{(k)}, w^{(k+1)}),$$

and Step 2 shows that

$$E(x^{(k)}, w^{(k+1)}) = \min_{w} E(x^{(k)}, w),$$

the corollary holds.

5. Modified Method

The iteration method for finding the optimal solution $\bar{x}$ of (P) is proposed in Section 2, and its local convergence is proved in Section 4. Step 3 in the proposed method requires the minimization of $E(x, w^{(k+1)})$. However, it seems that solving the associated unconstrained minimization problem requires much time because of the double iterations. In this section, we consider a modified method which determines $x^{(k+1)}$ without an iteration in Step 3.

The proposed modified method is as follows:

Step 1. Given $x^{(0)}$, set $k=0$ and choose a positive number $\varepsilon>0$.

Step 2: Solve the system of $(m+l)$ linear equations

$$A_2(x^{(k)})^* A_2(x^{(k)}) w^* = -A_2(x^{(k)})^* (f_x(x^{(k)}), 0, g(x^{(k)}))^*$$

and put $w^{(k+1)}$ as the solution.

Step 3: If $k=0$, then find $x^{(1)}$ that minimizes $E(x, w^{(1)})$. Otherwise, calculate $x^{(k+1)}$ by

$$x^{(k+1)} = x^{(k)} - \frac{\alpha}{\|A_1(x^{(k)}, w^{(k+1)})\|_E^2} y(x^{(k)}, w^{(k+1)}) A_1(x^{(k)}, w^{(k+1)}).$$

Step 4: Stop if $\max_j |x_j^{(k+1)} - x_j^{(k)}| < \varepsilon$. Otherwise set $k=k+1$ and return to Step 3.

Define an n-dimensional vector valued-function $\eta(x, w)$ as follows:

$$\eta(x, w) = x - \alpha \|A_1(x, w)\|_E^{-2} y(x, w) A_1(x, w).$$

Then the operation in Step 3 can be rewritten as

$$x^{(k+1)}=\eta(x^{(k)}, w^{(k+1)}).$$

As noted in Section 3, the operation in Step 2 is represented by

$$w^{(k+1)}=\xi(x^{(k)}).$$

For suitably chosen neighbourhoods $V_0(\bar{x})$ and $W_0(\bar{w})$, let

$$\hat{L}=\sup_{(x,w)\in V_0(\bar{x})\times W_0(\bar{w})}\|\partial\eta(x,w)/\partial x\|,$$

$$\hat{M}=\sup_{w\in W_0(\bar{w})}\|\partial\eta(\bar{x},w)/\partial w\|$$

and

$$\hat{N}=\sup_{x\in V_0(\bar{x})}\|\partial\xi(x)/\partial x\|.$$

The following theorem shows the local convergence of the modified method.

Theorem 2. If conditions (1)-(7) are satisfied, and the inequality

$$\hat{L}+\hat{M}\hat{N}<1$$

holds, then there exists a neighbourhood $V_0(\bar{x})$ such that for any initial point $x^{(0)} \in V_0(\bar{x})$, the sequence $x^{(k)}$ remains in $V_0(\bar{x})$ and converges to $\bar{x}$.

Proof.

$$\begin{aligned}\|x^{(k+1)}-\bar{x}\|&=\|\eta(x^{(k)},\xi(x^{(k)}))-\eta(\bar{x},\xi(x^{(k)}))\\&\quad+\eta(\bar{x},\xi(x^{(k)}))-\bar{x}\|\\&\leq\|\eta(x^{(k)},\xi(x^{(k)}))-\eta(\bar{x},\xi(x^{(k)}))\|\\&\quad+\|\eta(\bar{x},\xi(x^{(k)}))-\eta(\bar{x},\xi(\bar{x}))\|\\&\leq(\hat{L}+\hat{M}\hat{N})\|x^{(k)}-\bar{x}\|.\end{aligned}$$

This completes the proof.

6. Numerical Example

The Rosen-Suzuki Test Problem[7] was solved as a numerical example by using the proposed method and its modified version.

The Rosen-Suzuki Test Problem[7]:

Minimize

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

subject to

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

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

Table 1. Computation resutls of the proposed method.

$x^{(0)}=(0.6, 0.6, 0.6, 0.6)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec)	Number of iterations
0.9	0.005355	1.000254	2.002725	−0.985844	−43.99194	0.366	339
1.3	0.003690	1.000240	2.001867	−0.990165	−43.99415	0.308	271
1.9	0.002531	1.000216	2.001271	−0.993213	−43.99585	0.238	211

$x^{(0)}=(1.0, 1.0, 1.0, 1.0)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec)	Number of iterations
0.9	0.005343	1.000446	2.002683	−0.985861	−43.99200	0.381	342
1.3	0.003713	1.000337	2.001860	−0.990099	−43.99414	0.311	271
1.9	0.002533	1.000264	2.001263	−0.993204	−43.99585	0.250	220

$x^{(0)}=(1.1, 1.1, 1.1, 1.1)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec)	Number of iterations
0.9	0.005328	1.000748	2.002619	−0.985876	−43.99208	0.503	492
1.3	0.003699	1.000505	2.001821	−0.990120	−43.99418	0.403	371
1.9	0.002514	1.000355	2.001236	−0.993246	−43.99588	0.337	314

Table 2. Computation results of the modified method.

$x^{(0)}=(0.6, 0.6, 0.6, 0.6)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec)	Number of iterations
0.9	0.001715	1.000183	2.000855	−0.995379	−43.99711	0.517	501
1.3	0.001171	1.000149	2.000580	−0.996826	−43.99799	0.391	385
1.9	0.000800	1.000118	2.000392	−0.997832	−43.99861	0.298	290

$x^{(0)}=(1.0, 1.0, 1.0, 1.0)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec)	Number of iterations
0.9	0.001700	1.000212	2.000841	−0.995417	−43.99714	0.522	505
1.3	0.001170	1.000165	2.000575	−0.996836	−43.99800	0.384	349
1.9	0.000786	1.000124	2.000384	−0.997868	−43.99864	0.303	292

$x^{(0)}=(1.1, 1.1, 1.1, 1.1)$

α	x_1	x_2	x_3	x_4	f	CPU time (sec)	Number of iterations
0.9	0.001673	1.000257	2.000818	−0.995485	−43.99718	0.656	657
1.3	0.001155	1.000190	2.000563	−0.996874	−43.99802	0.483	488
1.9	0.000778	1.000137	2.000377	−0.997889	−43.99865	0.362	357

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

Optimal solution is $\bar{x} = (0, 1, 2, -1)$ with $f(\bar{x}) = -44$.

Computations with $\varepsilon = 10^{-4}$ were carried out on an M-190 computer of Kyoto University Computation Center. The results are shown in Tables 1 and 2.

7. Conclusions

In this paper, we proposed an iteration method and its modified version for solving Problem (P), and proved their local convergence. Compared with the previous method[5], the size of the system of equations solved for finding the optimal solution $\bar{x}$ is reduced from $(n+m+l)$ to n. Therefore, these methods seem favorable from the computational viewpoint.

References

1) A. V. Fiacco and G. P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization Technique, John Wiley, New York, 1968.
2) A. V. Fiacco and G. P. McCormick, "Computational Algorithm for the Sequential Unconstrained Minimization Technique for Nonlinear Programming," Management Sci., **10**, 601-617 (1964).
3) D. W. Marquardt, "An Algorithm for Least Squares Estimation of Nonlinear Parameters," SIAM J. Appl. Math. **11**, 431-441 (1963).
4) H. Mine, K. Ohno and T. Noda, "An Iteration Method for Nonlinear Programming Problems," J. Operations Res. Soc. Japan, **19**, 137-146 (1976).
5) H. Mine, K. Ohno and T. Noda, "An Iteration Method for Nonlinear Programming Problems: II," J. Operations Res. Soc. Japan, **20**, 132-138 (1977).
6) J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.
7) J. B. Rosen and S. Suzuki, "Construction of Nonlinear Programming Test Problems," Comm. ACM, **8**, 113 (1965).