Computers Math. Applic. Vol. 27, No. 1, pp. 25–31, 1994
Printed in Great Britain.
0898-1221/94 $6.00 + 0.00

A Hypoquadratic Convergence Method for Lagrange Multipliers

T. Altman
Department of Computer Science & Engineering
University of Colorado at Denver
Denver, CO 80204, U.S.A.

P. F. Boulos
Department of Computer Aided Engineering
Montgomery Watson
Pasadena, CA 91109, U.S.A.

(Received December 1991; accepted February 1993)

Abstract—In this paper, we investigate a class of hypoquadratically convergent methods for minimizing an objective function subject to equality constraints via the Lagrange multipliers method. The above class of inexact Newton methods has already been successfully applied for solving systems of nonlinear algebraic equations.

1. INTRODUCTION

In [1], we investigated a class of inexact Newton methods which were hypoquadratically convergent. These methods were designed for solving nonlinear systems of algebraic equations. In this paper, we apply the above methods to the problem of minimizing an objective function subject to equality constraints via the Lagrange multipliers method.

Consider the following nonlinear programming problem

$$\min f(x_1, x_2, \dots, x_n), \tag{1}$$

subject to the equality constraints

$$g_i(x_1, x_2, \dots, x_n) = 0, \qquad i = 1, 2, \dots, m; \quad m \le n, \tag{2}$$

where the independent variables and the values of the objective and the constraining functions are real.

The classical method of Lagrange multipliers yields the following necessary conditions: If $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^\top$ satisfies (2) and is a minimum or maximum point of the objective function f, then

$$g_i(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) = 0, \tag{3}$$

and

$$\frac{\partial}{\partial x_k} f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) + \sum_{i=1}^{m} \lambda_i \frac{\partial}{\partial x_k} g_i(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) = 0, \tag{4}$$

for $k = 1, 2, \dots, n$, and some $\lambda_1, \lambda_2, \dots, \lambda_m$ which are called the Lagrange multipliers.

A point $(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^\top$ satisfying Conditions (3) and (4) is called a stationary point. It is well-known that in order to determine if the stationary point is a maximum, minimum, or a saddle

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX point of the objective function, one has to examine the behavior of the second derivatives of the objective function in its immediate neighborhood. The aim of this paper is to present a method for finding such stationary points, which can then be subject to such further investigations [2,3].

2. DESCRIPTION OF THE METHOD

Let us introduce the following notation. Put

$$\Phi = f + \sum_{i=1}^{m} \lambda_i g_i,$$

then

$$\frac{\partial \Phi}{\partial x_k} = \frac{\partial f}{\partial x_k} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_k}.$$

Now put

$$F = \left(g_1, g_2, \dots, g_m, \frac{\partial \Phi}{\partial x_1}, \frac{\partial \Phi}{\partial x_2}, \dots, \frac{\partial \Phi}{\partial x_n}\right)^{\mathsf{T}}$$

or

$$F(\boldsymbol{x}) = (F_i(\boldsymbol{x}^{\mathsf{T}})), \qquad \text{for} \quad i = 1, 2, \dots, m, m+1, m+2, \dots, m+n, \tag{5}$$

where

$$F_i = g_i, \qquad \text{for} \quad i = 1, 2, \dots, m \quad \text{and} \quad F_{m+j} = \frac{\partial \Phi}{\partial x_j}, \qquad j = 1, 2, \dots, n.$$

Now put

$$\boldsymbol{x} = (x_1, x_2, \dots, x_{m+n})^{\mathsf{T}}, \tag{6}$$

where

$$x_{n+i} = \lambda_i, \qquad \text{for } i = 1, 2, \dots, m.$$

Equations (3) and (4) can now be rewritten as

$$F(\bar{\boldsymbol{x}}) = 0. \tag{7}$$

Denote by $F'(\boldsymbol{x})$ the Jacobian of F at $\boldsymbol{x}$:

$$F'(\boldsymbol{x}) = \frac{d(F_1(\boldsymbol{x}), \dots, F_{n+m}(\boldsymbol{x}))}{d(x_1, x_2, \dots, x_{n+m})} = \left(\frac{\partial F_j}{\partial x_k}\right), \qquad \begin{matrix} j = 1, \dots, n+m, \\ k = 1, \dots, n+m, \end{matrix} \tag{8}$$

where

$$\begin{aligned}
\frac{\partial F_i}{\partial x_j} &= \frac{\partial g_i}{\partial x_j}, && \text{for } i = 1, 2, \dots, m, \text{ and } j = 1, 2, \dots, n, \\
\frac{\partial F_i}{\partial x_j} &= \frac{\partial g_i}{\partial x_j} = 0, && \text{for } i = 1, 2, \dots, m, \text{ and } j = n+1, n+2, \dots, n+m, \\
\frac{\partial F_i}{\partial x_j} &= \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{k=1}^{m} \lambda_k \frac{\partial^2 g_k}{\partial x_i \partial x_j}, && \text{for } i = m+1, m+2, \dots, m+n; \ j = 1, 2, \dots, n, \\
\frac{\partial F_i}{\partial x_j} &= \frac{\partial g_k}{\partial x_i}, && \text{for } i = m+1, m+2, \dots, m+n; \ j = n+k, k = 1, 2, \dots, n.
\end{aligned}$$

Thus, the Jacobian in (8) can be represented by the following $(m+n) \times (m+n)$ matrix

$$F'(x) = \begin{pmatrix} \frac{\partial g_1(x)}{\partial x_1} & \frac{\partial g_1(x)}{\partial x_2} & \cdots & \frac{\partial g_1(x)}{\partial x_n} & 0 & 0 & \cdots & 0 \\ \frac{\partial g_2(x)}{\partial x_1} & \frac{\partial g_2(x)}{\partial x_2} & \cdots & \frac{\partial g_2(x)}{\partial x_n} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_m(x)}{\partial x_1} & \frac{\partial g_m(x)}{\partial x_2} & \cdots & \frac{\partial g_m(x)}{\partial x_n} & 0 & 0 & \cdots & 0 \\ \frac{\partial^2 \Phi(x)}{\partial x_1 \partial x_1} & \frac{\partial^2 \Phi(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 \Phi(x)}{\partial x_1 \partial x_n} & \frac{\partial g_1(x)}{\partial x_1} & \frac{\partial g_2(x)}{\partial x_1} & \cdots & \frac{\partial g_m(x)}{\partial x_1} \\ \frac{\partial^2 \Phi(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 \Phi(x)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 \Phi(x)}{\partial x_2 \partial x_n} & \frac{\partial g_1(x)}{\partial x_2} & \frac{\partial g_2(x)}{\partial x_2} & \cdots & \frac{\partial g_m(x)}{\partial x_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 \Phi(x)}{\partial x_n \partial x_1} & \frac{\partial^2 \Phi(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 \Phi(x)}{\partial x_n \partial x_n} & \frac{\partial g_1(x)}{\partial x_n} & \frac{\partial g_2(x)}{\partial x_n} & \cdots & \frac{\partial g_m(x)}{\partial x_n} \end{pmatrix}$$

Our problem now is to solve approximately (3) and (4) or, equivalently, equation (7). To this aim we shall use the hypoquadratic convergence method first presented in [1].

The following is the iterative method under consideration:

$$\boldsymbol{x}_{i+1} := \boldsymbol{x}_i + h_i, \tag{9}$$

where h_i is a solution of the equation

$$F'(\boldsymbol{x}_i)h_i + F(\boldsymbol{x}_i) = r_i, \tag{10}$$

$\boldsymbol{x}_0$ being the initial approximate solution. If $r_i = 0$ for $i = 0, 1, \dots$, then the iterative method (9) is the *exact* Newton method.

Denote by $B(\boldsymbol{x}_0, R)$ the ball with center $\boldsymbol{x}_0$ and radius R for some given $R > 0$:

$$B(\boldsymbol{x}_0, R) = \{\boldsymbol{x} \mid \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq R\}.$$

We assume that F is defined on a domain containing the ball $B(\boldsymbol{x}_0, R)$, $\boldsymbol{x}_0 \in \mathbf{R}^{n+m}$.

DEFINITION 2.1. [1] *Consider the numerical series*

$$\sum_{i=0}^{\infty} a_i, a_i > 0. \tag{11}$$

Assume that the series (11) is convergent and set

$$R_i = \sum_{k=1}^{\infty} a_k.$$

Then the series (11) is hypoquadratically convergent if

$$R_{i+1} \leq a^{t^i} R_i,$$

for some $0 < a < 1$ and $1 < t < 2$.

Suppose that the iterative method (9) converges to a solution $\boldsymbol{x}$ of the equation (7). Then, we have

$$\boldsymbol{x}_i = \boldsymbol{x}_0 + \sum_{k=0}^{i-1} h_k \quad \text{and } \boldsymbol{x} = \boldsymbol{x}_0 + \sum_{i=0}^{\infty} h_i. \tag{12}$$

Now, assume that the series in equation (12) is dominated by the series in Condition (11), that is,

$$\|h_i\| \leq D\, a_i, \qquad i = 0, 1, \dots,$$

for some constant $D > 0$. Then the series

$$\sum_{i=0}^{\infty} a_i$$

is called the iterative majorant for method (9), and we have

$$\|\boldsymbol{x} - \boldsymbol{x}_i\| \leq R_i. \tag{13}$$

We may assume that relation (13) holds for almost all i.

DEFINITION 2.2. [1] *The iterative method (9) is convergent hypoquadratically, if there exists an iterative majorant series which is convergent hypoquadratically.*

We shall now show how to make the inexact Newton method (9) hypoquadratically convergent. Consider the sequence

$$a_i = a_0^{t^i}, \tag{14}$$

where $1 < t < 2$ and $a_0 = \|F(\boldsymbol{x}_0)\|$ is sufficiently small, and set

$$R = \sum_{i=0}^{\infty} a_0^{t^i}. \tag{15}$$

Assume that the following conditions are satisfied:

(A_1) There exists a constant $C > 0$ such that

$$\|F(\boldsymbol{x} + h) - F(\boldsymbol{x}) - F'(\boldsymbol{x})h\| \leq C\|h\|^2 \tag{16}$$

for all $\boldsymbol{x}$ and $\boldsymbol{x} + h$ belonging to $B(\boldsymbol{x}_0, R)$.

(A_2) For every $\boldsymbol{x} \in B(\boldsymbol{x}_0, R)$ and $y \in \mathbf{R}^{n+m}$, if h is a solution of the equation

$$F'(\boldsymbol{x})h + y = 0, \tag{17}$$

then

$$\|h\| \leq K\|y\| \tag{18}$$

for some constant $K > 0$.

Moreover, assume that equation (7) can be solved with arbitrary accuracy, that is, $\|r_i\|$ in (10) can be made arbitrarily small. Thus, we assume that

$$\|r_i\| \leq a_i^p, \tag{19}$$

for $i = 0, 1, \dots,$ and $p > 2$ is fixed.

THEOREM 2.1. *In addition to the assumptions $(A_1), (A_2)$, and Condition (19), suppose that $a_0 = \|F(\boldsymbol{x}_0)\| < 1$ is so small that*

$$C_1 a_0^2 \leq a_0^t \quad \text{or } a_0 < C_1^{-1/(2-t)}, \tag{20}$$

where $1 < t < 2$ and $C_1 = 4CK^2 + 1$. Then the sequence of approximate solutions $\{\boldsymbol{x}_i\}$ determined by equation (9) converges to a solution $\boldsymbol{x}$ of the equation (7) and $\boldsymbol{x}_i, \boldsymbol{x} \in B(\boldsymbol{x}_0, R)$ with R given by equation (15). The convergence is hypoquadratic.

PROOF. The proof results from the proof of Theorem 2.1 in [1] . ∎

REMARK 2.1. Condition (16) is satisfied if $F'(\boldsymbol{x})$ is Lipschitz continuous. If $F'(\boldsymbol{x})$ is Hölder continuous, then Condition (16) is replaced by the following.

$$\|F(\boldsymbol{x} + h) - F(\boldsymbol{x}) - F'(\boldsymbol{x})h\| \leq C\|h\|^{1+\alpha},$$

for all $\boldsymbol{x}$ and $\boldsymbol{x} + h$ belonging to $B(\boldsymbol{x}_0, R)$ and some $0 < \alpha < 1$. Then the condition imposed on the parameter t is

$$1 < t < 1 + \alpha,$$

and we also assume that p in Condition (19) satisfies the following condition

$$p > 1 + \alpha.$$ ∎

REMARK 2.2. Condition (16) is satisfied if the second derivatives of f and g_i are Lipschitz continuous. ∎

3. SAMPLE APPLICATION OF THE METHOD

The plane

$$g_1 = lx + my + nz = 0 \tag{21}$$

is passing through the center of the ellipsoid

$$g_2 = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0, \qquad (a > b > c). \tag{22}$$

The problem of finding the halfaxes obtained through the intersection of the plane and the ellipsoid is equivalent to finding the extremum values of the function

$$f = r^2 = x^2 + y^2 + z^2, \tag{23}$$

provided its independent variables are subject to the constraints (21) and (22). Putting

$$\begin{aligned}\Phi(x,y,z) &= f(x,y,z) + \lambda_1 g_1(x,y,z) + \lambda_2 g_2(x,y,z)\\ &= x^2 + y^2 + z^2 + 2\mu(lx + my + nz) + \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right),\end{aligned}$$

where $\lambda_1 = 2\mu$ and $\lambda_2 = \lambda$, we get by equation (4) the partial derivatives of Φ

$$x + \mu l + \lambda\frac{x}{a^2} = 0, \quad y + \mu m + \lambda\frac{y}{b^2} = 0, \quad \text{and} \quad z + \mu n + \lambda\frac{z}{c^2} = 0. \tag{24}$$

By adding the equations (24) multiplied by x, y, z, respectively, we get

$$\lambda = -r^2.$$

Hence, we obtain from (24)

$$x = -\mu\frac{la^2}{a^2 - r^2}, \quad y = -\mu\frac{mb^2}{b^2 - r^2}, \quad \text{and} \quad z = -\mu\frac{nc^2}{c^2 - r^2}. \tag{25}$$

By adding the equations (25) multiplied by l, m, n, respectively, we obtain

$$\frac{l^2a^2}{a^2 - r^2} + \frac{m^2b^2}{b^2 - r^2} + \frac{n^2c^2}{c^2 - r^2} = 0. \tag{26}$$

Thus, the extremum values for r^2 can be obtained from equation (26) which yields

$$\begin{aligned}(l^2a^2 + m^2b^2 + n^2c^2)r^4 - (l^2a^2(b^2 + c^2) + m^2b^2(a^2 + c^2)\\ + n^2c^2(a^2 + b^2))r^2 + a^2b^2c^2(l^2 + m^2 + n^2) = 0.\end{aligned} \tag{27}$$

On the other hand, equation (25) implies

$$x^2 + y^2 + z^2 = r^2 = \mu^2\left(\frac{l^2a^4}{(a^2 - r^2)^2} + \frac{m^2b^4}{(b^2 - r^2)^2} + \frac{n^2c^4}{(c^2 - r^2)^2}\right).$$

Hence, we get

$$\mu = -r\left(\frac{l^2a^4}{(a^2 - r^2)^2} + \frac{m^2b^4}{(b^2 - r^2)^2} + \frac{n^2c^4}{(c^2 - r^2)^2}\right)^{-1/2}. \tag{28}$$

Let us now deal with specific numbers by putting

$$l = m = n = 1; \qquad a = 3,\ b = 2,\ c = 1. \tag{29}$$

Thus, equation (27) becomes

$$14r^4 - 98r^2 + 108 = 0.$$

Hence, we get for r^2 the approximate value

$$r^2 = 5.63, \tag{30}$$

and by virtue of equation (28), we obtain

$$\mu = -0.64. \tag{31}$$

Consequently, we determine from equations (25) and (31) the following values

$$x = 1.71, \quad y = -1.57, \quad z = -0.14. \tag{32}$$

Having found the numerical solution to the problem (21)–(23) with specific numerical values from (29), we are now in a position to illustrate the iterative method (9) by using the same problem as presented above, i.e., find an extremum value of the function

$$f(x,y,z) = x^2 + y^2 + z^2 \tag{33}$$

subject to constraints

$$g_1(x,y,z) = x + y + z = 0, \tag{34}$$

$$g_2(x,y,z) = \frac{x^2}{9} + \frac{y^2}{4} + z^2 - 1 = 0. \tag{35}$$

Now, using the Lagrange multipliers $\lambda_1 = \mu$, instead of $\lambda_1 = 2\mu$ as in equation (24), and $\lambda_2 = \lambda$, the problem reduces to solving the system (7) or (34), (35) and

$$\begin{aligned} \frac{\partial\Phi}{\partial x} &= 2x + \mu + \lambda\frac{2x}{9} = 0,\\ \frac{\partial\Phi}{\partial y} &= 2y + \mu + \lambda\frac{2y}{4} = 0,\\ \frac{\partial\Phi}{\partial z} &= 2z + \mu + \lambda 2z = 0. \end{aligned} \tag{36}$$

Thus, the system $F(x,y,z,\mu,\lambda) = 0$ (see equation (7)) consists of the equations (34)–(36), and its Jacobian is the following

$$F'(x,y,z,\mu,\lambda) = \begin{bmatrix} 1 & 1 & 1 & 0 & 0\\ \frac{2}{9}x & \frac{2}{4}y & 2z & 0 & 0\\ 2\left(1+\frac{\lambda}{9}\right) & 0 & 0 & 1 & 2\frac{x}{9}\\ 0 & 2\left(1+\frac{\lambda}{4}\right) & 0 & 1 & 2\frac{y}{4}\\ 0 & 0 & 2(1+\lambda) & 1 & 2z \end{bmatrix}.$$

Since the determinant of F' at the stationary point is not zero, the Jacobian is nonsingular in some neighborhood of that point. It follows that the equation (10) has a solution if $\boldsymbol{x}_0$ belongs to that neighborhood. The other assumptions of Theorem 2.1 can also be satisfied, since all functions involved are continuous, hence bounded in any closed neighborhood of the initial point $\boldsymbol{x}_0$.

Consider the same problem (1),(2) where we look for a stationary point $\bar{\boldsymbol{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ with $\bar{x}_1 \geq 0,\ \bar{x}_2 \geq 0, \dots, \bar{x}_n \geq 0$. We may still use the iterative method (9) with

$$\boldsymbol{x}_i = (x_{i_1}, x_{i_2}, \dots, x_{i_n}, x_{i_{n+1}}, x_{i_{n+2}}, \dots, x_{i_{n+m}}), \tag{37}$$

where the first n coordinates are supposed to be nonnegative. In the case when some of them are negative, we replace them with 0 and denote the new iterate with y_i. The next iteration step is then defined as in equation (9), that is,

$$\boldsymbol{x}_{i+1} = y_i + h_i, \tag{38}$$

where

$$F'(y_i)h_i + F(y_i) = r_i. \tag{39}$$

Now suppose that the iteration sequence (9) contains an infinite subsequence of the form given by equation (38). Denote by $\boldsymbol{x}^*$ the limit of the iteration sequence $\{\boldsymbol{x}_i\}$ obtained by implementing the changes made in equation (38). If

$$\boldsymbol{x}^* = ({x_1}^*, {x_2}^*, \dots, {x_n}^*, {x_{n+1}}^*, {x_{n+2}}^*, \dots, {x_{n+m}}^*),$$

then ${x_1}^* \geq 0$, ${x_2}^* \geq 0, \dots, {x_n}^* \geq 0$. In this fashion, we obtain a critical point whose first n coordinates are nonnegative. Moreover, if for some $1 \leq j \leq n$, we would need to use the formula (38) infinitely many times, then in the limit ${x_j}^* = 0$.

REMARK 3.1. In using the above modification, one has to make sure that $\|F(y_i)\| \leq 1$. But this is feasible for as $i \to \infty$ the sequence $\{F(\boldsymbol{x}_i)\}$ converges to 0. ∎

Let us modify the example from the beginning of this section by replacing equation (34) by

$$g_1(x, y, z) = x - y - z = 0, \tag{40}$$

and with an additional condition

$$x \geq 0, \quad y \geq 0, \quad z \geq 0. \tag{41}$$

Let us choose our initial guess $\boldsymbol{x}_0$ so that the conditions of Theorem 2.1 are satisfied. If

$$\boldsymbol{x}_0 = (x, y, z, \mu, \lambda) = (1.5, 1.3, 0.1, -1.25, -5), \tag{42}$$

then substituting these values into the equation (40) and replacing equation (36) by

$$\begin{aligned} \frac{\partial \Phi}{\partial x} &= 2x + \mu + \lambda \frac{2x}{9} = 0, \\ \frac{\partial \Phi}{\partial y} &= 2y - \mu + \lambda \frac{2y}{4} = 0, \\ \frac{\partial \Phi}{\partial z} &= 2z - \mu + \lambda 2z = 0, \end{aligned} \tag{43}$$

we get

$$\|F(\boldsymbol{x}_0)\|^2 = (0.1)^2 + (-0.32)^2 + (0.6)^2 + (0.45)^2 = 0.67 < 1.$$

Thus, our initial guess (42) satisfies the condition

$$a_0 = \|F(\boldsymbol{x}_0)\| < 1,$$

of Theorem 2.1.

REFERENCES

1. T. Altman, A hypoquadratic convergence method for systems of nonlinear equations, *Computers Math. Applic.* **20** (1), 57–60 (1990).
2. M. Avriel, *Nonlinear Programming: Analysis and Methods*, Prentice Hall, Englewood Cliffs, N.J., (1976).
3. P.F. Boulus and T. Altman, A graph-theoretic approach to explicit nonlinear pipe network optimization, *Journal of Applied Mathematical Modelling* **15**, 459–466 (1991).