第25卷第1期 数 学 研 究 与 评 论 Vol.25, No.1
2005年2月 JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION Feb., 2005

Article ID: 1000-341X(2005)01-0064-08 **Document code:** A

A New Huang Class and Its Properties for Unconstrained Optimization Problems

WEI Zeng-xin[1], *LI Qiao-xing*[2]

(1. School of Math. Info. Sci., Guangxi University, Nanning 530004, China;
2. College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Jiangsu 210016, China)
(E-mail: zxwei@gxu.edu.cn)

Abstract: This paper presents a new class of quasi-Newton methods for solving unconstrained minimization problems. The methods can be regarded as a generalization of Huang class of quasi-Newton methods. We prove that the directions and the iterations generated by the methods of the new class depend only on the parameter ρ if the exact line searches are made in each steps.

Key words: unconstrained optimization; quasi-Newton equation; quasi-Newton method.
MSC(2000): 90C30
CLC number: O221.2

1. Introduction

The restricted Huang class of quasi-Newton methods are very famous for solving the following unconstrained optimization problem

$$\min\{f(x) \mid x \in R^n\},$$

where $f : \Re^n \to R$ is a continuously differentiable function whose gradient is denoted by g. The updated matrix H_k is generated based on the following quasi-Newton equation

$$H_{k+1}y_k = s_k,$$

where $y_k = g_{k+1} - g_k$ and $s_k = x_{k+1} - x_k$.

Recently, [1] (or [2]) gave a formula

$$A_k(1) = c\|g_k\| + \max\{0, -\frac{y_k^T s_k}{\|s_k\|^2}\}, \tag{1.2}$$

where c is a positive number, then we have a modified BFGS method as follows:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k^a (y_k^a)^T}{s_k^T y_k^a},$$

where $y_k^a = y_k + A_k(1)s_k$.

Received date: 2003-01-13
Foundation item: NNSF of China (10161002) and NSF of Guangxi (0135004)

Just like $A_k(1)$, we can construct some formulas such as

$$A_k(2) = \mu_k a_k, \tag{1.3}$$

where $a_k = \frac{\|g_{k+1}^T(g_{k+1}-g_k)\|}{\|s_k\|}$, $\mu_k \geq 0$, and

$$A_k(3) = \frac{2(f_k - f_{k+1}) + (g_{k+1} + g_k)^T s_k}{\|s_k\|^2}, \tag{1.4}$$

where $f_k = f(x_k)$.

We denote $A_k(1)$, $A_k(2)$, $A_k(3)$ and so on, as A_k, and modify the famous methods such as BFGS, DFP and Broyden as follows

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k^*(y_k^*)^T}{s_k^T y_k^*},$$

$$B_{k+1} = B_k - (1 + \frac{s_k^T B_k s_k}{s_k^T y_k^*})\frac{y_k(y_k^*)^T}{s_k^T y_k^*} - \frac{y_k^* s_k^T B_k + B_k s_k (y_k^*)^T}{s_k^T y_k^*},$$

and

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k^*(y_k^*)^T}{s_k^T y_k^*} + \varphi(s_k^T B_k s_k)^{\frac{1}{2}} \omega_k \omega_k^T,$$

where $\omega_k = \frac{y_k^*}{s_k^T y_k^*} - \frac{B_k s_k}{s_k^T B_k s_k}$ and φ is a scalar parameter, $y_k^* = y_k + A_k s_k$.

Using the Huang formula

$$H_{k+1} = H_k + s_k u_k^T + H_k y_k v_k^T,$$

where

$$u_k = a_{11} s_k + a_{12} H_k^T y_k,$$

$$v_k = a_{21} s_k + a_{22} H_k^T y_k,$$

u_k and v_k satisfy

$$\begin{cases} u_k^T y_k = \rho, \\ v_k^T y_k = -1, \end{cases}$$

and H_{k+1} satisfies $H_{k+1} y_k = \rho s_k$, we propose the following new Huang formula:

$$H_{k+1} = H_k + s_k (u_k^*)^T + H_k y_k^* (v_k^*)^T, \tag{1.5}$$

where

$$y_k^* = y_k + A_k(i) s_k, i = 1, 2, 3, \tag{1.6}$$

$$u_k^* = a_{11} s_k + a_{12} H_k^T y_k^*, \tag{1.7}$$

$$v_k^* = a_{21} s_k + a_{22} H_k^T y_k^*. \tag{1.8}$$

u_k^* and v_k^* satisfy

$$\begin{cases} (u_k^*)^T y_k^* = \rho, \\ (v_k^*)^T y_k^* = -1. \end{cases} \tag{1.9}$$

We can easily deduce that H_{k+1} satisfies the equation

$$H_{k+1}y_k^* = \rho s_k. \tag{1.10}$$

We shall call the methods which satisfy (1.5)–(1.9) modified Huang class (M-Huang class).

The remainders of this paper are organized as follows. In Section 2, we present the M-Huang methods with exact line search. The verification of the properties of the M-Huang class is given in Section 3. We have a discussion in Section 4.

2. Statement of the algorithm

In this section, we give an algorithm which is based on the M-Huang class formula and exact line search.

Algorithm 2.1

Step 0: Choose an initial point $x_0 \in R^n$, an initial matrix $H_0 \in R^{n*n}$ and a parameter ρ. Set $k := 0$.

Step 1: If $g_k = 0$, stop.

Step 2: Solve the problem

$$d_k = -H_k^T g_k \tag{2.1}$$

to get a search direction d_k.

Step 3: Find α_k^* with exact line search:

$$f(x_k + \alpha_k^* d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k).$$

Step 4: Set $x_{k+1} = x_k + \alpha_k^* d_k$. Update H_{k+1} by the following formula:

$$H_{k+1} = H_k + s_k(u_k^*)^T + H_k y_k^*(v_k^*)^T$$

where y_k^*, u_k^* and v_k^* satisfy the equations (1.6)–(1.9).

Step 5: Set $k := k+1$ and go to step 1.

In Section 3, we will study the properties of the above algorithm.

3. Some properties of M-Huang class

In this section, we will show that Algorithm 2.1 generates the same directions and the iterates which depend only on the parameter ρ for each A_k. In doing so, we first give the following lemma.

Lemma 3.1 *Suppose that $\{(x_k, H_k, d_k)\}$ is gerenated by Algorithm 2.1 and for all k,*

$$g_k^T s_k \neq 0, \tag{3.1}$$

then (2.1) can be denoted as

$$d_{k+1} = -\{1 + a_{22}(y_k^*)^T H_k^T g_{k+1}\}[I - \frac{s_k (y_k^*)^T}{s_k^T y_k^*}]H_k^T g_{k+1} - \\ a_{22} A_k (y_k^*)^T H_k^T g_{k+1} [I - \frac{s_k (y_k^*)^T}{s_k^T y_k^*}] H_k^T s_k. \tag{3.2}$$

Proof By using the exact line search, we obtain for all k that

$$g_{k+1}^T s_k = 0. \tag{3.3}$$

Using the equalities (2.1) and (3.3), we have

$$\begin{aligned} d_{k+1} &= -H_{k+1}^T g_{k+1} \\ &= -[H_k + s_k (u_k^*)^T + H_k y_k^* (v_k^*)^T]^T g_{k+1} \\ &= -H_k^T g_{k+1} - u_k^* s_k^T g_{k+1} - v_k^* (y_k^*)^T H_k^T g_{k+1} \\ &= -H_k^T g_{k+1} - [a_{21} s_k + a_{22} H_k^T (y_k^*)](y_k^*)^T H_k^T g_{k+1} \\ &= -H_k^T g_{k+1} - [a_{21} s_k + a_{22} H_k^T g_{k+1} - a_{22} H_k^T g_k + a_{22} A_k H_k^T s_k](y_k^*)^T H_k^T g_{k+1} \\ &= -\{1 + a_{22}(y_k^*)^T H_k^T g_{k+1}\} H_k^T g_{k+1} - a_{21}(y_k^*)^T H_k^T g_{k+1} s_k + \\ &\quad a_{22}(y_k^*)^T H_k^T g_{k+1} H_k^T g_k - a_{22} A_k (y_k^*)^T H_k^T g_{k+1} H_k^T s_k. \end{aligned}$$

From $\frac{1}{\alpha_k^*} s_k = d_k = -H_k^T g_k$, we get

$$d_{k+1} = -\{1 + a_{22}(y_k^*)^T H_k^T g_{k+1}\} H_k^T g_{k+1} - [a_{21} + \frac{a_{22}}{\alpha_k^*}](y_k^*)^T H_k^T g_{k+1} s_k - \\ a_{22} A_k (y_k^*)^T H_k^T g_{k+1} H_k^T s_k. \tag{3.4}$$

On the other hand,

$$\begin{aligned} -(a_{21} + \frac{a_{22}}{\alpha_k^*}) s_k^T y_k^* &= -[a_{21} s_k^T y_k^* - \frac{a_{22}}{\alpha_k^*}(y_k^*)^T \alpha_k^* H_k^T g_k] \\ &= -[a_{21} s_k^T y_k^* - a_{22}(y_k^*)^T H_k^T g_k] \\ &= -[a_{21} s_k^T y_k^* + a_{22}(y_k^*)^T H_k^T y_k^*] + a_{22}(y_k^*)^T H_k^T g_{k+1} + a_{22} A_k (y_k^*)^T H_k^T s_k \\ &= -[a_{21} s_k + a_{22} H_k^T y_k^*]^T y_k^* + a_{22}(y_k^*)^T H_k^T g_{k+1} + a_{22} A_k (y_k^*)^T H_k^T s_k, \end{aligned}$$

which combining with (1.7) and (1.9) yields

$$-(a_{21} + \frac{a_{22}}{\alpha_k^*}) s_k^T y_k^* = 1 + a_{22}(y_k^*)^T H_k^T g_{k+1} + a_{22} A_k (y_k^*)^T H_k^T s_k. \tag{3.5}$$

By using (3.5) and (3.4), we obtain

$$\begin{aligned}
d_{k+1} =& -\{1+a_{22}(y_k^*)^T H_k^T g_{k+1}\}H_k^T g_{k+1}+\\
& \frac{1+a_{22}(y_k^*)^T H_k^T g_{k+1}+a_{22}A_k(y_k^*)^T H_k^T s_k}{s_k^T y_k^*}(y_k^*)^T H_k^T g_{k+1}s_k-\\
& a_{22}A_k(y_k^*)^T H_k^T g_{k+1}H_k^T s_k\\
=& -\{1+a_{22}(y_k^*)^T H_k^T g_{k+1}\}H_k^T g_{k+1}+\frac{1+a_{22}(y_k^*)^T H_k^T g_{k+1}}{s_k^T y_k^*}(y_k^*)^T H_k^T g_{k+1}s_k+\\
& \frac{a_{22}A_k(y_k^*)^T H_k^T s_k}{s_k^T y_k^*}(y_k^*)^T H_k^T g_{k+1}s_k - a_{22}A_k(y_k^*)^T H_k^T g_{k+1}H_k^T s_k\\
=& -\{1+a_{22}(y_k^*)^T H_k^T g_{k+1}\}[I-\frac{s_k(y_k^*)^T}{s_k^T y_k^*}]H_k^T g_{k+1}-\\
& a_{22}A_k(y_k^*)^T H_k^T g_{k+1}[I-\frac{s_k(y_k^*)^T}{s_k^T y_k^*}]H_k^T s_k.
\end{aligned}$$

From the above formula of d_{k+1}, we can see that the search direction d_{k+1} is only dependent on A_k but is not dependent on ρ. For example, If we use

$$y_k^* = y_k + A_k(2)s_k,$$

then we have

$$\begin{aligned}
d_{k+1} =& -\{1+a_{22}(y_k^*)^T H_k^T g_{k+1}\}[I-\frac{s_k(y_k^*)^T}{s_k^T y_k^*}]H_k^T g_{k+1}-\\
& a_{22}\mu_k a_k(y_k^*)^T H_k^T g_{k+1}[I-\frac{s_k(y_k^*)^T}{s_k^T y_k^*}]H_k^T s_k.
\end{aligned}$$

The following Theorem 3.1 is our main result in this paper.

Theorem 3.1 *Suppose that $\{x_k\}$ is generated by Algorithm 2.1. If for all k (3.1) holds, then $\{x_k\}$ is dependent only on the parameter ρ.*

Proof It is easy to verify that, for a given M-Huang class, if x_0 and H_0 are the same for all members in this class, then x_1, s_0 and y_0 are also the same. From Lemma 3.1, we can deduce that d_1, x_2, s_1 and y_1 are the same. Supposing that we have proved that $x_0, x_1, \ldots, x_{k+1}$ are the same, we will prove that x_{k+2} is also the same.

From Lemma 3.1, we can denote that

$$d_{k+1} = -\{1+a_{22}(y_k^*)^T H_k^T g_{k+1}\}R_k H_k^T g_{k+1} - a_{22}A_k(y_k^*)^T H_k^T g_{k+1}R_k H_k^T s_k, \tag{3.6}$$

where $R_k = I - \frac{s_k(y_k^*)^T}{s_k^T y_k^*}$.

Since for all k, $R_k s_k = 0$, so we get

$$\begin{aligned}
R_{k+1}s_{k+1} =& \alpha_{k+1}^* R_{k+1}d_{k+1}\\
=& -\alpha_{k+1}^*\{1+a_{22}(y_k^*)^T H_k^T g_{k+1}\}R_{k+1}R_k H_k^T g_{k+1}-\\
& \alpha_{k+1}^* a_{22}A_k(y_k^*)^T H_k^T g_{k+1}R_{k+1}R_k H_k^T s_k\\
=& 0.
\end{aligned}$$

Thus,

$$R_{k+1}R_kH_k^Tg_{k+1} = -\frac{a_{22}A_k(y_k^*)^TH_k^Tg_{k+1}}{1+a_{22}(y_k^*)^TH_k^Tg_{k+1}}R_{k+1}R_kH_k^Ts_k. \tag{3.7}$$

From $R_kH_k^Tg_k = -\frac{1}{\alpha_k^*}R_ks_k = 0$, we have

$$R_ku_k^* = R_k(a_{11}s_k + a_{12}H_k^Ty_k^*) = a_{11}R_ks_k + a_{12}R_kH_k^T(y_k + A_ks_k).$$

So

$$R_ku_k^* = a_{12}R_kH_k^Tg_{k+1} + a_{12}A_kR_kH_k^Ts_k.$$

Similarly, we also obtain

$$R_kv_k^* = R_k(a_{21}s_k + a_{22}H_k^Ty_k^*) = a_{22}R_kH_k^Tg_{k+1} + a_{22}A_kR_kH_k^Ts_k. \tag{3.9}$$

Again, using (1.7), we have

$$R_ku_k^* = [I - \frac{s_k(y_k^*)^T}{s_k^Ty_k^*}]u_k^* = u_k^* - \rho\frac{s_k}{s_k^Ty_k^*} \tag{3.10}$$

and

$$R_kv_k^* = [I - \frac{s_k(y_k^*)^T}{s_k^Ty_k^*}]v_k^* = v_k^* + \frac{s_k}{s_k^Ty_k^*}. \tag{3.11}$$

Combining (3.8) with (3.10) and (3.7), we have

$$\begin{aligned} R_{k+1}u_k^* =& R_{k+1}(R_ku_k^* + \rho\frac{s_k}{s_k^Ty_k^*}) \\ =& a_{12}R_{k+1}R_kH_k^Tg_{k+1} + a_{12}A_kR_{k+1}R_kH_k^Ts_k + \rho\frac{R_{k+1}s_k}{s_k^Ty_k^*}. \end{aligned}$$

So

$$R_{k+1}u_k^* = a_{12}\frac{A_k}{1+a_{22}(y_k^*)^TH_k^Tg_{k+1}}R_{k+1}R_kH_k^Ts_k + \rho\frac{R_{k+1}s_k}{s_k^Ty_k^*}. \tag{3.12}$$

Combining (3.9) with (3.11) and (3.7), we also have

$$R_{k+1}v_k^* = R_{k+1}(R_kv_k^* - \frac{s_k}{s_k^Ty_k^*}) = a_{22}\frac{A_k}{1+a_{22}(y_k^*)^TH_k^Tg_{k+1}}R_{k+1}R_kH_k^Ts_k - \frac{R_{k+1}s_k}{s_k^Ty_k^*}. \tag{3.13}$$

From (1.5), (3.12) and (3.13), we have

$$\begin{aligned}R_{k+1}H_{k+1}^T =& R_{k+1}(H_k + s_k(u_k^*)^T + H_k y_k^*(v_k^*)^T)^T\\ =& R_{k+1}H_k^T + R_{k+1}u_k^* s_k^T + R_{k+1}v_k^*(y_k^*)^T H_k^T\\ =& R_{k+1}H_k^T + [a_{12}\frac{A_k}{1+a_{22}(y_k^*)^T H_k^T g_{k+1}}R_{k+1}R_kH_k^T s_k + \rho\frac{R_{k+1}s_k}{s_k^T y_k^*}]s_k^T +\\ & [a_{22}\frac{A_k}{1+a_{22}(y_k^*)^T H_k^T g_{k+1}}R_{k+1}R_kH_k^T s_k - \frac{R_{k+1}s_k}{s_k^T y_k^*}](y_k^*)^T H_k^T\\ =& R_{k+1}H_k^T - R_{k+1}\frac{s_k(y_k^*)^T}{s_k^T y_k^*}H_k^T + \rho\frac{R_{k+1}s_k s_k^T}{s_k^T y_k^*} +\\ & \frac{A_k}{1+a_{22}(y_k^*)^T H_k^T g_{k+1}}R_{k+1}R_kH_k^T s_k(a_{12}s_k + a_{22}H_k y_k^*)^T\\ =& R_{k+1}R_kH_k^T + \rho\frac{R_{k+1}s_k s_k^T}{s_k^T y_k^*} +\\ & \frac{A_k}{1+a_{22}(y_k^*)^T H_k^T g_{k+1}}R_{k+1}R_kH_k^T s_k(a_{12}s_k + a_{22}H_k y_k^*)^T\end{aligned}$$

Hence,

$$R_{k+1}H_{k+1}^T = R_{k+1}R_kH_k^T B_k + \rho R_{k+1}\frac{s_k s_k^T}{s_k^T y_k^*}, \tag{3.14}$$

where

$$B_k = I + \frac{A_k}{1+a_{22}(y_k^*)^T H_k^T g_{k+1}}s_k(a_{12}s_k + a_{22}H_k y_k^*)^T, \quad k = 0,1,2,\cdots$$

By letting $B_{-1} = I$ and I be the identity matrix, we obtain

$$\begin{aligned}R_kH_k^T =& R_kR_{k-1}H_{k-1}^T B_{k-1} + \rho R_k\frac{s_{k-1}s_{k-1}^T}{s_{k-1}^T y_{k-1}^*}\\ =& \cdots\\ =& \prod_{j=0}^{k}(R_jH_0^T B_{j-1}) + \rho R_k\sum_{j=0}^{k-2}[\prod_{i=j+1}^{k-1}(R_i\frac{s_js_j^T}{s_j^T y_j^*}B_i)] + \rho R_k\frac{s_{k-1}s_{k-1}^T}{s_{k-1}y_{k-1}^*},\end{aligned}$$

which combining with (3.6) yields that the search direction d_{k+1} is dependent only on the parameter ρ. This complete the proof.

For different A_k, $B_k = I + \frac{A_k}{1+a_{22}(y_k^*)^T H_k^T g_{k+1}}s_k(a_{12}s_k + a_{22}H_k y_k^*)^T$ is different. Consequently, $R_kH_k^T$ is also different. This makes the sequence of $\{x_k\}$ different. For example, if let $A_k = A_k(2) = \mu_k a_k$, then we have

$$\begin{aligned}R_kH_k^T =& \prod_{j=0}^{k}\{R_jH_0^T[I + \frac{\mu_{j-1}a_{j-1}}{1+a_{22}(y_{j-1}^*)^T H_{j-1}^T g_j}s_{j-1}(a_{12}s_{j-1} + a_{22}H_{j-1}(y_{j-1}^*)^T)]\} +\\ & R_k\sum_{j=0}^{k-2}\{\prod_{i=j+1}^{k-1}[R_i\frac{s_j(s_j)^T}{s_j^T y_j^*}(I + \frac{\mu_i a_i}{1+a_{22}(y_i^*)^T H_i^T g_{i+1}}s_i(a_{12}s_i + a_{22}H_i y_i^*)^T)]\} +\\ & \rho R_k\frac{s_{k-1}s_{k-1}^T}{s_{k-1}y_{k-1}^*}.\end{aligned}$$

4. Disscussion

In this paper, we prove that all of the M-Huang class have the same search direction and that the iteration is dependent only on the parameter ρ if the exact line searches are made in each steps. But the property that Algorithm 2.1 stops in finite steps for convex quadratic programming has not been proved yet. This might be an important topic of further research.

References:

[1] LI Dong-hui, FUKUSHIMA M. *A modified BFGS method and its global convergence in nonconvex minimization* [J]. J. Comput. Appl. Math., 2001, **129**: 15–35.
[2] LI Dong-hui, FUKUSHIMA M. *A global and superlinear convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations* [J]. SIAM J. Numer. Anal., 1981, **37**: 17–41.

无约束最优化问题中一类新黄族及其性质

韦增欣 [1]，李桥兴 [2]
(1. 广西大学数学与信息科学学院，广西 南宁 530004;
2. 南京航空航天大学经济与管理学院，江苏 南京 210016)

摘要: 本文提出了一类新的用于解决无约束最优化问题的拟牛顿方法，并证明了这样的性质，在精确线性搜索条件下，每一步该族所有方法所产生的迭代方向和迭代点列仅依赖于参数 ρ. 该方法可视为拟牛顿方法中黄族的推广.

关键词: 无约束最优化；拟牛顿方程；拟牛顿方法.