

## 变尺度法与 Huang 族公式

杨丽君, 王保保

(西安电子科技大学 计算机网络与计算研究所, 陕西 西安 710071)

**摘 要:** 变尺度算法是求解无约束优化问题的有效而著名的算法, 算法易于实现, 计算量较小, 并形成了完整的算法体系, 对工程应用有重要的影响。文中简述了变尺度算法的基本思想, 介绍了 Huang 算法族和 Broyden 族及它们的基本性质, 分析比较了近年来对变尺度算法不同角度的研究。通过数值实验证明了变尺度算法的有效性。

**关键词:** 变尺度算法; Huang 算法族; Broyden 族

**中图分类号:** TP301.6

**文献标识码:** A

**文章编号:** 1005-3751(2006)02-0160-03

## Variable Metric Method and Huang Family Formula

YANG Li-jun, WANG Bao-bao

(Computer Network and Computing Institute, Xidian University, Xi'an 710071, China)

**Abstract:** The variable metric method is an effective and famous method, and can be easily computed. The variable metric method had formed an algorithm system and had a very important sense on the application. This paper discusses the variable metric method, deals with the algorithm of Huang family and its important properties with an analysis on the variable metric methods.

**Key words:** variable metric method; Huang family algorithm; Broyden family

## 0 引言

变尺度算法最早由 Davidon 提出, 简称 DFP 法。这种方法的出现, 使最优化方法的研究发生了一个飞跃, 随后出现了很多优秀的算法。20 世纪 70 年代初形成了 Huang 族算法和 Broyden 族算法, Broyden 族变尺度法可以看作 Huang 族变尺度法的一个子族。Huang 族变尺度算法族在无约束优化中有着重要的地位, 由于工业流程中的问题大多是连续可导的, 将变尺度法与同属于导数下降方法的公认较为有效的几种优化算法进行比较, 证明了兼有共轭方向法和牛顿法的特性的变尺度法是广泛使用的有效的无约束最优化方法。

## 1 变尺度法与 Huang 族公式

考虑无约束优化问题

$$\min f(x) \quad (1)$$

其中  $x \in R^n$ ,  $f: R^n \rightarrow R$  是一个连续可微函数。

为了避免计算目标函数  $f(x)$  的 Hesse 阵  $H(x^k)$  及其逆阵  $H(x^k)^{-1}$ , 考虑构造一个新的矩阵  $H_k$  (仅用目标函数的一阶导数)。用它去直接逼近 Hesse 阵的逆阵  $H(x^k)^{-1}$ , 这种方法称为拟牛顿法, 由于这类方法每一轮的迭代相当

于该变尺度后的最速下降法, 故又称为变尺度法或大步梯度法。算法有超线性收敛速度, 又不需要计算 Hesse 矩阵, 是目前最受重视、理论上最成熟的方法之一。

用不同的方法去构造不同的逼近矩阵序列  $H_k$  ( $k = 0, 1, \dots, n-1$ ), 就得到不同的变尺度法。1970 年, Huang 提出了一类算法族, 该算法族包含 3 个自由参数, 当算法中的  $H_k$  满足对称性要求且参数中的  $\omega = 1$  时, Huang 算法族就变为含一个参数的 Broyden 族, 因此 Broyden 算法族是 Huang 算法族的子族。记校正矩阵为  $H_k$ , 迭代方向为:

$$z^k = -H_k g_k \quad (2)$$

按照上述条件, 导出了 Huang 族校正公式的形式及关系式如下:

$$H_{k+1} = H_k + \Delta x_k u_k^T + H_k \Delta g_k v_k^T \quad (3)$$

$$\text{其中: } u_k = a_k \Delta x_k + b_k H_k^T \Delta g_k \quad (4)$$

$$v_k = c_k \Delta x_k + d_k H_k^T \Delta g_k \quad (5)$$

$$\text{且满足: } u_k^T \Delta g_k = w \quad (6)$$

$$v_k^T \Delta g_k = -1 \quad (7)$$

由于关系式(6)、(7)中的 5 个参数  $a_k, b_k, c_k, d_k, w$  中只有 3 个自由参数, 故在公式(3)中, 利用条件(6)、(7), 可得:

$$H_{k+1} \Delta g_k = w \Delta x_k \quad (8)$$

当  $w = 1$  时, 式(8)变为拟牛顿方程, 因而  $w \neq 1$  时, Huang 族不是拟牛顿算法族, 通常称为 Huang 变尺度算法族。现在, 如果在式(3)中, 假定  $H_k$  对称, 也要求  $H_{k+1}$  对称, 并令  $w = 1$ , 由  $u, v$  的(4)、(5)的表示式及式(3)可

收稿日期: 2005-05-17

作者简介: 杨丽君(1980—), 女, 云南人, 硕士研究生, 研究方向为图像处理、智能计算; 王保保, 副教授, 研究方向为图像处理、数据库和智能计算等。

知,若 $H_k$ 对称, $H_{k+1}$ 也对称,则必须有 $b_k = c_k$ ,又 $w = 1$ ,因而3个参数变为只有一个自由参数。现在,令 $b_k = c_k, w = 1$ ,将式(4)、(5)代入式(6)、(7)中,令 $a_k$ 为自由参数,解出其余参数的表达式,再将 $u, v$ 代入式(3),以得到 $H_{k+1}$ 的公式表示。将 $u, v$ 代入式(6)、(7),可得方程组:

$$\begin{cases} a_k \Delta x_k^T \Delta g_k + b_k \Delta g_k^T H_k \Delta g_k = 1 \\ b_k \Delta x_k^T \Delta g_k + d_k \Delta g_k^T H_k \Delta g_k = -1 \end{cases} \quad (9)$$

化简可得:

$$d_k = \frac{a_k (\Delta x_k^T \Delta g_k)^2 - \Delta x_k^T \Delta g_k - \Delta g_k^T H_k \Delta g_k}{(\Delta g_k^T H_k \Delta g_k)^2} \quad (10)$$

$$\begin{aligned} b_k = c_k &= \frac{1}{\Delta x_k^T \Delta g_k} (-1 - d_k \Delta g_k^T H_k \Delta g_k) \\ &= \frac{1 - a_k \Delta x_k^T \Delta g_k}{\Delta g_k^T H_k \Delta g_k} \end{aligned} \quad (11)$$

将式(4)、(5)、(10)、(11)代入式(3)中可得:

$$H_{k+1} = \left[ H_k + \frac{\Delta x_k \Delta x_k^T}{\Delta x_k^T \Delta g_k} - \frac{H_k \Delta g_k \Delta g_k^T H_k}{\Delta g_k^T H_k \Delta g_k} \right] + \Delta \bar{H}_k \quad (12)$$

经过整理简化得:

$$\begin{aligned} \Delta \bar{H}_k &= \frac{a_k (\Delta x_k^T \Delta g_k)^2 - \Delta x_k^T \Delta g_k}{\Delta g_k^T H_k \Delta g_k} \Delta g_k^T H_k \Delta g_k^T \cdot \\ &\quad \left[ \frac{\Delta x_k}{\Delta x_k^T \Delta g_k} - \frac{H_k \Delta g_k}{\Delta g_k^T H_k \Delta g_k} \right] \left[ \frac{\Delta x_k}{\Delta x_k^T \Delta g_k} - \frac{H_k \Delta g_k}{\Delta g_k^T H_k \Delta g_k} \right] \end{aligned} \quad (13)$$

由此,令 $b_k = c_k = \beta$ ,得到Broyden族拟牛顿公式:

$$\begin{aligned} H_{k+1} &= H_k + \frac{\Delta x_k \Delta x_k^T}{\Delta g_k^T \Delta x_k} - \frac{H_k \Delta g_k \Delta g_k^T H_k}{\Delta g_k^T H_k \Delta g_k} - \beta \cdot \\ &\quad \left[ \frac{\Delta g_k^T H_k \Delta g_k}{\Delta g_k^T \Delta x_k} \Delta x_k \Delta x_k^T - (\Delta x_k^T H_k \Delta g_k + \frac{\Delta g_k^T \Delta x_k}{\Delta g_k^T H_k \Delta g_k} H_k \Delta g_k \Delta g_k^T H_k) \right] \end{aligned} \quad (14)$$

此时,令 $\beta = 0$ ,即 $b_k = c_k = 0$ ,就得到了著名的DFP(Davidon - Fletcher - Powell)变尺度公式:

$$H_{k+1} = H_k + \frac{\Delta x_k \Delta x_k^T}{\Delta g_k^T \Delta x_k} - \frac{H_k \Delta g_k \Delta g_k^T H_k}{\Delta g_k^T H_k \Delta g_k} \quad (15)$$

在DFP变尺度公式中,如果目标函数为二次函数,令 $H_0 = I$ (单位矩阵),得到了共轭梯度法公式。

在Huang族公式中,令 $w = 1, b_k = c_k, d_k = 0$ ,即在子族Broyden族公式中令 $\beta = \frac{1}{\Delta g_k \Delta x_k^T}$ ,即得到BFGS变尺度公式:

$$\begin{aligned} H_{k+1} &= H_k + \left[ 1 + \frac{\Delta g_k^T H_k \Delta g_k}{\Delta g_k^T \Delta x_k} \right] \cdot \frac{\Delta x_k \Delta x_k^T}{\Delta g_k^T \Delta g_k} - \\ &\quad \frac{H_k \Delta g_k \Delta x_k^T}{\Delta g_k^T \Delta x_k} - \frac{\Delta x_k^T \Delta g_k H_k}{\Delta g_k^T \Delta x_k} \end{aligned} \quad (16)$$

在Huang族公式中,令 $w = 1, b_k = c_k = -\frac{1}{(\Delta x_k - H_k \Delta g_k)^T \Delta g_k}$ ,得Murtagh - Sargent公式:

$$H_{k+1} = H_k + \frac{(\Delta x_k - H_k \Delta g_k)(\Delta x_k - H_k \Delta g_k)^T}{(\Delta x_k - H_k \Delta g_k)^T \Delta g_k} \quad (17)$$

在Huang族公式中,令 $a_k = c_k = 0, b_k = -d_k$ ,得Pearson公式:

$$H_{k+1} = H_k + \frac{(\Delta x_k - H_k \Delta g_k) \Delta g_k^T H_k}{\Delta g_k^T H_k \Delta g_k} \quad (18)$$

1972年,Dixon发表了关于Huang算法族的性质的著名定理,深刻揭示了Huang算法族的重要特性。从Dixon定理可知,Broyden拟牛顿算法族,当 $x^1, H_1$ 相同时,算法都产生相同的迭代点列。除上述的性质外,拟牛顿算法具有可逆线性变换的不变性。

## 2 变尺度算法与几种优化算法的比较

最速下降法、牛顿法、共轭梯度法同变尺度法同属于导数下降法中几种有效的优化算法。最速下降法是以函数的一次近似而提出的算法,刚开始几步收敛较快,往后却越来越慢。牛顿法、共轭梯度法、变尺度算法则是以函数的二次近似为基础的,有较快的收敛速度。但牛顿法需要计算 $[\Delta^2 f(x)]^{-1}$ ,计算量和存储量都很大,在工程应用中并不实用。共轭梯度法中的搜索方向是与二次函数系数矩阵有关的所谓共轭方向,根据一维搜索确定步长,不需要存储二阶导数,二次函数理论上不超过 $n$ 次即可收敛,但对于一般非线性函数,共轭梯度法的收敛速度仅为线性。变尺度法是求解无约束极值问题的一种有效算法。它避免了计算二阶导数矩阵及其求逆计算,又比梯度法的收敛速度快,特别是对高维问题具有显著的优越性。DFP和BFGS变尺度算法是Huang算法族中的两个著名算法。下面用数值实验方法来证明变尺度算法的有效性。给出一个典型的算例(它是E. M. L. Beale (1958年)提出的,一再被用来比较不同算法的相对好坏,其最优值为 $[3, 0.5]$ ):

$$f(x) = [1.5 - x_1(1 - x_2)]^2 + [2.25 - x_1(1 - x_2^2)]^2 + [2.625 - x_1(1 - x_2^3)]^2 \quad (19)$$

四种算法的求解结果如表1所示。

表1 四种算法的求解结果

算法	迭代次数	$\ g\ $	计算时间(s)
最速下降法	273	1.445e-007	6.730
共轭梯度法	1600	7.925e-08	3.62
DFP法	47	2.105e-07	0.21
BFGS法	24	1.601e-08	0.11

文中所述的四种算法全部采用了数值逼近的方法来计算函数的梯度,使用的寻优方法是黄金分割法。采用不同的寻优方法,也将对算法的有效性产生极大的影响。由实验结果可以看出,对于中小规模的无约束优化问题,变尺度算法尤其是BFGS是十分有效的。基于BFGS方法的有效性,学者们对变尺度算法进行了更深的研究,提出了更多有效的算法。

## 3 Huang族公式的推广应用

在Huang族公式中,当 $w = 1$ 时,此时的Huang族是拟牛顿算法族。虽然拟牛顿算法的理论已经十分成熟,但如何使拟牛顿法更快地修正矩阵,更适合于并行计算,仍

值得研究。拟牛顿算法的推广有了很多研究成果,一类是从算法的性质与结构分析入手进行推广,一类从目标函数的近似表示形式入手。

### 3.1 具有 $n+1$ 个参数的拟牛顿算法

在 Huang 算法族的基础上,吴方一、桂湘云<sup>[1]</sup>根据 Huang 族的公式结构与性质进行了推广,得到了含有  $n+1$  个参数的新算法族, Huang 算法族的公式可写成:

$$\begin{aligned} H_{k+1} = & H_k + \Delta x_k [a_k \Delta x_k + b_k H_k^T \Delta g_k]^T + \\ & H_k \Delta g_k [c_k \Delta x_k + d_k H_k^T \Delta g_k]^T \end{aligned} \quad (20)$$

### 3.2 修正拟牛顿算法

(1) 1971 年, Biggs<sup>[2]</sup> 给出了拟牛顿算法的一种修正形式,修正公式为:

$$H_{k+1} = H_k - \frac{H_k \Delta x_k \Delta x_k^T H_k}{\Delta x_k^T H_k \Delta x_k} + t_k \frac{\Delta g_k \Delta g_k^T}{\Delta x_k^T \Delta g_k}$$

其中  $t_k$  为参数,

$$t_k = \frac{6}{\Delta x_k^T \Delta g_k} [f(x^k) - f(x^{k+1}) + \Delta x_k^T g_{k+1}] - 2 \quad (21)$$

这个修正公式与 BFGS 公式比较,只差参数  $t_k$ , 当  $t_k = 1$  时,式(21)即为 BFGS 公式。

(2) 1991 年,袁亚湘<sup>[3]</sup>根据  $f(x)$  的二次近似与某些插值条件,得到与式(20)类似的修正公式,他的公式中  $t_k$  的取值为:

$$t_k = \frac{2}{\Delta x_k^T \Delta g_k} [f(x^k) - f(x^{k+1}) + \Delta x_k^T g_{k+1}] \quad (22)$$

(3) 焦宝聪<sup>[4]</sup>基于目标函数局部二次模型近似,亦提出了一个类似公式,其中  $t_k$  改取为:

$$\begin{aligned} t_k = & \frac{2(1-\theta)}{\Delta x_k^T \Delta g_k} [f(x^k) - f(x^{k+1}) - \Delta x_k^T g_{k+1}] + \theta \\ & \theta \in [0, 1] \end{aligned} \quad (23)$$

(4) Lalee 和 Nocedal<sup>[5]</sup>在 1991 年提出了一列自调节 BFGS 方法,并在非精确搜索的假定下证明了该方法对于一致凸函数是收敛的,而且局部收敛速度是超线性的。

### 3.3 Huang 族变尺度算法与其他算法的综合应用

(1) 有限内存 BFGS 方法<sup>[5]</sup>。

对于中小规模的无约束优化问题, BFGS 是十分有效的。但对于大规模问题,算法所需内存相当重要。有限内存 BFGS 最早由 Perry(1997)和 Shanno(1978)提出。它可以看成是共轭梯度法的一种推广。

(2) 加速混沌变尺度混合优化算法<sup>[6]</sup>。

由于 DFP 法的有效性受初始选择点的影响,而加速混沌法往往造成优化时间的浪费。为弥补两算法的不足,胡云昌等人将加速混沌化算法与 DFP 变尺度算法相结合,利用了混沌寻优的性质(遍历性、随机性和规律性)、变尺度法的迅速下降的性质,提出了称之为加速混沌变尺度混合优化的算法,使得算法的寻优效率有了进一步提高。

(3) 一类既约变尺度算法<sup>[7]</sup>。

赖炎连等人将既约梯度法与 Huang 族变尺度法结合,提出了一类解非线性约束凸规划的既约变尺度算法。算法具有全局收敛性和超线性收敛速度。

## 4 结 论

Huang 变尺度算法族在无约束优化中的地位举足轻重,变尺度法是广泛使用的有效的无约束最优化方法。Huang 族公式的推广应用,变尺度算法与其他算法更有效的结合将有更大的应用前景,值得更进一步的研究和探讨。

### 参考文献:

- [1] 吴方,桂湘云. 一类具有  $n+1$  个参数的变尺度算法[J]. 数学学报, 1981(24): 921-930.
- [2] Biggs M C. Minimization algorithms making use of non-quadratic properties of the objective Function[J]. J. of institute of Mathematics and Its Application, 1987(8): 315-327.
- [3] Yuan Y. A modified BFGS algorithms for unconstrained optimization[J]. IMA, J. of Numerical Analysis, 1991(11): 325-332.
- [4] 焦宝聪. 一类改进 BFGS 算法及其收敛性分析[J]. 数学的实践与认识, 1994, 24(2): 143-149.
- [5] 袁亚湘. 非线性规划数值方法[M]. 上海: 上海科学出版社, 1993. 114-116.
- [6] 胡云昌. 加速混沌变尺度混合优化算法[J]. 天津大学学报, 2002, 35(1): 68-70.
- [7] 简金宝. 一类超线性收敛的既约变尺度法[J]. 系统科学与数学, 1996, 16(1): 78-80.

(上接第 159 页)

### 参考文献:

- [1] Perkowski M, Etzioni O. Towards adaptive Web sites: Conceptual framework and case study[J]. Artificial Intelligence, 2000, 118(1): 245-275.
- [2] Xing D, Shen J. Efficient data mining for web navigation patterns[J]. Information and Software Technology, 2004, 46(1): 55-63.
- [3] Mobasher B, Cooley R. Creating adaptive Web sites through usage-based clustering of URLs[A]. Proc of the 1999 IEEE Knowledge and Data Engineering Exchange Workshop (KDEX'99)[C]. New York: IEEE Press, 1999. 32-37.
- [4] Mobasher B, Dai H, Luo T, et al. Integrating Web Usage and Content Mining for More Effective Personalization[A]. Proceedings of the International Conference on E-commerce and Web Technologies (ECWeb2000) [C]. Greenwich, UK: Springer-Verlag, 2000. 165-176.
- [5] Dunham M H. Data Mining: Introductory and Advanced Topics[M]. [s.l.]: Prentice Hall, 2003.