中国农业大学学报　2001, 6(6): 23～26
Journal of China Agricultural University

迭代算法的广义 Q-收敛阶和效率[①]

钟　萍[②]　张春华　张海斌
(中国农业大学工程基础科学部)

摘　要　推广了迭代算法收敛分析中的 Q-收敛阶的概念, 据此给出了算法效率的一种一般的度量。分析了新效率定义与已有的 Ostrowski 效率和 Brent 效率之间的关系。这种度量适用于任何迭代算法, 因而为分析算法的优劣提供了一个理论依据。

关键词　广义 Q-收敛阶; Ostrowski 效率; Brent 效率
中图分类号　O 221.2

Extension Definition on Q-rates of Convergence and Efficiency

Zhong Ping　Zhang Chunhua　Zhang Haibin
(College of Applied Engineering Sciences, CAU)

Abstract　The definition of Q-rates of convergence in iterative algorithms is extened and the efficiency measure is developed. The relationship between the new definition of efficiency and the old two definitions of efficiency——Ostrowski efficiency and Brent efficiency was also analysed. Since this measure is suitable to every iterative algorithm, it provide a theoretical justification for the algorithm.

Key words　generalized Q-rates of convergence; Ostrowski efficiency; Brent efficiency

效率是最优化理论中用来辨别迭代算法优劣的重要指标之一, 主要由以下 2 个方面决定: 1) 迭代序列收敛的快慢, 即一个算法的收敛阶。2) 迭代过程中每步的计算量。显然, 收敛速度快而计算量小的算法是效率高的算法。早在 1960 年, Ostrowski 给出了一个较为直观的效率定义[1]:

定义 1　效率 Γ_O。设 $\{x^k\}$ 是由某个算法产生的迭代序列, 并且以 Q-收敛阶 α 的速度收敛于点 x^*。若由 x^k 计算 x^{k+1} 所需的计算量 $Q[x^k, x^{k+1}] = Q$, 其中 Q 是一个常数, 那么这个算法的效率 Γ_O 为

$$\Gamma_O = \frac{\ln\alpha}{Q} \tag{1}$$

由于式(1)要求每步迭代的计算量为常数, 显然效率 Γ_O 有很大的局限性。这种定义方式势必会使算法的效率度量受到限制。1973 年, Brent 给出了一个更为一般的定义[2]:

定义 2　效率 Γ_B。设 $\{x^k\}$ 是由某个算法产生的迭代序列, 并且收敛于点 x^*, 那么这个算法的效率 Γ_B 为

收稿日期: 2000-07-09
①国家自然科学基金资助项目
②钟　萍, 北京清华东路 17 号 中国农业大学(东校区) 71 信箱, 100083

$$\Gamma_B = \liminf_{k\to\infty} \frac{\ln(-\ln\|x^k - x^*\|)}{\sum_{i=1}^{k} Q_i[x^{i-1}, x^i]} \tag{2}$$

其中$Q[x^{i-1}, x^i]$是由x^{i-1}计算x^i所需的计算量。

效率Γ_B是Γ_O的一种拓广。当存在$Q = \lim_{k\to\infty} Q_k > 0$时, Γ_B就是Γ_O[2]。然而在Γ_B定义中, $\{x^k\}$的收敛阶是以R-收敛阶$\rho = \liminf_{k\to\infty}(-\ln\|x^k - x^*\|)^{1/k}$为基准的。$R$-收敛阶体现了$\{x^k\}$的平均收敛速度, 它比$Q$-收敛阶要弱, 因而在优化算法的收敛分析中大都采用Q-收敛阶[3]。这样, 在进行效率分析时, 可以有效地借助于优化算法收敛分析中已有的深刻结论, 得到与优化中常用的Q-收敛阶相一致的结果。本文中给出了一个较Ostrowski的效率定义更为一般的关于Q-收敛阶的效率定义, 并阐明它与以上 2 种效率定义之间的关系。

1 广义Q-收敛阶

为了给出一个一般的Q-收敛阶的效率定义, 首先引进"进展指标"和"广义Q-收敛阶"的概念。

定义 3 进展指标。设迭代点x_+和x_c在点x^*的$\delta(\delta > 0)$邻域内, 从x_c到x_+的关于x^*的进展指标ν定义为

$$\nu = \nu[x_c, x_+, x^*] = \frac{\ln\|x_+ - x^*\|}{\ln\|x_c - x^*\|} \tag{3}$$

定义 4 广义Q-收敛阶。设$\{x^k\}$是由某个算法产生的迭代序列, 并且收敛于点x^*, 那么它的广义Q-收敛阶定义为

$$\rho(\{x^k\}) = \liminf_{k\to\infty} \nu[x^{k-1}, x^k, x^*] \tag{4}$$

其中$\nu[\cdot, \cdot, \cdot]$是由式(3)定义的进展指标。

下面的定理表明广义Q-收敛阶是通常Q-收敛阶的拓广。

定理 1 设$\{x^k\}$至少以Q-收敛阶α的速度收敛于点x^*, 即存在$\delta(\delta > 0)$和常数C_1 $(C_1 > 0)$, 使得当$\|x^k - x^*\| < \delta$时

$$\|x^{k+1} - x^*\| \le C_1\|x^k - x^*\|^\alpha \tag{5}$$

则

$$\rho(\{x^k\}) \ge \alpha \tag{6}$$

特别地, 当$\{x^k\}$恰以Q-收敛阶α的速度收敛于点x^*, 即存在$\delta(\delta > 0)$和常数$C_3(C_2 > C_3 > 0)$, 使得当$\|x^k - x^*\| < \delta$时

$$C_3\|x^k - x^*\|^\alpha \le \|x^{k+1} - x^*\| \le C_2\|x^k - x^*\|^\alpha \tag{7}$$

则

$$\rho(\{x^k\}) = \alpha \tag{8}$$

证明 由式(5)可得

$$\frac{\ln\|x^{k+1} - x^*\|}{\ln\|x^k - x^*\|} \ge \alpha + \frac{\ln C_1}{\ln\|x^k - x^*\|} \tag{9}$$

由定义 3, 式(6)成立; 由式(7)可以类似证明式(8)。

2　效率 Γ 与 Γ_O, Γ_B 之间的关系

首先给出与 Q-收敛阶相一致的效率定义 Γ, 然后阐明它与 Γ_O 以及 Γ_B 之间的关系。

定义 5　效率 Γ。设 $\{x^k\}$ 是由某个算法产生的迭代序列, 并且收敛于 x^*, 那么这个算法的效率 Γ 定义为

$$\Gamma = \inf_{\{x^k\}} \sup_{\{y^k\}\subset\{x^k\}} \frac{\ln(\rho(\{y^k\}))}{\limsup_{k\to\infty} Q[y^{k-1}, y^k]} \tag{10}$$

其中: $\{y^k\}$ 是 $\{x^k\}$ 的任意一个子序列, $\rho(\{y^k\})$ 是 $\{y^k\}$ 的广义 Q-收敛阶, $Q[y^{k-1}, y^k]$ 是由 y^{k-1} 计算 y^k 所需的计算量。

由定理 1 和定义 5 可以得到 Γ 与 Γ_O 之间的关系。

定理 2　设 $\{x^k\}$ 是由某个算法产生的迭代序列, 并且收敛于点 x^*。若存在 $Q = \lim_{k\to\infty} Q[x^{k-1}, x^k] > 0$, 且 $\{x^k\}$ 以 Q-收敛阶 α 收敛到 x^*, 则 $\Gamma \geq \Gamma_O$; 特别地, 若 $\{x^k\}$ 恰以 Q-收敛阶 α 收敛到 x^*, 则 $\Gamma = \Gamma_O$。

下面的定理阐述了 Γ 与 Γ_B 之间的关系。

定理 3　设 $\{x^k\}$ 是由某个算法产生的迭代序列, 并且收敛于点 x^*。若 $Q = \liminf_{k\to\infty} Q[x^{k-1}, x^k] > 0$, 则 $\Gamma_B \geq \Gamma$; 特别地, 若存在 $Q = \lim_{k\to\infty} Q[x^{k-1}, x^k] > 0$ 时, $\Gamma_B = \Gamma$。

证明　设 $\{y^k\}$ 是 $\{x^k\}$ 的任意一个子序列, 并且不妨设存在 $\delta \in (0, 1)$, 使得 $\|y^1 - x^*\| < \delta$, 由定义 3 可得

$$\ln\|y^k - x^*\| = \frac{\ln\|y^k - x^*\|}{\ln\|y^{k-1} - x^*\|} \cdot \frac{\ln\|y^{k-1} - x^*\|}{\ln\|y^{k-2} - x^*\|} \cdot \cdots\cdots \cdot \frac{\ln\|y^2 - x^*\|}{\ln\|y^1 - x^*\|} \cdot \ln\|y^1 - x^*\| =$$

$$v[y^{k-1}, y^k, x^*] \cdot v[y^{k-2}, y^{k-1}, x^*] \cdot \cdots\cdots \cdot v[y^1, y^2, x^*] \cdot \ln\|y^1 - x^*\| \tag{11}$$

所以

$$\frac{\ln(-\ln\|y^k - x^*\|)}{\sum_{i=1}^{k} Q[y^{i-1}, y^i]} = X_1 + X_2 \tag{12}$$

其中

$$X_1 = \frac{\sum_{i=1}^{k} \ln v[y^{i-1}, y^i, x^*]}{\sum_{i=1}^{k} Q[y^{i-1}, y^i]} \tag{13}$$

$$X_2 = \frac{\ln(-\ln\|y^1 - x^*\|)}{\sum_{i=1}^{k} Q[y^{i-1}, y^i]} \tag{14}$$

令 $Q_i = \dfrac{Q[y^{i-1}, y^i]}{Q}$, 则

$$\lim_{k\to\infty} \sum_{i=1}^{k} Q_i = \infty$$

从而

$$\lim_{k\to\infty} X_2 = \frac{\ln(-\ln\|y^1 - x^*\|)}{\lim\limits_{k\to\infty}\sum\limits_{i=1}^{k} Q_i} \cdot \frac{1}{Q} = 0 \tag{15}$$

由式(13)和定义 4 知

$$\liminf_{k\to\infty} X_1 \geq \frac{\ln\rho(\{y^k\})}{\limsup\limits_{k\to\infty} Q[y^{k-1}, y^k]} \tag{16}$$

由式(13), (15), (16)和定义 5 知

$$\Gamma_B = \liminf_{k\to\infty} \frac{\ln(-\ln\|y^k - x^*\|)}{\sum\limits_{i=1}^{k} Q[y^{i-1}, y^i]} = \liminf_{k\to\infty} X_1 + \lim_{k\to\infty} X_1 \geq \Gamma$$

当 $Q = \lim\limits_{k\to\infty} Q[x^{k-1}, x^k] > 0$ 时, 式(16)的等式成立。所以

$$\Gamma_B = \Gamma$$

本研究得到邓乃扬教授的指导, 谨致谢意。

参考文献

1 Ostrowski A. Solution of Equations and Systems of Equations. New York: Academic Press, 1960. 233～241

2 Brent R. Some efficient algorithms for solving systems of nonlinear equation. SIAM Journal on Numerical Analysis, 1973(10): 327～344

3 Nocedal J, Wright S J. Numerical Optimization. New York: Springer, 1999. 30