Advances in Applied Mathematics 应用数学进展, **2024, 13(6), 2930-2942**
Published Online June 2024 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2024.136280

多目标优化非单调线性加权牛顿算法

欧 杨

长沙理工大学数学与统计学院，湖南 长沙

收稿日期：2024年5月28日；录用日期：2024年6月22日；发布日期：2024年6月29日

摘 要

本文提出了一种求解多目标优化问题的非单调线性加权牛顿算法。在目标函数有下界且梯度Lipschitz连续的条件下证明了算法的全局收敛性。在目标函数满足二次连续可微且局部强凸的条件下，证明了该算法具有局部超线性收敛速度。数值实验结果表明，相比于单调线性加权牛顿算法，该算法能够更加高效地求解多目标优化问题。

关键词

多目标优化问题，非单调线搜索，Pareto最优解，局部超线性收敛

Nonmonotone Linear Weighted Newton Algorithm for Multi-Objective Optimization

Yang Ou

School of Mathematics and Statistics, Changsha University of Science and Technology, Changsha Hunan

Received: May 28th, 2024; accepted: Jun. 22nd, 2024; published: Jun. 29th, 2024

Abstract

This paper introduces a nonmonotone linear weighted Newton algorithm for solving multi-objective optimization problems. It is shown that the algorithm achieves global convergence under the assumption of a lower bound on the objective function and a Lipschitz continuous gradient. Furthermore, under the conditions of quadratic continuous differentiability and local strong convexity, the algorithm demonstrates local superlinear convergence. Numerical experiments demonstrate that this algorithm outperforms the monotone linear weighted Newton algorithm in efficiently solving multi-objective optimization problems.

文章引用: 欧杨. 多目标优化非单调线性加权牛顿算法[J]. 应用数学进展, 2024, 13(6): 2930-2942.
DOI: 10.12677/aam.2024.136280

Keywords

Multi-Objective Optimization Problem, Nonmonotone Line Search, Pareto Optimal Solution, Local Superlinear Convergence

1. 引言

多目标优化是最优化理论的一个分支，它与非线性规划、泛函分析以及矩阵论等有着密切的关系，在金融投资、物流网络及国防军事等方面有着广泛的应用需求。鉴于多目标优化在实际生活中的重要作用，学者们自上世纪七十年代以来，针对多目标优化问题开展了大量研究。到目前为止，多目标优化在理论方面已经取得很多重要成果，一套平行于单目标最优化的理论体系正日趋完善[1]，这为后续求解多目标优化问题提供了指导方向。

牛顿法具有较快的收敛速度，是求解单目标优化问题常用的方法之一。2009 年，Fliege [2]首次将牛顿法推广到多目标优化问题中。Fliege 通过求解一个极大极小问题来构造搜索方向，并且证明了当目标函数的二阶导数 Lipschitz 连续时，算法具有二阶收敛速度。虽然 Fliege 提出的多目标优化牛顿算法具有较快的收敛速度，但是在求解极大极小问题时的计算量较多，且计算难度大。Jiang [3]对 Fliege 提出的牛顿法进行改进，采用线性标量化方法，对目标函数的二次近似形式进行线性加权，将它们的和作为新的搜索方向，这样能够充分减少计算量，且算法在一定条件下同样具有二阶收敛速度。除牛顿法之外，最速下降法[4]、投影梯度法[5]、共轭梯度法[6]以及拟牛顿法[7]也已被扩展到多目标优化领域。

上述算法都是运用单调线搜索方法确定步长。对于单目标优化问题而言，单调线搜索方法非常有效，但是对于多目标优化问题来说，很难做到让每一个目标函数值在每一次迭代中都减小，且强制单调性会减慢求解方法的收敛速度，而非单调线搜索方法能有效避免这些不足。Zhang 和 Hager [8]提出一种平均值类型的非单调线搜索方法，并且证明了这种方法比单调线搜索方法和极大值非单调线搜索方法[9]更有效。

本文结合非单调线搜索技术，对 Jiang [3]提出的单调算法进行改进，提出了非单调线性加权牛顿法。在一定的假设条件下证明了算法的收敛性，数值实验结果表明，在相同的条件下，本文提出的算法在求解多目标优化问题时的迭代次数及时间更少，能更加有效地求解多目标优化问题。

本文考虑如下多目标优化问题：

$$\begin{aligned} \min\quad & F(x) \\ \text{s.t.}\quad & x\in U \end{aligned} \tag{1.1}$$

其中，$F:U\to\mathbb{R}^m$，$F(x)=\left(F_1(x),F_2(x),\cdots,F_m(x)\right)$，$U\subset\mathbb{R}^n$ 为开集。

在本文中，矩阵 $A\in\mathbb{R}^{m\times n}$ 的像空间用 Im(*A*)表示，$\operatorname{Im}(A)=\left\{Ax\mid x\in\mathbb{R}^n\right\}$。$DF(x)\in\mathbb{R}^{m\times n}$ 表示目标函数 F 在 x 处的 Jacobian 矩阵，$\nabla F_j(x)$ 表示第 j 个目标函数在 x 处的梯度，$\nabla^2 F_j(x)$ 表示第 j 个目标函数在 x 处的Hessian 矩阵，λ_j 表示目标函数的第 j 个分量对应的权重系数，$\sum_{j=1}^{m}\lambda_j=1$ 且 $\lambda_j>0$ $(j=1,2,\cdots,m)$。

在本文中，假设 F 在 U 上二次连续可微且目标函数每个分量 F_j 的 Hessian 矩阵都为正定矩阵，即

$$F\in C^2\left(U,\mathbb{R}^m\right),$$

$$\nabla^2 F_j\left(x\right)>0,\ \ \forall x\in U,\ j=1,2,\cdots,m.$$

在多目标优化问题中，考虑如下偏序关系：

$$\begin{aligned}\forall x,y\in\mathbb{R}^n, x=y &\Leftrightarrow x_i=y_i, i=1,2,\cdots,n.\\ x<y &\Leftrightarrow x_i<y_i, i=1,2,\cdots,n.\\ x\le y &\Leftrightarrow x_i\le y_i, \text{但 } x\ne y, i=1,2,\cdots,n.\end{aligned}$$

以下定义及引理均来源于文献[2]。

定义 1.1 设 $x^*\in U$，若不存在 $x\in U$，使得 $F\left(x\right)\le F\left(x^*\right)\left(F\left(x\right)<F\left(x^*\right)\right)$，则 x^* 是问题(1.1)的 Pareto 最优解或有效解(弱 Pareto 最优解或弱有效解)。

定义 1.2 设 $x^*\in U$，若存在 x^* 的一个邻域 $V\subseteq U$，且不存在 $x\in U\cap V$，使得 $F\left(x\right)\le F\left(x^*\right)\left(F\left(x\right)<F\left(x^*\right)\right)$，则 x^* 是问题(1.1)的局部 Pareto 最优解或局部有效解(局部弱 Pareto 最优解或局部弱有效解)。

定义 1.3 设 $\tilde{x}\in U$，若 $\mathrm{Im}\left(DF\left(\tilde{x}\right)\right)\cap\left(-\mathbb{R}^m_{++}\right)=\varnothing$，则 $\tilde{x}$ 为 F 的临界点(稳定点)。

引理 1.1 设 $\bar{x}\in U$，$F\in C^1\left(U,\mathbb{R}^m\right)$，则有如下结论：

1) 若 $\bar{x}$ 是局部弱 Pareto 最优解，则 $\bar{x}$ 是 F 的一个稳定点。

2) 若 U 是凸集，F 是凸向量函数，则当 $\bar{x}$ 是 F 的稳定点时，$\bar{x}$ 是弱 Pareto 最优解。

3) 若 $\forall x\in U$，有 $\nabla^2 F_j\left(x\right)>0\left(j=1,2,\cdots,m\right)$ 其中，U 是凸集，$F\in C^2\left(U,\mathbb{R}^m\right)$，则当 $\bar{x}\in U$ 是 F 的稳定点时，$\bar{x}$ 是 Pareto 最优解。

2. 非单调线性加权牛顿算法

考虑多目标优化问题(1.1)，在文献[2]中，定义 x 处的牛顿方向为 s_N，s_N 是如下无约束优化问题的最优解

$$\begin{cases}\min\ \max\limits_{j=1,2,\cdots,m} \nabla F_j\left(x\right)^{\mathrm{T}} s_N+\left(1/2\right)s_N^{\mathrm{T}}\nabla^2 F_j\left(x\right)s_N\\ \text{s.t.}\ \ s_N\in\mathbb{R}^n.\end{cases}\tag{2.1}$$

在上述问题中，$\nabla F_j\left(x\right)^{\mathrm{T}} s_N+\left(1/2\right)s_N^{\mathrm{T}}\nabla^2 F_j\left(x\right)s_N$ 是增量 $F_j\left(x+s_N\right)-F\left(x\right)$ 的二次近似，我们需要在增量最大的情况下求得最小的搜索方向 s_N。Fliege [2]将上述问题转化为约束优化问题，运用拉格朗日乘子法求得 s_N。对于多目标优化问题来说，这种求解搜索方向的方法较为复杂，所需的计算量也比较大。

本文采用加权系数法求解搜索方向 s，能够减少计算工作量，求解过程更加简便高效。考虑如下无约束优化问题：

$$\begin{cases}\min\sum\limits_{j=1}^{m}\lambda_j\left(\nabla F_j\left(x\right)^{\mathrm{T}} s+\left(1/2\right)s^{\mathrm{T}}\nabla^2 F_j\left(x\right)s\right)\\ \text{s.t.}\ \ s\in\mathbb{R}^n.\end{cases}\tag{2.2}$$

由假设，$\nabla^2 F_j\left(x\right)\left(j=1,2,\cdots,m\right)$ 正定，这说明 $\sum\limits_{j=1}^{m}\lambda_j\left(\nabla F_j\left(x\right)^{\mathrm{T}} s+\left(1/2\right)s^{\mathrm{T}}\nabla^2 F_j\left(x\right)s\right)$ 是强凸的，从而问题(2.2)存在唯一的极小值点。

用 $\theta\left(x\right)$ 表示问题(2.2)的最优值，则有

$$s\left(x\right)=\arg\min_{s\in\mathbb{R}^n}\sum_{j=1}^{m}\lambda_j\left(\nabla F_j\left(x\right)^{\mathrm{T}} s+\left(1/2\right)s^{\mathrm{T}}\nabla^2 F_j\left(x\right)s\right)\tag{2.3}$$

$$\theta(x)=\min_{s\in\mathbb{R}^n}\sum_{j=1}^{m}\lambda_j\left(\nabla F_j(x)^{\mathrm{T}}s+(1/2)s^{\mathrm{T}}\nabla^2F_j(x)s\right) \tag{2.4}$$

基于上述讨论，我们将给出求解多目标优化问题的非单调线性加权牛顿算法(Nonmonotone Linear Weighted Newton)，我们将它简称为 NLWN 算法，它的具体步骤如下：

算法 2.1 (NLWN 方法)

步 1 选取初始点 $x_0\in U$，选取参数 $\sigma\in(0,1)$，$\mu>0$，$\rho\in(0,1)$，$\eta\in[0,1]$，$k=0$，$C^0=F(x_0)$，$q_0=1$。

步 2 由(2.3)和(2.4)，分别求出 $s(x_k)$ 和 $\theta(x_k)$。

步 3 若 $\theta(x_k)=0$，终止算法；否则，转步 4。

步 4 记 h_k 是使得下式成立的最小非负整数 h：

$$F_j\left(x_k+\mu\rho^h s(x_k)\right)\le C_j^k+\sigma\mu\rho^h\theta(x_k),\ j=1,2,\cdots,m. \tag{2.5}$$

步 5 令 $\alpha_k=\mu\rho^{h_k}$，$x_{k+1}=x_k+\alpha_k s(x_k)$。

步 6 更新 q_k 和 C_j^k：

$$q_{k+1}=\eta q_k+1, \tag{2.6}$$

$$C_j^{k+1}=\eta q_k C_j^k+F_j(x_{k+1})/q_{k+1}. \tag{2.7}$$

令 $k=k+1$，转步 2。

注记 2.1 当 $\eta=0$ 时，则有 $C_j^{k+1}=F_j(x_{k+1})$。此时非单调线搜索退化成 Armijo 线搜索。

对于算法 2.1，我们有以下引理。

引理 2.1 [10]对于算法 2.1 的每一次迭代，总有 $F(x_k)\le C^k$。

引理 2.2 [10]设 $\{x_k\}$ 是由算法 2.1 产生的序列。若 x_k 不是稳定点，则存在步长 $\alpha_k>0$ 满足(2.5)中的非单调线搜索准则。

下面的引理给出 $\theta(x)$ 与 $s(x)$ 的性质。

引理 2.3 [2]在本文的假设下，有以下结论：

1) $\forall x\in U$，$\theta(x)\le 0$。

2) 以下三个说法是等价的：

a) x 不是稳定点。

b) $\theta(x)<0$。

c) $s(x)\ne 0$。

3) 由(2.3)给定的函数 $s:U\to\mathbb{R}^n$ 在紧集上有界，由(2.4)给定的函数 $\theta:U\to\mathbb{R}$ 在 U 上连续。

3. 全局收敛性

在这一小节，我们将证明算法 2.1 的全局收敛性。在证明算法 2.1 的全局收敛性之前，给出如下假设和引理：

假设 3.1 $\forall x\in\mathbb{R}^n$，$\nabla F_j(x)(j=1,2,\cdots,m)$ Lipschitz 连续。

引理 3.1 当假设 3.1 成立，且 α_k 满足非单调线搜索条件(2.5)时，下面的不等式成立：

$$\alpha_k\ge\min\left\{\rho,2\rho(1-\sigma)\left|\theta(x_k)\right|/L\left\|s(x_k)\right\|^2\right\}. \tag{3.1}$$

证 若 $\alpha_k\ge\rho$，则(3.1)成立。

若 $\alpha_k<\rho$，由 α_k 满足(2.5)式可知，$\exists j_0\in\{1,2,\cdots,m\}$，有

$$F_{j_0}\left(x_k+\left(\alpha_k/\rho\right)s\left(x_k\right)\right)>F_{j_0}\left(x_k\right)+\sigma\left(\alpha_k/\rho\right)\theta\left(x_k\right). \tag{3.2}$$

由假设 3.1，$\nabla F_j\left(j=1,2,\cdots,m\right)$ Lipschitz 连续，则有

$$\begin{aligned}&F_{j_0}\left[x_k+\left(\alpha_k/\rho\right)s\left(x_k\right)\right]-F_j\left(x_k\right)\\&=\left(\alpha_k/\rho\right)\nabla F_{j_0}\left(x_k\right)^{\mathrm{T}}s\left(x_k\right)+\int_0^{\alpha_k/\rho}\left[\nabla F_{j_0}\left(x_k+ts\left(x_k\right)\right)-\nabla F_{j_0}\left(x_k\right)\right]^{\mathrm{T}}s\left(x_k\right)\mathrm{d}t\\&\le\left(\alpha_k/\rho\right)\theta\left(x_k\right)+\int_0^{\alpha_k/\rho}tL\left\|s\left(x_k\right)\right\|^2\mathrm{d}t\\&=\left(\alpha_k/\rho\right)\theta\left(x_k\right)+L\alpha_k^2\left\|s\left(x_k\right)\right\|^2\Big/\left(2\rho^2\right).\end{aligned} \tag{3.3}$$

由(3.2)和(3.3)式可得

$$\sigma\left(\alpha_k/\rho\right)\theta\left(x_k\right)<\left(\alpha_k/\rho\right)\theta\left(x_k\right)+L\alpha_k^2\left\|s\left(x_k\right)\right\|^2\Big/\left(2\rho^2\right).$$

整理后可得

$$\alpha_k>2\rho\left(\sigma-1\right)\theta\left(x_k\right)\Big/L\left\|s\left(x_k\right)\right\|^2. \tag{3.4}$$

由引理 2.3 有 $\theta(x)\le 0$，从而 $\left|\theta(x)\right|=-\theta(x)$。此时(3.4)可写成

$$\alpha_k>2\rho\left(1-\sigma\right)\left|\theta\left(x_k\right)\right|\Big/L\left\|s\left(x_k\right)\right\|^2.$$

故引理 3.1 成立。

定理 3.1 假设 $F_j(x)$ 有下届 $\left(j=1,2,\cdots,m\right)$，$x\in\mathbb{R}^n$，$\eta<1$。当假设 3.1 成立，且存在一个常数 $c_1>0$ 使得下面的不等式成立时，

$$\left|\theta\left(x_k\right)\right|\ge c_1\left\|s\left(x_k\right)\right\|^2,\forall k=1,2,\cdots \tag{3.5}$$

由算法 2.1 产生的序列 $\{x_k\}$ 的任一极限点都是问题(1.1)的 Pareto 最优解。

证 首先证明下面的不等式，其中，$\beta=\min\left\{\sigma\rho,2\sigma\rho c_1\left(1-\sigma\right)/L\right\}$。

$$F_j\left(x_{k+1}\right)\le C_j^k-\beta\left|\theta\left(x_k\right)\right|\ \left(j=1,2,\cdots,m\right) \tag{3.6}$$

考虑下面两种情况：

(i) 使用非单调线搜索条件(2.5)及 $\alpha_k\ge\rho$，可得

$$F_j\left(x_{k+1}\right)\le C_j^k+\sigma\alpha_k\theta\left(x_k\right)=C_j^k-\sigma\alpha_k\left|\theta\left(x_k\right)\right|\le C_j^k-\sigma\rho\left|\theta\left(x_k\right)\right|\left(j=1,2,\cdots,m\right)$$

令 $\beta=\sigma\rho$，即得(3.6)。

(ii) 使用非单调线搜索条件(2.5)及 $\alpha_k<\rho$，由引理 3.1，可得

$$\alpha_k\ge 2\rho\left(1-\sigma\right)\left|\theta\left(x_k\right)\right|\Big/L\left\|s\left(x_k\right)\right\|^2.$$

再由(3.5)得

$$\begin{aligned}F_j\left(x_{k+1}\right)&\le C_j^k+\sigma\alpha_k\theta\left(x_k\right)=C_j^k-\sigma\alpha_k\left|\theta\left(x_k\right)\right|\\&\le C_j^k-2\sigma\rho\left(1-\sigma\right)\left|\theta\left(x_k\right)\right|^2\Big/L\left\|s\left(x_k\right)\right\|^2\\&\le C_j^k-2\sigma\rho c_1\left(1-\sigma\right)\left|\theta\left(x_k\right)\right|\Big/L\quad\left(j=1,2,\cdots,m\right)\end{aligned}$$

令 $\beta=2\sigma\rho c_1\left(1-\sigma\right)/L$，即得(3.6)。

设 $\hat{x}$ 是 $\{x_k\}$ 的极限点，接下来证明 $\hat{x}$ 收敛到问题(1.1)的 Pareto 最优解。

由 q_k 及 C_j^k 的更新公式(2.6)、(2.7)，以及(3.6)，有

$$q_{k+1}=1+\sum_{i=0}^{k}\eta^{i+1}. \tag{3.7}$$

$$\begin{aligned}C_j^{k+1}&=\left[\eta q_k C_j^k+F_j\left(x_{k+1}\right)\right]/q_{k+1}\\&\le\left[\eta q_k C_j^k+C_j^k-\beta\left|\theta\left(x_k\right)\right|\right]/q_{k+1}\\&=\left[\left(\eta q_k+1\right)C_j^k-\beta\left|\theta\left(x_k\right)\right|\right]/q_{k+1}\\&=\left[q_{k+1}C_j^k-\beta\left|\theta\left(x_k\right)\right|\right]/q_{k+1}\\&=C_j^k-\beta\left|\theta\left(x_k\right)\right|/q_{k+1}\quad\left(j=1,2,\cdots,m\right)\end{aligned} \tag{3.8}$$

由已知条件，$F_j(x)$ 有下届。由引理 2.1，有 $F_j\left(x_k\right)\le C_j^k$，从而 C_j^k 也有下界。$\left(j=1,2,\cdots,m\right)$

对(3.8)的前 $n+1$ 项求和，得到

$$\sum_{k=0}^{n}\left|\theta\left(x_k\right)\right|/q_{k+1}\le\left(1/\beta\right)\left(C_j^0-C_j^{n+1}\right).$$

令 $n\to+\infty$，得到

$$\sum_{k=0}^{+\infty}\left|\theta\left(x_k\right)\right|/q_{k+1}<\infty \tag{3.9}$$

$\hat{x}$ 是 $\{x_k\}$ 的极限点，设 K 是一个无限指标集，则有 $\lim\limits_{k\in K}x_k=\hat{x}$。我们要证明 $\hat{x}$ 是 Paroto 最优解，即证 $\theta(\hat{x})=0$，用反证法证明。

假设 $\theta(\hat{x})<0$，即

$$\exists\varepsilon_0>0，\ \delta_0>0，\ 0<\delta<\delta_0，\ \forall k\in K，\ \left\|x_k-\hat{x}\right\|\le\delta\text{ 时，有 }\left|\theta\left(x_k\right)\right|\ge\varepsilon_0。$$

从而有

$$\sum_{k=0}^{+\infty}\left|\theta\left(x_k\right)\right|/q_{k+1}\ge\sum_{k\in K,\left\|x_k-\hat{x}\right\|\le\delta}\varepsilon_0/q_{k+1}. \tag{3.10}$$

由(3.7)可得

$$q_{k+1}=1+\sum_{i=0}^{k}\eta^{i+1}<\sum_{i=0}^{+\infty}\eta^{i}=1/\left(1-\eta\right). \tag{3.11}$$

由(3.10)和(3.11)可得

$$\sum_{k=0}^{+\infty}\left|\theta\left(x_k\right)\right|/q_{k+1}\ge\sum_{k\in K,\left\|x_k-\hat{x}\right\|\le\delta}\varepsilon_0/q_{k+1}>\sum_{k\in K,\left\|x_k-\hat{x}\right\|\le\delta}\varepsilon_0\left(1-\eta\right)=+\infty.$$

这与(3.9)矛盾，从而假设不成立，即 $\theta(\hat{x})=0$。由引理 1.1 和引理 2.3 可知，$\hat{x}$ 是问题(1.1)的 Paroto 最优解，也即，算法 2.1 产生的序列 $\{x_k\}$ 的任一极限点都是问题(1.1)的 Pareto 最优解，故定理 3.1 成立。

4. 局部超线性收敛速度

我们在本小节给出算法 2.1 具有局部超线性收敛速度的证明。在证明之前先给出下面的引理及假设。

引理 4.1 [2]设 $x\in U\subset\mathbb{R}^n$，$a,b\in\mathbb{R}$，且 $0<a\le b$。若 $aI\le\nabla^2F_j\left(x\right)\le bI\ \left(j=1,2,\cdots,m\right)$，则

$$(a/2)\|s(x)\|^2 \le |\theta(x)| \le (b/2)\|s(x)\|^2 .$$

引理 4.2 [2]设 $x \in U \subset \mathbb{R}^n$， $a>0$。若 $aI \le \nabla^2 F_j(x)\,(j=1,2,\cdots,m)$，则

$$|\theta(x)| \le (1/2a)\left\|\sum_{j=1}^{m}\lambda_j \nabla F_j(x)\right\|^2 .$$

引理 4.3 [2]设 $\hat{x} \in U$，定义 $\hat{x}_+ = \hat{x} + s(\hat{x})$。若 $a,r,\delta,\varepsilon>0$，且满足

1) $\forall x \in B[\hat{x},r]$，有 $aI \le \nabla^2 F_j(x)\,(j=1,2,\cdots,m)$；

2) $\forall x,y \in B[\hat{x},r]$，$\|x-y\|<\delta$，有 $\|\nabla^2 F_j(x)-\nabla^2 F_j(y)\|<\varepsilon\,(j=1,2,\cdots,m)$；

3) $B[\hat{x},r] \subset U$；

4) $\|s(\hat{x})\| \le \min\{\delta,r\}$；

则 $\|s(\hat{x}_+)\| \le (\varepsilon/a)\|s(\hat{x})\|$。

假设 4.1 $\forall \varepsilon>0$，$\exists \delta>0$，$\forall x,y \in V$，$\|x-y\|<\delta$，有 $\|\nabla^2 F_j(x)-\nabla^2 F_j(y)\| \le \varepsilon$。

假设 4.2 $\forall x \in V$，有 $aI \le \nabla^2 F_j(x) \le bI$。

假设 4.3 $B[x_0,r] \subset V$。

假设 4.4 $\varepsilon/a \le 1-\sigma$。

假设 4.5 $\|s(x_0)\| \le \min\{\delta, r(1-\varepsilon/a)\}$。

在上面的假设中，$a,b,r,\delta,\varepsilon>0$，$V \subset U$，$j=1,2,\cdots,m$。

下面给出算法 2.1 的局部超线性收敛定理及证明。

定理 4.1 设 $\{x_k\}$ 是算法 2.1 产生序列，若假设 4.1~4.5 成立，则序列 $\{x_k\}$ 超线性收敛于问题(1.1)的局部 Pareto 最优解。

证 首先证明当 k 充分大时，$\alpha_k=1$。

将 $F_j(x_k+s(x_k))\,(j=1,2,\cdots,m)$ 在 x_k 处二阶泰勒展开，得到

$$F_j(x_k+s(x_k)) = F_j(x_k)+\nabla F_j(x_k)^{\mathrm{T}} s(x_k)+(1/2)s(x_k)^{\mathrm{T}}\nabla^2 F_j(x_k)s(x_k)+o\left(\|s(x_k)\|^2\right).$$

由引理 2.1 及 $\theta(x_k)$ 的定义，上式可写为

$$\begin{aligned} F_j(x_k+s(x_k)) &\le C_j^k+\theta(x_k)+o\left(\|s(x_k)\|^2\right) \\ &= C_j^k+\sigma\theta(x_k)+(1-\sigma)\theta(x_k)+(\varepsilon/2)\|s(x_k)\|^2 . \end{aligned}$$

又由引理 4.1，上式可写为

$$\begin{aligned} F_j(x_k+s(x_k)) &\le C_j^k+\sigma\theta(x_k)-(1-\sigma)(a/2)\|s(x_k)\|^2+(\varepsilon/2)\|s(x_k)\|^2 \\ &= C_j^k+\sigma\theta(x_k)+\left[\varepsilon-a(1-\sigma)\right]\|s(x_k)\|^2/2. \end{aligned} \tag{4.1}$$

由假设 4.4 可得 $\varepsilon-a(1-\sigma)\le 0$，

则(4.1)可写为

$$F_j(x_k+s(x_k)) \le C_j^k+\sigma\theta(x_k)\ \ (j=1,2,\cdots,m)$$

上式说明，当 k 充分大时，$\alpha_k=1$。

其次，证明算法 2.1 产生的序列 $\{x_k\}$ 收敛于问题(1.1)的局部 Pareto 最优解。

由第一部分的证明，当 k 充分大时，$\alpha_k=1$。即存在充分大的 k_0，$\forall k>k_0$，有

$$x_{k+1}=x_k+\alpha_k s(x_k)=x_k+s(x_k).$$

下面用归纳法证明：

1) $$\|x_k-x_0\|\le\|s(x_0)\|\left[1-(\varepsilon/a)^k\right]\Big/(1-\varepsilon/a) \tag{4.2}$$

2) $$\|s(x_k)\|\le(\varepsilon/a)^k\|s(x_0)\| \tag{4.3}$$

当 $k=0$ 时，经计算可知(4.2)，(4.3)都成立。

假设当 $k=n$ 时，(4.2)，(4.3)成立，即 $\|x_n-x_0\|\le\|s(x_0)\|\left[1-(\varepsilon/a)^n\right]\Big/(1-\varepsilon/a)$，$\|s(x_n)\|\le(\varepsilon/a)^n\|s(x_0)\|$。

对(4.2)，当 $k=n+1$ 时：

$$\begin{aligned}\|x_{n+1}-x_0\|&=\|x_n+s(x_n)-x_0\|\\&\le\|x_n-x_0\|+\|s(x_n)\|\\&=\|s(x_0)\|\left[1-(\varepsilon/a)^{kn}\right]\Big/(1-\varepsilon/a)+\|s(x_0)\|(\varepsilon/a)^n\\&=\|s(x_0)\|\left[1-(\varepsilon/a)^{kn+1}\right]\Big/(1-\varepsilon/a).\end{aligned}$$

从而(4.2)对 $\forall k\in\mathbb{N}$ 都成立。

由(4.2)及假设 4.5 可得 $x_k\in B[x_0,r]$。又由假设 4.1~4.5 及引理 4.3，得到

$$\|s(x_{k+1})\|\le(\varepsilon/a)\|s(x_k)\|. \tag{4.4}$$

对(4.3)，结合(4.4)，当 $k=n+1$ 时，有

$$\|s(x_{n+1})\|\le(\varepsilon/a)\|s(x_n)\|\le(\varepsilon/a)(\varepsilon/a)^n\|s(x_0)\|=\|s(x_0)\|(\varepsilon/a)^{n+1}.$$

从而(4.3)对 $\forall k\in\mathbb{N}$ 都成立。

由(4.2)，(4.3)可得

$$\sum_{k=0}^{+\infty}\|x_{k+1}-x_k\|=\sum_{k=0}^{+\infty}\|s(x_k)\|\le\|s(x_0)\|\sum_{k=0}^{+\infty}(\varepsilon/a)^k<\infty, \tag{4.5}$$

从而 $\{x_k\}$ 是柯西序列，存在 $x^*\in\mathbb{R}^n$，使得 $\lim\limits_{k\to\infty}x_k=x^*$。

对(4.2)中的不等式两边同时取极限可得

$$\|x^*-x_0\|\le\|s(x_0)\|/(1-\varepsilon/a). \tag{4.6}$$

由(4.5)可知，$\{\|s(x_k)\|\}$ 收敛到 0。由假设 4.2 及引理 4.1 可得 $\lim\limits_{k\to\infty}|\theta(x_k)|=0$。又由引理 2.3，得到 $|\theta(x^*)|=0$，即 x^* 是问题(1.1)的稳定点，由引理 1.1 可知，x^* 是问题(1.1)的局部 Pareto 最优解。

最后，我们证明 $\{x_k\}$ 超线性收敛到 x^*。

定义 $r_k=\|s(x_0)\|(\varepsilon/a)^k/(1-\varepsilon/a)$，$\delta_k=\|s(x_0)\|(\varepsilon/a)^k$，$k=0,1,2,\cdots$。$\forall\gamma>0$，令 $\hat{\varepsilon}=\min\{a\gamma/(1+2\gamma),\varepsilon\}$。$\hat{r}=r_k$，$\hat{\delta}=\delta_k$，$\hat{x}_0=x_k$。

$\forall x\in B[x_k,r_k]$，则 $\|x-x_k\|\le r_k=\|s(x_0)\|(\varepsilon/a)^k/(1-\varepsilon/a)$。

由(4.3)和假设 4.5 可得，$x_k\in B[x_0,r]$。

由三角不等式可得

$$
\begin{aligned}
\|x-x_0\| &\le \|x-x_k\|+\|x_k-x_0\| \\
&\le \|s(x_0)\|\left[(\varepsilon/a)^k+1-(\varepsilon/a)^k\right]/(1-\varepsilon/a) \\
&= \|s(x_0)\|/(1-\varepsilon/a) \\
&\le r(1-\varepsilon/a)/(1-\varepsilon/a)=r
\end{aligned}
$$

即 $x\in B[x_0,r]$。又 $x\in B[x_k,r_k]$ 且 $B[x_0,r]\subset V$，从而 $B[x_k,r_k]\subset B[x_0,r]\subset V$。

下面证明假设 4.1~4.5 对 $\hat{\varepsilon},\hat{r},\hat{\delta},\hat{x}_0$ 也成立。

假设 4.1 的结论对 $\forall x,y\in V,\|x-y\|\le\delta$ 成立，又 $B[x_k,r_k]\subset B[x_0,r]\subset V$，$\hat{\delta}<\delta$，故假设 4.1 对 $\hat{\varepsilon},\hat{\delta}$ 也成立。

假设 4.2 的结论对 $\forall x\in V$ 成立，而 $\hat{x}_0=x_k\in B[x_k,r_k]\subset B[x_0,r]\subset V$，故假设 4.2 对 $\hat{x}_0$ 也成立。

由 $\hat{\varepsilon}$ 的定义可知 $\hat{\varepsilon}\le\varepsilon$，从而 $\hat{\varepsilon}/a\le\varepsilon/a\le 1-\sigma$，故假设 4.3 对 $\hat{\varepsilon}$ 也成立。

已知 $B[x_0,r]\subset V$，又 $B[\hat{x}_0,\hat{r}]=B[x_k,r_k]\subset B[x_0,r]$，即 $B[\hat{x}_0,\hat{r}]\subset V$，故假设 4.4 对 $\hat{x}_0,\hat{r}$ 也成立。

对假设 4.5，由 $\hat{x}_0,\hat{\delta}$ 的定义及(4.3)可得

$$
\|s(\hat{x}_0)\|=\|s(x_k)\|\le\|s(x_0)\|(\varepsilon/a)^k=\delta_k=\hat{\delta}. \tag{4.7}
$$

由 $\hat{r},\hat{\varepsilon}$ 的定义可得

$$
\begin{aligned}
\hat{r}(1-\hat{\varepsilon}/a) &= r_k(1-\hat{\varepsilon}/a) \\
&= \|s(x_0)\|(\varepsilon/a)^k(1-\hat{\varepsilon}/a)/(1-\varepsilon/a) \\
&\ge \|s(x_0)\|(\varepsilon/a)^k(1-\varepsilon/a)/(1-\varepsilon/a) \\
&= \|s(x_0)\|(\varepsilon/a)^k
\end{aligned}
$$

又由(4.3)，有

$$
\|s(\hat{x}_0)\|=\|s(x_k)\|\le\|s(x_0)\|(\varepsilon/a)^k\le\hat{r}\left[1-(\hat{\varepsilon}/a)\right] \tag{4.8}
$$

结合(4.7)，有 $\|s(\hat{x}_0)\|\le\min\left\{\hat{\delta},\hat{r}\left[1-(\hat{\varepsilon}/a)\right]\right\}$，故假设 4.5 对 $\hat{\varepsilon},\hat{r},\hat{\delta},\hat{x}_0$ 成立。

综上，假设 4.1~4.5 对 $\hat{\varepsilon},\hat{r},\hat{\delta},\hat{x}_0$ 成立。

由假设 4.1~4.5 对 $\hat{\varepsilon},\hat{r},\hat{\delta},\hat{x}_0$ 成立可知，(4.6)对 $\hat{x}_0$ 和 $\hat{\varepsilon}$ 也成立，又 $\hat{x}_0=x_k$，则有

$$
\|x^*-x_k\|\le\|s(x_k)\|/(1-\hat{\varepsilon}/a). \tag{4.9}
$$

由(4.4)和(4.9)可得

$$
\|x^*-x_{k+1}\|\le\|s(x_{k+1})\|/(1-\hat{\varepsilon}/a)\le\|s(x_k)\|(\hat{\varepsilon}/a)/(1-\hat{\varepsilon}/a) \tag{4.10}
$$

由(4.10)及三角不等式可得

$$
\begin{aligned}
\|x^*-x_k\| &\ge \|x_{k+1}-x_k\|-\|x^*-x_{k+1}\| \\
&\ge \|s(x_k)\|-\|s(x_k)\|(\hat{\varepsilon}/a)/(1-\hat{\varepsilon}/a) \\
&= \|s(x_k)\|(1-2\hat{\varepsilon}/a)/(1-\hat{\varepsilon}/a).
\end{aligned} \tag{4.11}
$$

由 $\hat{\varepsilon}$ 的定义，(4.10)和(4.11)可得

$$
\|x^*-x_{k+1}\|\le\|x^*-x_k\|(\hat{\varepsilon}/a)/(1-2\hat{\varepsilon}/a)\le\gamma\|x^*-x_k\|
$$

又 γ 是任意一个大于 0 的常数，故 $\{x_k\}$ 超线性收敛到 x^*。

5. 数值实验

本小节给出一些数值结果，以表明所提出的非单调线性加权牛顿法对求解多目标优化问题的有效性。实验使用 MATLAB2016A 软件，使用 MATLAB 中的 FMINCON 函数求解子问题从而获取搜索方向。在终止准则为$\left|\theta\left(x_k\right)\right|<10^{-3}$或最大迭代次数为 500 的条件下，将本文中的算法 2.1 与文献[3]中的算法(为方便起见，将文献[3]的算法记为算法 2.2)作比较，在同样的参数设置下(具体的参数设置为：$\sigma=0.55$，$\mu=0.6$，$\rho=0.2$，$\eta=0.5$)，比较两个算法求解同一个多目标优化问题的迭代次数及时间。算法 2.1 与算法 2.2 的数值实验结果对比情况如表 1 所示。

Table 1. Comparison of numerical experimental results between Algorithm 2.1 and Algorithm 2.2
表 1. 算法 2.1 与算法 2.2 数值实验结果的对比情况

算例名称	初始点 x_0^{T}	算法 2.1		算法 2.2	
		迭代次数 n	CPU 时间 t	迭代次数 n	CPU 时间 t
DGO1 [11]	0	4	0.0310	13	0.0870
	π/6	5	0.0421	15	0.0986
	π/9	8	0.0930	18	0.1142
MHHM1 [12]	0	4	0.0511	8	0.0751
	0.3	4	0.0410	11	0.0776
	0.5	3	0.0317	7	0.0509
SSFYY2 [13]	0	3	0.0311	202	1.3550
	−1	4	0.0481	212	1.3723
	−0.25	4	0.0403	215	1.3910
BK1 [14]	(0, 2)	5	0.0368	23	0.2134
	(0, −1)	67	0.4232	500	4.9268
	(−1, 2)	5	0.0485	197	1.9371
LRS1 [15]	(1, 2)	5	0.0362	97	0.6231
	(9, 5)	7	0.0476	420	2.5258
	(−2, 4)	6	0.0422	500	2.9715
MHHM2 [12]	(0.4, 0.1)	4	0.0322	12	0.0836
	(1, 1)	3	0.0267	8	0.0546
	(0.5, 0.2)	4	0.0534	11	0.0797
MOP5 [16]	(1, 2)	52	0.5662	276	3.6353
	(π/6, π/6)	6	0.0617	500	4.6164
	(1, 1.5)	89	0.8493	445	3.8441
PNR [17]	(1, 0.7)	5	0.0828	17	0.1892
	(1.2, 1)	7	0.1264	23	0.2575
	(−1, 1)	7	0.0852	74	0.6424

续表

SP1 [18]	(2, 1)	5	0.0487	12	0.1026
	(−1, 1)	5	0.0569	37	0.3020
	(−3, 0)	6	0.0570	60	0.4823
T5 [19]	(1, 1)	5	0.0986	12	0.1172
	(2, 3)	10	0.1125	47	0.419
	(9, 4)	11	0.1616	57	0.5487
VFM1 [20]	(0.2, 0)	3	0.0297	283	1.8009
	(1, 0.8)	97	0.6974	500	3.2561
	(1, 1)	100	0.6881	500	3.4263
AP4 [21]	(1, 1, 2)	50	0.7091	500	6.3972
	(1, 0.5, 2)	89	1.2459	500	5.5488
	(0.5, 1, 2)	47	0.644	500	6.2969
T8 [19]	(1, 0, 1)	5	0.0921	16	0.2274
	(1, 0, 2)	7	0.1301	374	5.4864
	(2, 0, 5)	10	0.1596	25	0.3842
TRIDIA [10]	(0.1, −0.2, 0.4)	4	0.0770	181	2.8785
	(−0.1, 0.2, 0.5)	66	0.9319	378	5.3868
	(0, 0.1, 0.2)	70	1.1856	355	5.5894
ROSEN BROCK [10]	(0.97, 0.94, 0.99, 0.98)	24	0.6158	258	6.0115
	(1.25, 1.3, 1.35, 1.4)	9	0.2246	25	0.5857
	(2, 1.98, 1.96, 2)	14	0.3259	223	4.8018
SD [22]	$(1, \sqrt{2}, \sqrt{2}, \sqrt{2})$	11	0.2120	500	7.1499
	$(1, \sqrt{2}, \sqrt{2}, 1)$	11	0.1964	500	6.2174
	(1, 1.45, 1.45, 1)	15	0.2960	500	6.1492
Shifted-TRIDIA [10]	(−1, 1, −1, 1)	7	0.1571	326	6.3939
	(0.1, 0.3, 0.2, 0.1)	4	0.0979	78	1.7083
	(−0.4, −0.3, 0.5, 0.5)	96	2.0398	500	9.8446
TOINT [10]	(0, −1, −1, 4)	6	0.0958	25	0.3197
	(−1, −1, 3, 2)	5	0.0841	22	0.2644
	(3, 3, 1, 4)	7	0.1149	30	0.3474
DD1 [23]	(0, 0, 1, 1, 1)	5	0.0416	18	0.2402
	(1, 2, 3, 0, 0)	6	0.0477	85	0.6142
	(0, 2, 2, −1, −1)	5	0.0482	77	0.5736
JOS1 [24]	(0, −1, 1, 0, 0)	4	0.0681	94	2.6891
	(−0.3, 0.2, 0.1, 0.4, 0.5)	67	0.4709	160	1.1145
	(0.3, 0.3, −0.1, 0.8, 0.9)	4	0.0337	85	0.6184

表 1 列举出了算法 2.1 和算法 2.2 关于 20 个测试问题在不同初始点的数值结果。由表 1 可知，虽然算法 2.1 和算法 2.2 都能求解测试问题，但是算法 2.1 的数值表现优于算法 2.2。在相同的条件下，算法 2.1 求解多目标优化测试问题的迭代次数以及 CPU 时间都远小于算法 2.2，算法 2.1 能更高效地求解问题。

6. 结论

本文基于线性加权牛顿算法，结合非单调线性搜索技术，提出了求解无约束多目标优化问题的非单调线性加权牛顿算法(NLWN)。当目标函数二阶连续可微且局部强凸时，证明了由该算法产生的序列局部超线性收敛于多目标优化问题的 Pareto 最优解。最后，通过数值试验来评估算法的有效性，将它与单调线性加权牛顿算法进行比较，数值结果表明 NLWN 算法具有较好的数值表现。

参考文献

[1] 林锉云, 董加礼. 多目标优化的方法与理论[M]. 长春: 吉林教育出版社, 1992.

[2] Fliege, J., Drummond, L.M.G. and Svaiter, B.F. (2009) Newton's Method for Multiobjective Optimization. *SIAM Journal on Optimization*, **20**, 602-626. https://doi.org/10.1137/08071692x

[3] 江术兰. 多目标优化问题的标量化方法及牛顿算法[D]: [硕士学位论文]. 重庆: 重庆师范大学, 2022.

[4] Fliege, J. and Svaiter, B.F. (2000) Steepest Descent Methods for Multicriteria Optimization. *Mathematical Methods of Operations Research* (*ZOR*), **51**, 479-494. https://doi.org/10.1007/s001860000043

[5] Drummond, L.M.G. and Iusem, A.N. (2004) A Projected Gradient Method for Vector Optimization Problems. *Computational Optimization and Applications*, **28**, 5-29. https://doi.org/10.1023/b:coap.0000018877.86161.8b

[6] Lucambio Pérez, L.R. and Prudente, L.F. (2018) Nonlinear Conjugate Gradient Methods for Vector Optimization. *SIAM Journal on Optimization*, **28**, 2690-2720. https://doi.org/10.1137/17m1126588

[7] Povalej, Ž. (2014) Quasi-Newton's Method for Multiobjective Optimization. *Journal of Computational and Applied Mathematics*, **255**, 765-777. https://doi.org/10.1016/j.cam.2013.06.045

[8] Zhang, H. and Hager, W.W. (2004) A Nonmonotone Line Search Technique and Its Application to Unconstrained Optimization. *SIAM Journal on Optimization*, **14**, 1043-1056. https://doi.org/10.1137/s1052623403428208

[9] Grippo, L., Lampariello, F. and Lucidi, S. (1986) A Nonmonotone Line Search Technique for Newton's Method. *SIAM Journal on Numerical Analysis*, **23**, 707-716. https://doi.org/10.1137/0723046

[10] Mita, K., Fukuda, E.H. and Yamashita, N. (2019) Nonmonotone Line Searches for Unconstrained Multiobjective Optimization Problems. *Journal of Global Optimization*, **75**, 63-90. https://doi.org/10.1007/s10898-019-00802-0

[11] Dumitrescu, D., Grosan, C. and Oltean, M. (2000) A New Evolutionary Approach for Multiobjective Optimization. *Studia Universitatis Babes-Bolyai Informatica*, **45**, 51-68.

[12] Mao, J., Hirasawa, K., Hu, J. and Murata, J. (2001) Genetic Symbiosis Algorithm for Multiobjective Optimization Problems. *Transactions of the Society of Instrument and Control Engineers*, **37**, 893-901. https://doi.org/10.9746/sicetr1965.37.893

[13] Shim, M., Suh, M., Furukawa, T., Yagawa, G. and Yoshimura, S. (2002) Pareto-Based Continuous Evolutionary Algorithms for Multiobjective Optimization. *Engineering Computations*, **19**, 22-48. https://doi.org/10.1108/02644400210413649

[14] Custódio, A.L., Madeira, J.F.A., Vaz, A.I.F. and Vicente, L.N. (2011) Direct Multisearch for Multiobjective Optimization. *SIAM Journal on Optimization*, **21**, 1109-1140. https://doi.org/10.1137/10079731x

[15] Laumanns, M., Rudolph, G. and Schwefel, H.P. (1998) A Spatial Predator-Prey Approach to Multi-Objective Optimization: A Preliminary Study. *Lecture Notes in Computer Science*, **5**, 241-249.

[16] Huband, S., Hingston, P., Barone, L. and While, L. (2006) A Review of Multiobjective Test Problems and a Scalable Test Problem Toolkit. *IEEE Transactions on Evolutionary Computation*, **10**, 477-506. https://doi.org/10.1109/tevc.2005.861417

[17] Preuss, M., Naujoks, B. and Rudolph, G. (2006) Pareto Set and EMOA Behavior for Simple Multimodal Multiobjective Functions. *International Conference on Parallel Problem Solving from Nature*, Reykjavik, 9-13 September 2006, 513-522.

[18] Sefrioui, M. and Perlaux, J. (2000) Nash Genetic Algorithms: Examples and Applications. *Proceedings of the* 2000

Congress on Evolutionary Computation, La Jolla, 16-19 July 2000, 509-516.

[19] Thomann, J. and Eichfelder, G. (2019) Numerical Results for the Multiobjective Trust Region Algorithm MHT. *Data in Brief*, **25**, Article ID: 104103. https://doi.org/10.1016/j.dib.2019.104103

[20] Vlennet, R., Fonteix, C. and Marc, I. (1996) Multicriteria Optimization Using a Genetic Algorithm for Determining a Pareto Set. *International Journal of Systems Science*, **27**, 255-260. https://doi.org/10.1080/00207729608929211

[21] Ansary, M.A.T. and Panda, G. (2014) A Modified Quasi-Newton Method for Vector Optimization Problem. *Optimization*, **64**, 2289-2306. https://doi.org/10.1080/02331934.2014.947500

[22] Stadler, W. and Dauer, J. (1993) Multicriteria Optimization in Engineering: A Tutorial and Survey. *Progress in Astronautics and Aeronautics*, **150**, 209-249.

[23] Das, I. and Dennis, J.E. (1998) Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems. *SIAM Journal on Optimization*, **8**, 631-657. https://doi.org/10.1137/s1052623496307510

[24] Jin, Y., Olhofer, M. and Sendho, B. (2001) Dynamic Weighted Aggregation for Evolutionary Multi-Objective Optimization: Why Does It Work and How. *Proceedings of the* 3*rd Annual Conference on Genetic and Evolutionary Computation*, San Francisco, 7-11 July 2001, 1042-1049.