

# 解决向量优化问题的一种非单调投影梯度算法

刘 雪, 周犁文

西南石油大学理学院, 四川 成都

收稿日期: 2023年4月28日; 录用日期: 2023年5月21日; 发布日期: 2023年5月29日

## 摘 要

本文引入一种新的算法: 极大型投影梯度算法, 它是将解决带约束向量优化问题的投影梯度算法和极大型非单调线搜索技术相结合的一种算法。下降方向通过由解决单目标规划的投影梯度算法推广到向量优化的投影梯度算法来得到, 而在步长选择上采用经典的单调线搜索技术容易陷入局部收敛的困境, 与非单调技术结合以后, 可以摆脱这一困境。在合适的条件下, 证明了算法的全局收敛和线性收敛性。

## 关键词

向量优化, 投影梯度算法, 非单调线搜索技术, 全局收敛, 线性收敛

# A Nonmonotone Projected Gradient Algorithm for Solving Vector Optimization Problems

Xue Liu, Liwen Zhou

School of Science, Southwest Petroleum University, Chengdu Sichuan

Received: Apr. 28<sup>th</sup>, 2023; accepted: May 21<sup>st</sup>, 2023; published: May 29<sup>th</sup>, 2023

## Abstract

This paper introduces a new algorithm: the max-type projected gradient algorithm, which combines the projected gradient algorithm for solving constrained vector optimization problems with the max-type nonmonotone line search technology. The Descent direction is obtained by extending the projection gradient algorithm for solving single objective programming to the projection gradient algorithm for vector optimization, while the classical monotone line search technology in step size selection is easy to fall into the dilemma of local convergence, which can be overcome by combining with nonmonotone technology. Under milder conditions, the global and linear conver-

gences of the algorithm were demonstrated, and numerical experiments were conducted to verify its effectiveness.

## Keywords

Vector Optimization, Projected Gradient Algorithm, Nonmonotone Line Search Technology, Global Convergence, Linear Convergence

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

本文考虑如下约束向量优化问题(constraint vector optimization, 简记 CVOP):

$$\begin{cases} \min F(x), \\ \text{s.t. } x \in C, \end{cases}$$

$C$  是  $R^n$  中的非空闭凸子集,  $F: R^n \rightarrow R^m$  是连续可微的, 且  $F(x) = (f_1(x), \dots, f_m(x))$ ,  $f_i(x): R^n \rightarrow R$ ,  $\forall i \in \{1, \dots, m\}$ ,  $F(x)$  是连续可微的意味着它的每一个分量  $f_i(x)$  都是连续的。这种类型的问题在金融、环境分析和管理学等领域都有所涉及, 见文献[1] [2] [3]。

在求解向量优化问题的众多方法中, 主要分成两类: 标量化方法和下降方法。

标量化方法的基础思想是通过特定的参数将向量值问题转化为标量值问题。其中, 著名的加权求和方法是通过对各个目标函数进行线性组合来得到一个单目标优化, 新问题的最优解亦可看成是原始问题的最优解, 如[4]。

向量优化的下降方法不需要提前知晓任何的参数信息, 而且关于收敛性有理论的证明。它们是从单目标非线性规划的算法中拓展而来的。当向量优化无约束的时候, 即如下形式:

$$\min F(x)$$

有下面几种方法被提出:

Jörg Fliege 和 Benar Fux Svaiter [5]提出解决向量优化的最速下降法, 下降方向定义如下:

$$v(x) := \arg \min \left\{ \nabla f(v) + \frac{1}{2} \|v\|^2 \right\}$$

在解决约束向量优化问题时, 目标函数是 Lipschitz 连续的条件下, 证明了该算法收敛到一个满足某些一阶必要条件的点。

Fliege, J.等人[6]提出解决向量优化的牛顿法, 下降方向定义如下:

$$v(x) := \arg \min \max_{s \in R^n, j=1, \dots, m} \left\{ \nabla f_j(x_k)^T s + \frac{1}{2} s^T \nabla^2 f_j(x_k) s \right\}.$$

文中也证明了在目标函数满足二阶连续可微的条件下, 该算法生成的序列点是局部超线性收敛到最优点的。

对于带约束的向量优化问题, 2004年, Graña Drummond 和 Iusem [7]提出了解决向量优化的投影梯

度算法(projected gradient method, 简记为 PGM), 下降方向定义如下:

$$v(x) := \arg \min_{v \in C_x} \left\{ \|v\|^2 / 2 + \beta \max_{y \in G} (y^T JF(x)v) \right\}.$$

其中,  $JF(x)$  表示的是  $F$  在  $x$  处的雅可比矩阵,  $C_x = C - x = \{z - x : z \in C\}$ ,  $G = \{y \in K^* : \|y\| = 1\}$ 。上面几种方法可以通过选择适合的初始点来找到帕累托最优点, 并且它们都是迭代型的算法, 每次迭代的时候选择的下降方向都是通过解决一个目标函数是强凸的单目标优化来得到的, 这样保证了下降方向是唯一的。特别地, Graña Drummond 和 Iusem 也就投影梯度算法的两个方面给出了收敛性分析。一个是加上平方和步长和凸假设, 他们证明了算法收敛到一个弱帕累托解。另一个则是给定目标函数是凸函数的假设, 但是没有额外加步长, 此时, 他们证明了聚点是驻点。作为一种经典而有效的方法, 投影梯度算法被广泛研究和应用在各种多目标规划中。关于固定步长的投影梯度算法, 也有一些重要的工作。针对非可微凸向量优化问题, Bello Cruz 等人[8]提出投影次梯度算法, 并且证明了由算法生成的序列的聚点是弱帕累托解。最近, 通过使用量化的方法, Brito 等人[9]和 Bento 等人[10]使用带权重的次梯度算法来解决非可微凸向量优化问题, 并且获得了收敛到帕累托解的结果。虽然上述方法没有使用单调线搜索技术, 但是它们依然有着很强的收敛结果。

在迭代过程中, 改变步长能否对算法进行改进是一个值得研究的问题。本文受此启发, 引入非单调步长来得到非单调投影梯度算法, 并在合适的条件下证明算法的有效性。

## 2. 相关定义和引理

**定义 2.1** 集合  $C$  被称为是锥, 如果对任意的  $\alpha \in (0, +\infty)$  和  $x \in C$ , 都有  $\alpha x \in C$ 。进一步, 集合  $C$  被称为是尖锥, 如果它是锥, 并且满足  $-C \cap C = \{0\}$ 。

**定义 2.2** 令集合  $K$  是非空闭尖凸锥, 偏序 “ $\prec (\preceq)$ ” 的定义如下:

$$z \prec (\preceq) z' \Leftrightarrow z' - z \in K (\text{int } K).$$

**定义 2.3** 集合  $K^*$  称为集合  $K$  的极锥, 如果它满足:

$$K^* = \{y \in R^n : \langle y, x \rangle \geq 0, \forall x \in K\}.$$

**定义 2.4** 设集合  $C$  是一个非空的集合。

1)  $x^* \in C$  被称为是帕累托最优, 如果不存在  $y \in C$  使得

$$F(y) \preceq F(x^*) \text{ 和 } F(y) = F(x^*).$$

2)  $x^* \in C$  被称为是弱帕累托最优, 如果不存在  $y \in C$  使得

$$F(y) \prec F(x^*).$$

3)  $x^* \in C$  被称为是  $F$  的驻点, 如果满足

$$JF(x^*)(C - x^*) \cap (-\text{int } K) = \emptyset,$$

其中,  $JF(x^*)(C - x^*) := \{JF(x^*)(x - x^*) : x \in C\}$ 。

很多情况下, (弱)帕累托解很难被证明, 许多学者会考虑证明生成的序列收敛到驻点来取代证明收敛到(弱)帕累托解。根据驻点的定义, 我们知道如果  $\bar{x}$  不是驻点, 那么, 存在  $x \in C$  使得

$JF(\bar{x})(x - \bar{x}) \in -\text{int } K$ , 这等价于  $JF(\bar{x})(x - \bar{x}) \prec 0$ , 取  $K = R_+^m$  时, 这进一步等价于如下表达式:

$$(\nabla f_i(\bar{x}))^T (x - \bar{x}) < 0, \forall i \in \{1, \dots, m\}.$$

又  $F$  是连续可微的函数, 故

$$\lim_{\gamma \rightarrow 0^+} (f_i(\bar{x} + \gamma(x - \bar{x})) - f_i(\bar{x})) / \gamma = (\nabla f_i(\bar{x}))^T (x - \bar{x}) < 0,$$

此式说明  $x - \bar{x}$  在  $\bar{x}$  处是一个下降方向, 即不是驻点就存在下降方向, 反之, 不存在下降方向的点就是驻点。

**定义 2.5** 映射  $F$  在  $H$  上是 L-Lipschitz 连续映射, 当且仅当存在常数  $L > 0$  使得映射  $F$  满足

$$\|F(x) - F(y)\| \leq L\|x - y\|, \forall x, y \in H.$$

如果  $L = 1$ , 则称  $F$  是非扩张的; 如果  $L \in (0, 1)$ , 则称  $F$  是压缩的。

**定义 2.6** (收敛速度) 设序列  $\{x_k\}$  是  $R^n$  中的序列, 并且满足依范数收敛到  $x^* \in R^n$ , 称序列  $\{x_k\}$  具有 Q-线性收敛率, 如果存在  $\theta \in (0, 1)$  使得当下标  $k$  充分大时, 有

$$\|x_{k+1} - x^*\| \leq \theta \|x_k - x^*\|.$$

下面给出一些引理用于后面算法的收敛性证明。

**引理 2.7** [11] 令  $C$  是一个锥, 则对于任意的  $x, y \in C$ , 有  $x + y \in C$ 。

**引理 2.8** [12] 对于问题(CVOP), 如果  $x \in R^n$  是帕累托最优点, 则  $x$  也是驻点。

### 3. 算法及其收敛性

#### 3.1. 算法描述

**步骤 1** 初始化: 给定初始点  $x_0 \in C$ , 参数  $\sigma \in (0, 1)$ ,  $\rho > 1$ ,  $\mu \in (0, 1]$ ,  $\varepsilon > 0$  以及非负整数  $M$  和  $N$ , 令  $k = 0$  和  $m(k) = 0$ ;

**步骤 2** 确定寻找方向: 计算

$$v(x_k) := \arg \min_{v \in C_{x_k}} \left\{ \frac{\|v\|^2}{2} + \beta \max_{y \in G} (y^T JF(x_k)v) \right\}; \quad (3-1)$$

**步骤 3** 停止准则: 如果迭代次数  $k > N$  或者  $\|v(x_k)\| < \varepsilon$ , 停止; 否则, 继续步骤 4;

**步骤 4** 令  $m(k) = \min\{m(k-1) + 1, M\}$  和  $C_k$ :

$$C_{i,k} := \max_{0 \leq j \leq m(k)} f_i(x_{k-j}), i \in \{1, \dots, m\}; \quad (3-2)$$

**步骤 5** 确定步长: 取  $\gamma_k = \mu \rho^{-h_k}$ , 其中  $h_k$  是使得下面不等式成立的最小非负整数:

$$F(x_k + \gamma_k v(x_k)) \preceq C_k + \sigma \gamma_k JF(x_k)v(x_k); \quad (3-3)$$

**步骤 6** 令  $x_{k+1} = x_k + \gamma_k v(x_k)$ ,  $k = k + 1$ , 并返回到步骤 2。

**注 3.1:** 当  $\{x_k\}$  是由算法 1 生成的序列时, 由(3-2)可以推导出  $F(x_k) \preceq C_k$ 。

首先, 定义函数  $\varphi: R^m \rightarrow R$  如下:

$$\varphi(x) := \max_{y \in G} (y^T x). \quad (3-4)$$

结合极锥  $K^*$  和集合  $G$  的定义, 集合  $-K$  和  $-\text{int } K$  可以转化下面两个形式:

$$-K = \{y \in R^m : \varphi(y) \leq 0\}, \quad (3-5)$$

和

$$-\text{int } K = \{y \in \mathbb{R}^m : \varphi(y) < 0\}. \quad (3-6)$$

下面一个引理说明函数  $\varphi$  的一些性质。

**引理 3.1** ([7] Proposition 2)  $\varphi$  的定义如(3-4)所示, 下面几个性质成立:

- 1) 函数  $\varphi$  具有一阶正齐次性, 即  $\varphi(kx) = k\varphi(x), \forall x \in \mathbb{R}^m, k \geq 0$ ;
- 2) 对于所有的  $x, z \in \mathbb{R}^m$ ,  $\varphi(x+z) \leq \varphi(x) + \varphi(z)$  和  $\varphi(x) - \varphi(z) \leq \varphi(x-z)$  都成立;
- 3) 对于所有的  $x, z \in \mathbb{R}^m$ , 若  $x \prec z$  (或者  $x \preceq z$ ), 则  $\varphi(x) < \varphi(z)$  (或者  $\varphi(x) \leq \varphi(z)$ );
- 4) 函数  $\varphi: \mathbb{R}^m \rightarrow \mathbb{R}$  是 Lipschitz 连续的, 且 Lipschitz 常数  $L=1$ 。

其次, 将(3-1)中目标函数记为  $\theta: C \rightarrow \mathbb{R}$ :

$$\theta(x) := \frac{\|v(x)\|^2}{2} + \beta \max_{y \in G} (y^T JF(x)v(x)). \quad (3-7)$$

其中,  $v(x) \in C_x = \{z - x \mid z \in C\}$ , 且  $v(x)$  是使得  $\theta(x)$  取值最小的一个向量。

下面的引理则陈述关于此函数的一些性质, 这些性质对于算法的收敛性证明非常有用。

**引理 3.2** ([7] Proposition 3)  $v$  和  $\theta$  的定义分别如(3-1)和(3-7)所示, 下面几个结论成立:

- (1)  $\theta(x) \leq 0, \forall x \in C$ ;
- (2) 当  $x \in C$  时, 下面几个条件等价:
  - (a)  $x$  不是  $F$  的驻点;
  - (b)  $\theta(x) < 0$ ;
  - (c)  $v(x) \neq 0$ .

**注 3.2:** (1) 由命题 3.2(1)和  $\theta$  的定义知  $\varphi(JF(x)v(x)) \leq 0$ , 由  $\varphi$  的定义知  $JF(x)v(x) \preceq 0$ , 从而知  $v(x)$  是一个下降方向;

(2) 从引理 3.2(2)可以推导出  $x \in C$  是  $F$  的驻点  $\Leftrightarrow \theta(x) = 0 \Leftrightarrow v(x) = 0$ .

在步骤 3 的停止准则中, 其中一个要求  $\|v(x_k)\| < \varepsilon$ 。一般来说, 参数  $\varepsilon$  是任意小的, 所以从注 3.2(2)可知此条件意味着  $x_k$  是驻点。

**命题 3.3** 对于给定的点  $x \in C$ ,  $v(x)$  定义如(3-1), 则它满足下面两个不等式:

$$\|v(x)\|^2 \leq -2\beta\varphi(JF(x)v(x)), \quad (3-8)$$

$$\|v(x)\| \leq 2\beta\|JF(x)\|, \forall i \in \{1, \dots, m\}. \quad (3-9)$$

证明: 第一个不等式根据  $\theta(x)$  的定义和命题 3.2(1)可以得到。再利用  $\varphi$  的定义和柯西-施瓦茨不等式, 有

$$-2\beta\varphi(JF(x)v(x)) \leq 2\beta\|JF(x)\|\|v(x)\|.$$

将此与第一个不等式相结合即可得到第二个不等式。

下面的引理给出了函数  $\theta$  的连续性性质。

**引理 3.4** ([7], Proposition 4) 函数  $\theta$  在  $C$  中是连续的。

下面的引理说明极大型线搜索策略和投影梯度算法结合是可行的。

**引理 3.5** (1) 算法 1 是有意义的;

(2) 令  $\{x_k\}$  是算法 1 生成的序列, 则  $\{x_k\} \subseteq C$ 。

证明: (1) 显然函数  $\theta(x)$  是一个强凸函数, 则下降方向  $v(x)$  可以得到且唯一。由注 3.1 可知(3-3)是有意义的。故结论成立。

(2) 下面用数学归纳法来证明结论。

当下标  $k=0$  时, 由算法 1 的初始化步骤可知  $x_0 \in C$ , 结论成立;

假设当下标  $k=k'$  时,  $x_{k'} \in C$ , 由下降方向  $v(x_{k'})$  的定义知,  $v(x_{k'}) + x_{k'} \in C$ , 而

$$x_{k'+1} = x_{k'} + \gamma_{k'} v(x_{k'}),$$

以及  $\gamma_{k'} \in (0,1]$  和  $C$  是一个闭凸集合, 所以

$$x_{k'+1} = \gamma_{k'} (x_{k'} + v(x_{k'})) + (1 - \gamma_{k'}) x_{k'} \in C.$$

结论成立。

### 3.2. 全局收敛

**命题 3.6** 假设  $F$  连续可微有下界, 序列  $\{x_k\}$  是由算法 1 生成的, 则序列  $\{C_k\}$  是单调非增的, 并且存在极限。

证明: 因为  $v(x_k)$  是一个下降方向, 故  $JF(x_k)v(x_k) \leq 0$  对所有的  $k \in N$  均成立, 将其带入到(3-3)中, 有  $F(x_{k+1}) \leq C_k$ 。

当序列  $\{x_k\}$  是由算法 1 生成的序列时, 根据(3-2)和  $m(k+1) \leq m(k)+1$ , 对于向量的每一个分量  $i$  和迭代次数  $k$ , 有

$$\begin{aligned} C_{i,k+1} &= \max_{0 \leq j \leq m(k+1)} f_i(x_{k+1-j}) \\ &\leq \max_{0 \leq j \leq m(k)+1} f_i(x_{k+1-j}) \\ &= \max_{0 \leq j \leq m(k)+1} \{C_{i,k}, f_i(x_{k+1})\} = C_{i,k}. \end{aligned}$$

这也就说明了序列  $\{C_k\}$  是单调非增的。同时考虑到  $C_k \geq F(x_k)$  和  $F$  是有下界的, 根据单调有界定理知序列  $\{C_k\}$  存在极限。证毕。

**引理 3.7** 假设  $F$  有下界且连续可微, 序列  $\{x_k\}$  是由算法 1 生成的, 那么,

$$\lim_{k \rightarrow \infty} \gamma_k \varphi(JF(x_k)v(x_k)) = 0.$$

证明: 对于任意迭代次数  $k$  和  $i \in \{1, \dots, m\}$ , 令  $l_i(k)$  为使得下面两个条件成立的整数:

$$k - m(k) \leq l_i(k) \leq k, \quad (3-10)$$

$$C_{i,k} = f_i(x_{l_i(k)}). \quad (3-11)$$

利用(3-3)和  $x_k$  的迭代公式, 有

$$\begin{aligned} C_{i,k} &= f_i(x_{l_i(k)}) \\ &= f_i(x_{l_i(k)-1} + \gamma_{l_i(k)-1} v(x_{l_i(k)-1})) \\ &\leq C_{i,l_i(k)-1} + \sigma \gamma_{l_i(k)-1} \left( \nabla f_i(x_{l_i(k)-1}) \right)^T v(x_{l_i(k)-1}). \end{aligned} \quad (3-12)$$

对于条件(3-10), 可得  $k - m(k) - 1 \leq l_i(k) - 1 \leq k - 1 < k$ , 当  $k$  足够大时  $m(k)$  取常数, 再根据命题 3.6, 有

$$\lim_{k \rightarrow \infty} (C_{i,k} - C_{i,l_i(k)-1}) = 0.$$

又因为  $\gamma_{l_i(k)-1} > 0$  和  $\nabla f_i(x_{l_i(k)-1})v(x_{l_i(k)-1}) \leq 0$ , 结合(3-12)可得

$$\lim_{k \rightarrow \infty} \gamma_{l_i(k)-1} \left( \nabla f_i(x_{l_i(k)-1}) \right)^T v(x_{l_i(k)-1}) = 0. \quad (3-13)$$

因为  $i \in \{1, \dots, m\}$  是任意取的, 进一步可得

$$\lim_{k \rightarrow \infty} \gamma_{l_i(k)-1} JF(x_{l_i(k)-1})v(x_{l_i(k)-1}) = 0.$$

从而

$$\lim_{k \rightarrow \infty} \gamma_{l_i(k)-1} \varphi(JF(x_{l_i(k)-1})v(x_{l_i(k)-1})) = 0. \quad (3-14)$$

定义

$$\hat{l}_i(k) := l_i(k + M + 2).$$

显然  $\{\hat{l}_i(k)\}_k \subseteq \{l_i(k)\}_k$ , 在证明结论之前, 先证明下面两个结果: 对任意给定的  $j \geq 1$ , 有

$$\lim_{k \rightarrow \infty} \gamma_{\hat{l}_i(k)-j} \varphi(JF(x_{\hat{l}_i(k)-j})v(x_{\hat{l}_i(k)-j})) = 0, \quad (3-15)$$

$$\lim_{k \rightarrow \infty} C_{i,k} = \lim_{k \rightarrow \infty} f_i(x_{\hat{l}_i(k)-j}). \quad (3-16)$$

接下来用数学归纳法来证明这两个结果。

当  $j=1$  时, 结合(3-14)与  $\{\hat{l}_i(k)\}_k \subseteq \{l_i(k)\}_k$ , 可得(3-15)。考虑到(3-9)和  $\{x_k\}$  的迭代公式, 有

$$\|x_{\hat{l}_i(k)} - x_{\hat{l}_i(k)-1}\| = \|\gamma_{\hat{l}_i(k)-1} v(x_{\hat{l}_i(k)-1})\| \leq \left\| -2\beta \gamma_{\hat{l}_i(k)-1}^2 \varphi(JF(x_{\hat{l}_i(k)-1})v(x_{\hat{l}_i(k)-1})) \right\|^{1/2}. \quad (3-17)$$

因为  $\gamma_{\hat{l}_i(k)-1} \in (0, \mu]$ , 在等式两端同时对  $k$  取极限, 并将(3-15)带入可得

$$\lim_{k \rightarrow \infty} \|x_{\hat{l}_i(k)} - x_{\hat{l}_i(k)-1}\| = 0.$$

又  $f_i$  是连续函数, 故

$$\lim_{k \rightarrow \infty} f_i(x_{\hat{l}_i(k)-1}) = \lim_{k \rightarrow \infty} f_i(x_{\hat{l}_i(k)}) = \lim_{x \rightarrow \infty} C_{i,k},$$

后一等式的成立利用了  $C_k$  的定义(3-2)。因此, (3-16)成立。

假设当  $j = \bar{j}$  时, (3-15)和(3-16)成立。利用(3-3), 有

$$f_i(x_{\hat{l}_i(k)-\bar{j}}) \leq C_{i,\hat{l}_i(k)-\bar{j}-1} + \sigma \gamma_{\hat{l}_i(k)-\bar{j}-1} \left( \nabla f_i(x_{\hat{l}_i(k)-\bar{j}-1}) \right)^T v(x_{\hat{l}_i(k)-\bar{j}-1}), \quad (3-18)$$

注意到  $\lim_{k \rightarrow \infty} (\hat{l}_i(k) - \bar{j} - 1) = \infty$  和(3-18)在  $j = \bar{j}$  时成立, 利用命题 3.6,  $\sigma > 0$ ,  $\gamma_{\hat{l}_i(k)-\bar{j}-1} > 0$  和

$\left( \nabla f_i(x_{\hat{l}_i(k)-\bar{j}-1}) \right)^T v(x_{\hat{l}_i(k)-\bar{j}-1}) \leq 0$ , 由(3-18)可以推导出

$$\lim_{k \rightarrow \infty} \gamma_{\hat{l}_i(k)-\bar{j}-1} \left( \nabla f_i(x_{\hat{l}_i(k)-(\bar{j}+1)}) \right)^T v(x_{\hat{l}_i(k)-(\bar{j}+1)}) = 0.$$

由  $i$  的任意性有

$$\lim_{k \rightarrow \infty} \gamma_{\hat{l}_i(k)-\bar{j}-1} JF(x_{\hat{l}_i(k)-(\bar{j}+1)})v(x_{\hat{l}_i(k)-(\bar{j}+1)}) = 0.$$

利用命题 3.1, 上式可以进一步变为

$$\lim_{k \rightarrow \infty} \gamma_{\hat{l}_i(k)-\bar{j}-1} \varphi \left( JF \left( x_{\hat{l}_i(k)-(\bar{j}+1)} \right) v \left( x_{\hat{l}_i(k)-(\bar{j}+1)} \right) \right) = 0, \tag{3-19}$$

则(3-15)成立。

根据命题 3.3 和  $x_k$  的迭代公式, 有

$$\begin{aligned} \left\| x_{\hat{l}_i(k)-\bar{j}} - x_{\hat{l}_i(k)-(\bar{j}+1)} \right\| &= \left\| \gamma_{\hat{l}_i(k)-(\bar{j}+1)} v \left( x_{\hat{l}_i(k)-(\bar{j}+1)} \right) \right\| \\ &\leq \left\| -2\beta \gamma_{\hat{l}_i(k)-(\bar{j}+1)}^2 \varphi \left( JF \left( x_{\hat{l}_i(k)-(\bar{j}+1)} \right) v \left( x_{\hat{l}_i(k)-(\bar{j}+1)} \right) \right) \right\|^{1/2}. \end{aligned}$$

将(3-19)与上式相结合, 可推导出

$$\lim_{k \rightarrow \infty} \left\| x_{\hat{l}_i(k)-\bar{j}} - x_{\hat{l}_i(k)-(\bar{j}+1)} \right\| = 0.$$

又因为  $f_i$  是连续函数, 故

$$\lim_{k \rightarrow \infty} f_i \left( x_{\hat{l}_i(k)-(\bar{j}+1)} \right) = \lim_{k \rightarrow \infty} f_i \left( x_{\hat{l}_i(k)-\bar{j}} \right) = \lim_{x \rightarrow \infty} C_{i,k}.$$

于是, (3-15)和(3-16)对任意给定的  $j \geq 1$  都成立。

对所有的迭代次数  $k > 0$ , 有

$$x_{k+1} = x_{\hat{l}_i(k)} - \sum_{j=1}^{\hat{l}_i(k)-k-1} \gamma_{\hat{l}_i(k)-j} v \left( x_{\hat{l}_i(k)-j} \right). \tag{3-20}$$

注意到  $\hat{l}_i(k) - k - 1 \leq M + 1$ , 利用命题 3.3 和式(3-15)、(3-20), 整理有

$$\lim_{k \rightarrow \infty} \left\| x_{k+1} - x_{\hat{l}_i(k)} \right\| = 0.$$

再一次用到  $f_i$  的连续性, 得  $\lim_{k \rightarrow \infty} f_i \left( x_{\hat{l}_i(k)} \right) = \lim_{k \rightarrow \infty} f_i \left( x_{k+1} \right) = \lim_{k \rightarrow \infty} C_{i,k}$ , 将此带入到不等式(3-3)中即可得到结论。证毕。

**定理 3.8** 假设  $F$  是有下界且连续可微的。序列  $\{x_k\}$  是算法 1 生成的序列, 如果序列  $\{x_k\}$  存在聚点, 那么此聚点是  $F$  的驻点。

证明: 记  $\bar{x}$  是序列  $\{x_k\}$  的聚点, 因为  $\{x_k\} \subseteq C$  以及集合  $C$  是一个闭凸集, 故  $\bar{x} \in C$ 。同时, 序列  $\{x_k\}$  存在一个子序列  $\{x_k\}_{k \in I}$  收敛到  $\bar{x}$ , 其中  $I \subseteq N$  是一个指标集。由引理 3.7 可以得到

$$\lim_{k \rightarrow \infty} \gamma_k \varphi \left( JF \left( x_k \right) v \left( x_k \right) \right) = 0. \tag{3-21}$$

此式可以拆解成两种情况:

- (1)  $\lim_{k \rightarrow \infty} \varphi \left( JF \left( x_k \right) v \left( x_k \right) \right) = 0$ ;
- (2)  $\lim_{k \rightarrow \infty} \gamma_k = 0$ 。

当情况(1)成立时, 由引理 3.4 知  $\theta$  是连续的, 则  $\theta(\bar{x}) = \lim_{k \in I, k \rightarrow \infty} \theta(x_k)$ 。又

$$0 \geq \lim_{k \rightarrow \infty} \theta(x_k) = \lim_{k \rightarrow \infty} \left\{ \frac{\|v(x_k)\|^2}{2} + \beta \varphi \left( JF \left( x_k \right) v \left( x_k \right) \right) \right\} = \lim_{k \rightarrow \infty} \frac{\|v(x_k)\|^2}{2} \geq 0.$$

则

$$\theta(\bar{x}) = \lim_{k \in I, k \rightarrow \infty} \theta(x_k) = 0.$$

再根据注 3.2(2)可知  $\bar{x}$  是驻点。

当情况(2)成立时, 存在一个指标  $\bar{k}$  使得  $\gamma_k < \mu$  对所有的  $k \geq \bar{k}$  和  $k \in I$  成立。  $\gamma_k$  是满足条件(3-3)的步长, 故

$$F(x_k + \rho\gamma_k v(x_k)) \not\leq C_k + \sigma\rho\gamma_k JF(x_k)v(x_k),$$

也就是

$$C_k + \sigma\rho\gamma_k JF(x_k)v(x_k) - F(x_k + \rho\gamma_k v(x_k)) \notin K. \quad (3-22)$$

而左边又可以写成如下形式:

$$\begin{aligned} & C_k + \sigma\rho\gamma_k JF(x_k)v(x_k) - F(x_k + \rho\gamma_k v(x_k)) \\ &= [C_k - F(x_k)] + [F(x_k) + \sigma\rho\gamma_k JF(x_k)v(x_k) - F(x_k + \rho\gamma_k v(x_k))]. \end{aligned} \quad (3-23)$$

又因为  $F(x_k) \leq C_k$ , 即

$$C_k - F(x_k) \in K,$$

而  $K$  是一个锥, 结合引理 2.7 可得

$$F(x_k) + \sigma\rho\gamma_k JF(x_k)v(x_k) - F(x_k + \rho\gamma_k v(x_k)) \notin K. \quad (3-24)$$

根据中值定理, 可知存在  $u_k = x_k + \frac{\omega_k}{\|v(x_k)\|} \rho\gamma_k v(x_k)$  ( $\frac{\omega_k}{\|v(x_k)\|} \in [0, 1]$ ) 满足

$$F(x_k) - F(x_k + \rho\gamma_k v(x_k)) = -\rho\gamma_k JF(u_k)v(x_k),$$

将此带入到(3-24)中, 并在两边同时除以  $\rho\gamma_k$ , 整理得

$$\sigma JF(x_k)v(x_k) - JF(u_k)v(x_k) \notin K. \quad (3-25)$$

序列  $\left\{ \frac{v(x_k)}{\|v(x_k)\|} \right\}$  是一个有界序列, 则该序列存在一个收敛的子序列, 即存在指标集  $\bar{K} \subseteq I$  使得

$$\lim_{k \in \bar{K}, k \rightarrow \infty} x_k = \bar{x}, \quad \lim_{k \in \bar{K}, k \rightarrow \infty} \frac{v(x_k)}{\|v(x_k)\|} = \bar{v}.$$

利用  $\lim_{k \rightarrow \infty} \gamma_k = 0$ , 有

$$\lim_{k \in \bar{K}, k \rightarrow \infty} u_k = \lim_{k \in \bar{K}, k \rightarrow \infty} x_k + \lim_{k \in \bar{K}, k \rightarrow \infty} \rho\gamma_k \omega_k \frac{v(x_k)}{\|v(x_k)\|} = \bar{x}.$$

在(3-25)两端同时除以  $\|v(x_k)\|$ , 并对  $k \in \bar{K}$  中的  $k$  取极限, 可得

$$(\sigma - 1)JF(\bar{x})\bar{v} \notin \text{int } K.$$

考虑到  $\sigma \in (0, 1)$ , 故  $JF(\bar{x})\bar{v} \notin -\text{int } K$ , 又  $v(x_k) \neq 0$  是一个下降方向, 故  $JF(x_k)v(x_k) < 0$ , 即  $JF(x_k)v(x_k) \in -\text{int } K$ , 而  $K$  是一个闭凸尖锥, 可得  $JF(\bar{x})\bar{v} = 0$ , 从而  $\varphi(JF(\bar{x})\bar{v}) = 0$ , 再利用  $\theta$  的定义, 有

$$0 = \beta\varphi(JF(\bar{x})\bar{v}) \leq \theta(\bar{x}).$$

另一方面, 由命题 3.2(1)知  $\theta(\bar{x}) \leq 0$ , 故  $\theta(\bar{x}) = 0$ , 因此  $\bar{x}$  是  $F$  的驻点, 证毕。

### 3.3. 线性收敛

本小节假设以下条件成立:

(A1)  $\nabla f_i (i \in \{1, \dots, m\})$  是 L-Lipschitz 连续的;

(A2) 存在一个正数  $c$  满足下面的不等式成立:

$$(\nabla f_i(x))^T v(x) \leq -c \|\nabla f_i(x)\|^2, i \in \{1, \dots, m\};$$

(A3) 令  $\bar{x}$  是  $F$  的驻点, 存在正常数  $\lambda_1$  和  $\lambda_2$  满足:

$$\lambda_1 \|\nabla f_i(x)\|^2 \leq f_i(x) - f_i(\bar{x}) \leq \lambda_2 \|\nabla f_i(x)\|^2, i \in \{1, \dots, m\}.$$

下面的引理证明了极大型和平均型非单调线搜索生成的步长是有下界的。

**引理 3.9** 假设(A1)成立, 则对于任意的迭代次数  $k$  都存在一个正常数  $\gamma_{\min}$  使得  $\gamma_k > \gamma_{\min}$ 。

证明: 任取迭代次数  $k \in N$  并将其固定, 再根据  $\gamma_k$  的相关定义可得

$$F(x_k + \rho\gamma_k v(x_k)) \leq C_k + \sigma\rho\gamma_k JF(x_k) v(x_k).$$

故存在  $i(k) \in \{1, \dots, m\}$  使得

$$\begin{aligned} f_{i(k)}(x_k + \rho\gamma_k v(x_k)) &> C_{i(k),k} + \sigma\rho\gamma_k (\nabla f_{i(k)}(x_k))^T v(x_k) \\ &\geq f_{i(k)}(x_k) + \sigma\rho\gamma_k (\nabla f_{i(k)}(x_k))^T v(x_k), \end{aligned} \quad (3-26)$$

第二个不等式的成立是依据(3-2)和注 3.1 和注 3.3。再利用(A1)和  $F$  是连续可微的函数的条件, 可得

$$\begin{aligned} & f_{i(k)}(x_k + \rho\gamma_k v(x_k)) - f_{i(k)}(x_k) \\ &= \rho\gamma_k (\nabla f_{i(k)}(x_k))^T v(x_k) + \int_0^{\rho\gamma_k} (\nabla f_{i(k)}(x_k + tv(x_k)) - \nabla f_{i(k)}(x_k))^T v(x_k) dt \\ &\leq \rho\gamma_k (\nabla f_{i(k)}(x_k))^T v(x_k) + \int_0^{\rho\gamma_k} tL \|v(x_k)\|^2 dt \\ &= \rho\gamma_k (\nabla f_{i(k)}(x_k))^T v(x_k) + \frac{1}{2} \rho^2 \gamma_k^2 L \|v(x_k)\|^2. \end{aligned}$$

将上式与(3-26)相结合, 可得

$$\gamma_k \geq \frac{2(\sigma-1)(\nabla f_{i(k)}(x_k))^T v(x_k)}{\rho L \|v(x_k)\|^2}. \quad (3-27)$$

考虑到  $\sigma \in (0,1)$  和  $(\nabla f_{i(k)}(x_k))^T v(x_k) \leq 0$ , 再将(3-27)和(3-8)相结合可得

$$\gamma_k \geq \frac{(1-\sigma)(\nabla f_{i(k)}(x_k))^T v(x_k)}{\beta\rho L \max_{y \in G} \{y^T JF(x_k)v(x_k)\}} \geq \frac{1-\sigma}{\beta\rho L},$$

这也就证明了结论。

在单目标规划中, Y.H. Dai 等人[13]证明了极大型线搜索算法的线性收敛性质, 其中要求目标函数是光滑和一致凸。受此启发, 将这一性质扩展到多目标规划中并建立相关的线性收敛性, 依然在目标函数是连续可微和一致凸的条件下。下面是关于算法 1 的线性收敛性分析。

**定理 3.10** 设条件(A1)、(A2)和(A3)成立, 则由算法 1 生成的序列  $\{x_k\}$  满足: 存在  $a \in (0,1)$  使得

$$F(x_k) - F(\bar{x}) \leq a^k (F(x_0) - F(\bar{x})), \quad (3-28)$$

其中  $\bar{x}$  表示的是序列  $\{x_k\}$  的聚点。

证明: (3-28)等价于下面的不等式:

$$f_i(x_k) - f_i(\bar{x}) \leq a^k (f_i(x_0) - f_i(\bar{x})), \forall i \in \{1, \dots, m\}. \quad (3-29)$$

在证明结论之前, 先证明下面的不等式对所有的  $j=0, \dots, m(k+1)$  都成立:

$$f_i(x_{k+1-j}) \leq C_{i,k} + \sigma\gamma_{k-j} (\nabla f_i(x_{k-j}))^T v(x_{k-j}), \forall i \in \{1, \dots, m\}. \quad (3-30)$$

下面用数学归纳法来证明:

当  $j=0$  时, (3-30)可以依据(3-3)得到。

假设(3-30)在  $j \in [0, m(k+1)-1]$  处成立, 考虑到  $(\nabla f_i(x_{k-j}))^T v(x_{k-j}) \leq 0$ 、 $\sigma > 0$  和  $\gamma_{k-j} > 0$ , 有

$$\max_{0 \leq l \leq j} f_i(x_{k+1-l}) \leq C_{i,k} = \max_{0 \leq l \leq m(k)} f_i(x_{k-l}), \forall i \in \{1, \dots, m\}. \quad (3-31)$$

将(3-3), (3-2)以及  $m(k)$  的定义与上式相结合, 可得

$$\begin{aligned} f_i(x_{k+1-(j+1)}) &\leq C_{i,k-j-1} + \sigma\gamma_{k-j-1} (\nabla f_i(x_{k-j-1}))^T v(x_{k-j-1}) \\ &= \max_{0 \leq l \leq m(k-j-1)} f_i(x_{k-j-1-l}) + \sigma\gamma_{k-j-1} (\nabla f_i(x_{k-j-1}))^T v(x_{k-j-1}) \\ &\leq \max \left\{ \max_{0 \leq l \leq j} f_i(x_{k+1-l}), \max_{0 \leq l \leq m(k)} f_i(x_{k-l}) \right\} + \sigma\gamma_{k-j-1} (\nabla f_i(x_{k-j-1}))^T v(x_{k-j-1}) \\ &= \max_{0 \leq l \leq m(k)} f_i(x_{k-l}) + \sigma\gamma_{k-j-1} (\nabla f_i(x_{k-j-1}))^T v(x_{k-j-1}) \\ &= C_{i,k} + \sigma\gamma_{k-j-1} (\nabla f_i(x_{k-j-1}))^T v(x_{k-j-1}), \forall i \in \{1, \dots, m\}. \end{aligned}$$

即可证明了(3-30)对所有的  $j=0, \dots, m(k+1)$  都成立。在式(3-30)两边同时关于  $j \in [0, m(k+1)]$  取最大值, 该式即可变成

$$\begin{aligned} C_{i,k+1} &= \max_{0 \leq j \leq m(k+1)} f_i(x_{k+1-j}) \\ &\leq C_{i,k} + \max_{0 \leq j \leq m(k+1)} \sigma\gamma_{k-j} \nabla f_i(x_{k-j}) v(x_{k-j}), \forall i \in \{1, \dots, m\}. \end{aligned}$$

利用假设条件(A2)和引理 3.9, 可推导出

$$C_{i,k+1} \leq C_{i,k} - \sigma\gamma_{\min} c \left\| \nabla f_i(x_{k-j_0}) \right\|^2, \forall i \in \{1, \dots, m\},$$

其中  $\left\| \nabla f_i(x_{k-j_0}) \right\|^2 = \max_{0 \leq j \leq m(k+1)} \left\| \nabla f_i(x_{k-j}) \right\|^2$ , 再在上式两端同时减去  $f_i(\bar{x})$ , 可得

$$C_{i,k+1} - f_i(\bar{x}) \leq C_{i,k} - f_i(\bar{x}) - \sigma\gamma_{\min} c \left\| \nabla f_i(x_{k-j_0}) \right\|^2, \forall i \in \{1, \dots, m\}. \quad (3-32)$$

此时, 将证明(3-29)转变成证明下面的不等式:

$$C_{i,k+1} - f_i(\bar{x}) \leq a(C_{i,k} - f_i(\bar{x})), \forall i \in \{1, \dots, m\}. \quad (3-33)$$

如果上述不等式成立, 利用注 3.1, 有

$$\begin{aligned}
f_i(x_{k+1}) - f_i(\bar{x}) &\leq C_{i,k+1} - f_i(\bar{x}) \\
&\leq a(C_{i,k} - f_i(\bar{x})) \\
&\leq \dots \\
&\leq a^{k+1}(C_{i,0} - f_i(\bar{x})) \\
&= a^{k+1}(f_i(x_0) - f_i(\bar{x})), \forall i \in \{1, \dots, m\}.
\end{aligned}$$

这一不等式也就是(3-29)。

下面定义如下的两个数:

$$\begin{aligned}
a &:= \frac{\lambda_2}{c\sigma\gamma 2_{\min}}, \\
b &:= \frac{1}{c\sigma\gamma 2_{\min}}.
\end{aligned}$$

依据(3-2), 有

$$C_{i,k+1} - f_i(\bar{x}) = \max_{0 \leq j \leq m(k+1)} f_i(x_{k+1-j}) - f_i(\bar{x}) = f_i(x_{l_i(k+1)}) - f_i(\bar{x}),$$

$l_i(k)$  的定义如(3-10)和(3-11)。

下面分成两种情况来证明(3-33):

- (1)  $\|\nabla f(x_{k-j_0})\|^2 \geq b(C_{i,k} - f_i(\bar{x}))$  将此带入到不等式(3-32)中即可得到(3-33);
- (2)  $\|\nabla f(x_{k-j_0})\|^2 < b(C_{i,k} - f_i(\bar{x}))$  由假设条件(A3)和  $j_0$  的相关定义可以获得

$$C_{i,k+1} - f_i(\bar{x}) \leq \lambda_2 \|\nabla f_i(x_{l_i(k+1)})\|^2 \leq \lambda_2 \|\nabla f_i(x_{k-j_0})\|^2.$$

再将此与假设条件相结合即可得到(3-33)。证毕。

## 4. 结论

本文提出一种新的算法: 极大型投影梯度算法, 具体而言, 将解决约束向量优化问题的投影梯度算法与极大型非单调线搜索技术相结合, 这样做的目的是为了摆脱单调线搜索技术陷入局部收敛的可能。在适当的条件下, 证明所提算法确实不再是局部收敛的, 而是全局收敛的。进一步, 加强条件: 目标函数各分量的梯度函数是满足 Lipschitz 条件的, 并且和下降方向与目标函数满足某些关系。在此基础上证明了所提算法是线性收敛的。

## 参考文献

- [1] Wiecek, M.M. (2007) Advances in Cone-Based Preference Modeling for Decision Making with Multiple Criteria. *Decision Making in Manufacturing and Services*, **1**, 153-173. <https://doi.org/10.7494/dmms.2007.1.2.153>
- [2] De, P. and Wells, G.C.E. (1992) On the Minimization of Completion Time Variance with a Bicriteria Extension. *Operations Research*, **40**, 1148-1155. <https://doi.org/10.1287/opre.40.6.1148>
- [3] White, D.J. (1998) Epsilon-Dominating Solutions in Mean-Variance Portfolio Analysis. *European Journal of Operational Research*, **105**, 457-466. [https://doi.org/10.1016/S0377-2217\(97\)00056-8](https://doi.org/10.1016/S0377-2217(97)00056-8)
- [4] Quiroz, E.A.P., Apolinario, H.C.F., Villacorta, K.D. and Oliveira, P.R. (2019) A Linear Scalarization Proximal Point Method for Quasiconvex Multiobjective Minimization. *Journal of Optimization Theory and Applications*, **183**, 1028-1052. <https://doi.org/10.1007/s10957-019-01582-z>
- [5] Fliege, J. and Svaiter, B.F. (2000) Steepest Descent Methods for Multicriteria Optimization. *Mathematical Methods of*

- 
- Operations Research*, **51**, 479-494. <https://doi.org/10.1007/s001860000043>
- [6] Fugee, J., Grana Drummond, L.M. and Svaiter, B.F. (2010) Newton's Method for Multiobjective Optimization. *SIAM Journal on Optimization: A Publication of the Society for Industrial and Applied Mathematics*, **20**, 602-626. <https://doi.org/10.1137/08071692X>
- [7] Drummond, L. and Iusem, A.N. (2004) A Projected Gradient Method for Vector Optimization Problems. *Computational Optimization & Applications*, **28**, 5-29. <https://doi.org/10.1023/B:COAP.0000018877.86161.8b>
- [8] Yunier, J. and Cruz, B. (2013) A Subgradient Method for Vector Optimization Problems. *SIAM Journal on Optimization*, **23**, 2169-2182. <https://doi.org/10.1137/120866415>
- [9] Brito, A.S., Neto, J., Santos, P., et al. (2017) A Relaxed Projection Method for Solving Multiobjective Optimization Problems. *European Journal of Operational Research*, **256**, 17-23. <https://doi.org/10.1016/j.ejor.2016.05.026>
- [10] Bento, G.C., Cruz Neto, J.X., Santos, P.S.M. and Souza, S.S. (2018) A Weighting Subgradient Algorithm for Multiobjective Optimization. *Optimization Letters*, **12**, 399-410. <https://doi.org/10.1007/s11590-017-1133-x>
- [11] Boyd, S. and Vandenberghe, L. (2004) *Convex Optimization*. Cambridge University Press, 727. <https://doi.org/10.1017/CBO9780511804441>
- [12] Mahdavi-Amiri, N. and Salehi Sadaghiani, F. (2020) A Superlinearly Convergent Nonmonotone Quasi-Newton Method for Unconstrained Multiobjective Optimization. *Optimization Methods and Software*, **35**, 1223-1247. <https://doi.org/10.1080/10556788.2020.1737691>
- [13] Dai, Y.H. (2002) On the Nonmonotone Line Search. *Journal of Optimization Theory and Applications*, **112**, 315-330. <https://doi.org/10.1023/A:1013653923062>