

doi:10.13624/j.cnki.issn.1001-7445.2021.0216

一个共轭梯度算法与随机两人零和博弈分析

李向荣¹, 黎鹏², 卢俊宇¹, 袁功林^{*1}

(1. 广西大学 数学与信息科学学院, 广西 南宁 530004; 2. 广西大学 商学院, 广西 南宁 530004)

摘要: 为研究一种新的关于求解无约束优化问题的三项共轭梯度算法, 通过构造新的 β_k 和应用修正线搜索技术的方法, 证明了算法的性质满足充分下降性; 具有信赖域特征; 对于非凸函数满足全局收敛性, 图像处理实验表明新算法比经典 PRP 算法更具竞争力; 通过对随机两人零和博弈模型的求解, 验证了算法对实际问题的有效性。

关键词: 共轭梯度算法; 最优化; 线搜索; 收敛性; 信赖域

中图分类号: O224 **文献标识码:** A **文章编号:** 1001-7445(2021)01-0216-10

A conjugate gradient algorithm and analysis of two-player zero-sum stochastic game

LI Xiang-rong¹, LI Peng², LU Jun-yu¹, YUAN Gong-lin^{*1}

(1. College of Mathematics and Information Science, Guangxi University, Nanning 530004, China;

2. Business School, Guangxi University, Nanning 530004, China)

Abstract: To study a new three-term conjugate gradient algorithm that solves unconstrained optimization problems, by constructing new β_k formula and applying modified line search technology, the following properties are proved: the sufficient descent property is satisfied; the algorithm has the characteristics of trust region; the algorithm has the global convergence for nonconvex functions. The numeral results of the images restoration proves proposed algorithm is outstanding than classical PRP algorithm. The given algorithm is effective for a two-player zero-sum stochastic game model.

Key words: conjugate gradient algorithm; optimization; line search; convergence; trust region

0 引言

考虑如下的最优化问题:

$$\min \{f(x) \mid x \in \mathbf{R}^n\}, \quad (1)$$

收稿日期: 2018-12-01; 修订日期: 2019-01-11

基金项目: 国家自然科学基金资助项目(11661009); 广西自然科学基金重点基金资助项目(2017GXNSFDA198046); 广西高校高水平创新团队及卓越学者计划项目(桂教师范[2019]52); 中央引导地方科技发展专项资金资助项目(ZY20198003); 国家社科重点基金资助项目(17AJL012)

通讯作者: 袁功林(1976—), 男, 河南商丘人, 广西大学教授, 博士生导师; E-mail: yuangl0417@126.com。

引文格式: 李向荣, 黎鹏, 卢俊宇, 等. 一个共轭梯度算法与随机两人零和博弈分析[J]. 广西大学学报(自然科学版), 2021, 46(1): 216-225.

其中 $f:\mathbf{R}^n\rightarrow\mathbf{R}$ 并且 $f\in C^2$ 。求解该模型的方法有牛顿法、拟牛顿法、信赖域方法和共轭梯度法等,其中共轭梯度法因结构简单,存储量小被广泛使用。常见的求解问题式(1)的共轭梯度算法迭代公式为

$$x_{k+1}=x_k+\alpha_k d_k, \tag{2}$$

其中, x_k 是当前迭代点, d_k 是搜索方向, α_k 为沿着搜索方向的步长, $k=0,1,2,\cdots$ 。搜索方向 d_k 定义如下:

$$d_{k+1}=\begin{cases} -g_{k+1}, & k=0 \\ -g_{k+1}+\beta_k g_k, & k\geq 1 \end{cases}, \tag{3}$$

其中, $\beta_k\in\mathbf{R}$,不同的 β_k 决定不同的共轭梯度算法^[1-9]。著名的 PRP 算法公式^[10-11]中 β_k 有如下形式:

$$\beta_k^{\text{PRP}}=\frac{g_{k+1}^T(g_{k+1}-g_k)}{\|g_k\|^2}, \tag{4}$$

其中, $g_k=g(x_k)=\nabla f(x_k)$, $g_{k+1}=g(x_{k+1})=\nabla f(x_{k+1})$ 分别是 x_k 和 x_{k+1} 处的梯度。PRP 算法能有效处理大规模优化问题,并且数值表现优越,但其收敛性不理想。如在下列弱 Wolfe-Powell(WWP)线搜索下,它不能满足非凸函数的全局收敛性。

$$f(x_k+\alpha_k d_k)\leq f(x_k)+\delta \alpha_k g_k^T d_k, \tag{5}$$

和

$$g(x_k+\alpha_k d_k)^T d_k\geq \sigma g_k^T d_k, \tag{6}$$

其中, $\delta\in(0,\frac{1}{2})$, $\sigma\in(\delta,1)$ 。为了解决这个问题,YUAN 等^[12]提出了一个修正的 WWP 线搜索方法:

$$f(x_k+\alpha_k d_k)\leq f(x_k)+\delta \alpha_k g_k^T d_k+\alpha_k \min\left(-\delta_1 g_k^T d_k,\delta \frac{\alpha_k}{2}\|d_k\|^2\right), \tag{7}$$

和

$$g(x_k+\alpha_k d_k)^T d_k\geq \sigma g_k^T d_k+\min(-\delta_1 g_k^T d_k,\delta \alpha_k\|d_k\|^2), \tag{8}$$

其中, $\delta\in(0,\frac{1}{2})$, $\delta_1\in(0,\delta)$, $\sigma\in(\delta,1)$ 。式(7),(8)被称为 Yuan-Wei-Lu(YWL)线搜索技术,该技术能解决关于 BFGS 拟牛顿算法对一般函数的全局收敛性问题。基于该线搜索技术和 PRP 算法的特征,笔者进行更深入研究,并提出一个三项共轭梯度算法,将其命名为 LL 算法,它有如下的性质:

- ① 搜索方向满足充分下降性;
- ② 在 YWL 线搜索下新算法对一般函数具有全局收敛性;
- ③ 对大规模优化问题来说新算法比 PRP 算法优秀。

本文所用记号如下: g_k 为目标函数 $f(x)$ 在点 x_k 处的梯度, $\|\cdot\|$ 为向量的欧氏范数。

1 目标与算法

PRP 算法对一般方程来说不满足全局收敛性的一个重要原因在于它不具有充分下降性,TOUATI-AHMED 等^[13],AL-BAALI^[14],GILBERT 等^[15],HU 等^[16]表明了对共轭梯度法来说充分下降性是保证全局收敛性的关键。而王博朋等^[17]提出了如下的一种三项共轭梯度法公式的搜索方向:

$$d_{k+1}=\begin{cases} -q_{k+1}, & k=0 \\ -q_{k+1}+\frac{\left[q_{k+1}^T t_k-\min\left(q_{k+1}^T d_k,\frac{a\|t_k\|^2 q_{k+1}^T d_k}{\|q_k\|^2}\right)\right]d_k}{\max\left(b\|d_k\|\|t_k\|,\|q_k\|^2+\frac{ba\|t_k\|^2\|d_k\|^2}{\|q_k\|^2}\right)}, & k\geq 1 \end{cases}, \tag{9}$$

其中, $a>\frac{1}{4}$, $b>0$, $q_k=q(x_k)$, $q_{k+1}=q(x_{k+1})$, $t_k=q_{k+1}-q_k$ 。

上述算法能求解如下的非线性方程组:

$$q(x)=0,x\in\mathbf{R}^n, \tag{10}$$

其中, $q:\mathbf{R}^n\rightarrow\mathbf{R}^n$ 连续可微且单调。笔者发现其中的 q_k 是一个 n 维向量,是该非线性方程组的解,使得

$q(x) = 0$ 。考虑如式(1)无约束优化问题中的 $g(x) \rightarrow 0$, 这两者之间是否有联系, 下面就来证明一下。

算法(1)步骤如下:

Step 1: 给定初始点 $x_0 \in \mathbf{R}^n$, $\varepsilon \in (0, 1)$, $\delta \in (0, 1/2)$, $\delta_1 \in (0, \delta)$, $\sigma \in (\delta, 1)$, 令 $k=1$ 。

Step 2: 若 $\|g_k\| \leq \varepsilon$, 则算法停止。

Step 3: 用 YWL 线搜索选出 α_k 。

Step 4: 令迭代公式为 $x_{k+1} = x_k + \alpha_k d_k$ 。

Step 5: 如果 $\|g_{k+1}\| \leq \varepsilon$ 成立, 则算法停止。

Step 6: 通过:

$$d_{k+1} = \begin{cases} -g_{k+1}, & k=0 \\ -g_{k+1} + \frac{\left[g_{k+1}^T y_k - \min \left(g_{k+1}^T d_k, \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2} \right) \right] d_k}{\max \left(\lambda \|d_k\| \|y_k\|, \|g_k\|^2 + \frac{\gamma \mu \|y_k\|^2 \|d_k\|^2}{\|g_k\|^2} \right)}, & k \geq 1 \end{cases}, \quad (11)$$

来修正搜索方向, 其中, $\mu > \frac{1}{4}$, $\gamma > 0$, $g_k = g(x_k)$, $g_{k+1} = g(x_{k+1})$, $y_k = g_{k+1} - g_k$ 。

Step 7: 令 $k := k+1$ 并且转到 step 2。

2 下降性和收敛性的证明

为了证明算法(1)的全局收敛性, 现提出必要的假设条件:

假设 1: ① 水平集 $\Omega = \{x \in \mathbf{R}^n \mid f(x) \leq f(x_0)\}$ 是有界的。

② 函数 f 有下界, 可微, 还满足 Lipschitz 连续, 即

$$\|g(x) - g(y)\| \leq L\|x - y\|, \text{ 任意 } x, y \in \Omega, \quad (12)$$

成立, 并且 $L > 0$ 是一个常数。现在需证明关于式(11)的搜索方向 d_k 的充分下降性。

引理 1 搜索方向 d_k 由式(11)定义, 则下面的结论成立:

$$d_{k+1}^T g_{k+1} \leq -\eta \|g_{k+1}\|^2, \quad (13)$$

其中, $0 < \eta \leq 1$ 是一个常数。

证明 情形 1 如果 $\min \left(g_{k+1}^T d_k, \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2} \right) = g_{k+1}^T d_k$, 则

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2, \quad (14)$$

得证。

情形 2 如果 $\min \left(g_{k+1}^T d_k, \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2} \right) = \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2}$, 则

$$\begin{aligned} d_{k+1}^T g_{k+1} &= -\|g_{k+1}\|^2 + \frac{\left[g_{k+1}^T y_k - \min \left(g_{k+1}^T d_k, \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2} \right) \right] d_k}{\max \left(\lambda \|d_k\| \|y_k\|, \|g_k\|^2 + \frac{\gamma \mu \|y_k\|^2 \|d_k\|^2}{\|g_k\|^2} \right)} g_{k+1} \leq \\ &= -\|g_{k+1}\|^2 + \frac{\left(g_{k+1}^T y_k - \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2} \right) d_k g_{k+1}}{\|g_k\|^2 + \frac{\gamma \mu \|y_k\|^2 \|d_k\|^2}{\|g_k\|^2}} \leq \\ &= -\|g_{k+1}\|^2 + \frac{\left(g_{k+1}^T y_k - \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2} \right) d_k g_{k+1}}{\|g_k\|^2} = \end{aligned}$$

$$\frac{u^T v - \frac{1}{2}(u^2 + v^2)}{\|g_k\|^2} + \frac{-(1 - \frac{1}{4\mu}) \|g_{k+1}\|^2 \|g_k\|^2}{\|g_k\|^2} \leq -\left(1 - \frac{1}{4\mu}\right) \|g_{k+1}\|^2, \quad (15)$$

其中, $u = \frac{\|g_k\|^2}{\sqrt{2\mu}}$, $v = \frac{\sqrt{2\mu} g_{k+1}^T d_k}{\|g_k\|} y_k$ 。又不等式 $u^T v \leq \frac{1}{2}(u^2 + v^2)$ 成立, 综上所述引理, 得证。

引理 2 由式(11)定义的 d_k 有下面的结论:

$$\|d_k\| \leq c \|g_k\|, \quad (16)$$

其中, c 是一个常数。

证明

$$\begin{aligned} \|d_{k+1}\| &= \|-g_{k+1} + \beta_k d_k\| \leq \|g_{k+1}\| + \|\beta_k d_k\| \leq \|g_{k+1}\| + \|\beta_k\| \|d_k\| = \\ &\|g_{k+1}\| + \left\| \frac{\left[g_{k+1}^T y_k - \min\left(g_{k+1}^T d_k, \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2}\right) \right]}{\max\left(\lambda \|d_k\| \|y_k\|, \|g_k\|^2 + \frac{\gamma\mu \|y_k\|^2 \|d_k\|^2}{\|g_k\|^2}\right)} \right\| \|d_k\| = \\ &\|g_{k+1}\| + \left\| \frac{\max(0, g_{k+1}^T y_k - \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2})}{\max\left(\lambda \|d_k\| \|y_k\|, \|g_k\|^2 + \frac{\gamma\mu \|y_k\|^2 \|d_k\|^2}{\|g_k\|^2}\right)} \right\| \|d_k\| \leq \\ &\|g_{k+1}\| + \left\| \frac{g_{k+1}^T y_k}{\max\left(\lambda \|d_k\| \|y_k\|, \|g_k\|^2 + \frac{\gamma\mu \|y_k\|^2 \|d_k\|^2}{\|g_k\|^2}\right)} \right\| \|d_k\| + \\ &\left\| \frac{g_{k+1}^T y_k - \frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2}}{\max\left(\lambda \|d_k\| \|y_k\|, \|g_k\|^2 + \frac{\gamma\mu \|y_k\|^2 \|d_k\|^2}{\|g_k\|^2}\right)} \right\| \|d_k\| \leq \\ &\|g_{k+1}\| + \left\| \frac{g_{k+1}^T y_k}{\lambda \|d_k\| \|y_k\|} \right\| \|d_k\| + \left\| \frac{\frac{\mu \|y_k\|^2 g_{k+1}^T d_k}{\|g_k\|^2}}{\|g_k\|^2 + \frac{\gamma\mu \|y_k\|^2 \|d_k\|^2}{\|g_k\|^2}} \right\| \|d_k\| \leq \left(1 + \frac{1}{\lambda} + \frac{1}{\gamma}\right) \|g_{k+1}\|, \quad (17) \end{aligned}$$

由上知引理 2 得证。接下来的定理是关于算法的全局收敛性。

定理 1 令假设 1 成立, 并由引理 2 可知 $\|d_k\| \leq c \|g_k\|$, c 是一个大于 0 的常数。故有

$$\lim_{k \rightarrow \infty} \|g_k\| = 0. \quad (18)$$

证明 通过式(8)、(13), 可以有

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k + \min(-\delta_1 g_k^T d_k, \delta \alpha_k \|d_k\|^2) \geq \sigma g_k^T d_k, \quad (19)$$

通过式(12), 有

$$-(1 - \sigma) g_k^T d_k \leq g_{k+1}^T d_k - g_k^T d_k \leq \|g_{k+1} - g_k\| \|d_k\| \leq \alpha_k \|d_k\|^2, \quad (20)$$

所以可以得到

$$\alpha_k \geq \frac{-(1 - \sigma) g_k^T d_k}{\|d_k\|^2} \geq \eta \frac{(1 - \sigma) \|g_k\|^2}{\|d_k\|^2} \geq \frac{\eta(1 - \sigma)}{c^2}. \quad (21)$$

通过假设 1, 式(7)、(12)、(18)得到

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k + \alpha_k \min\left(-\delta_1 g_k^T d_k, \delta \frac{\alpha_k}{2} \|d_k\|^2\right) \leq$$

$$\begin{aligned} f(x_k) + (\delta - \delta_1) \alpha_k g_k^T d_k &\leq f(x_k) - \eta(\delta - \delta_1) \alpha_k \|g_k\|^2 \\ &\leq f(x_k) - \frac{\eta^2(1 - \sigma)(\delta - \delta_1)}{c^2} \|g_k\|^2. \end{aligned} \tag{22}$$

总结上述不等式从 $k=0$ 到 ∞ ,可以得到

$$\frac{\eta^2(1 - \sigma)(\delta - \delta_1)}{c^2} \sum_{k=0}^n \|g_k\|^2 \leq f(x_k) - f(x_k + \alpha_k d_k) , \tag{23}$$

$$\frac{\eta^2(1 - \sigma)(\delta - \delta_1)}{c^2} \sum_{k=0}^n \|g_k\|^2 \leq (f_0 - f_1) + (f_1 - f_2) + \cdots < +\infty , \tag{24}$$

$$\sum_{k=0}^n \|g_k\|^2 \leq f_0 - f_\infty < +\infty , \tag{25}$$

因为由假设 1 可以得出 $f(x)$ 有界,本文得到 $\|g_k\| \rightarrow 0$,完成证明。

3 数值试验

数值实验由两部分组成,一是图像修复处理实验,二是验证随机两人零和博弈模型。实验具体设置如下:

实验参数: $\delta = 0.2, \delta_1 = 0.5, \sigma = 0.85, \mu = 0.5, \gamma = 0.8$ 。

停止条件: $\frac{\|x_{k+1} - x_k\|}{\|x_k\|} < 10^{-3}$,或者 $\frac{\|f(x_{k+1})\| - \|f(x_k)\|}{\|f(x_k)\|} < 10^{-3}$ 。

数值实验中所有的代码都在 Matlab R2014b 上编写,运行环境为 8 G 内存, 2.30 GHz CPU Windows 10 操作系统。

3.1 图像处理

在图像处理实验中, LL 算法和 PRP 算法将对被脉冲噪声破坏的图像恢复原始图像。图像修复工作在优化领域具有重要的现实意义,能体现出共轭梯度法在实际应用中起到重要的作用。

实验选择 Lena (512×512)、Baboon (512×512) 和 Man (512×512)作为测试图像。图 1、图 2 给出了详细的性能结果。可以看出 LL 算法、PRP 算法都对测试图像的影像复原有效。表 1 中列出了 CPU 时间的消耗,以比较目的在于 LL 算法和 PRP 算法在图像处理方面的优劣。



(a) 测试图像 Lena



(b) LL 算法所修复图像



(c) PRP 算法所修复图像



(d) 测试图像 Man



(e) LL 算法所修复图像



(f) PRP 算法所修复图像

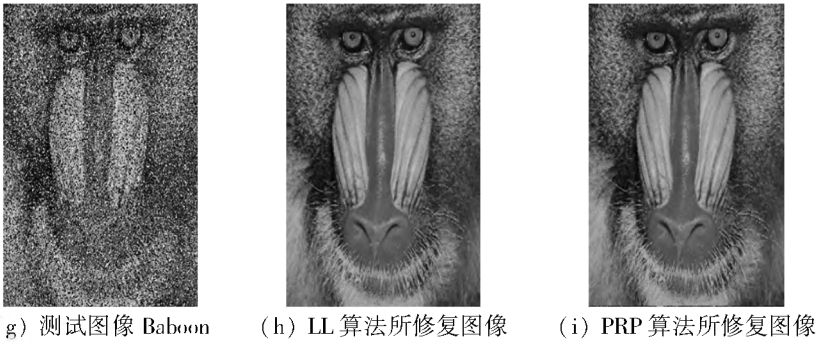


图 1 30 % 噪音测试图像与经 LL 算法, PRP 算法修复图像

Fig.1 30 % noise test image and LL algorithm, PRP algorithm to repair the image

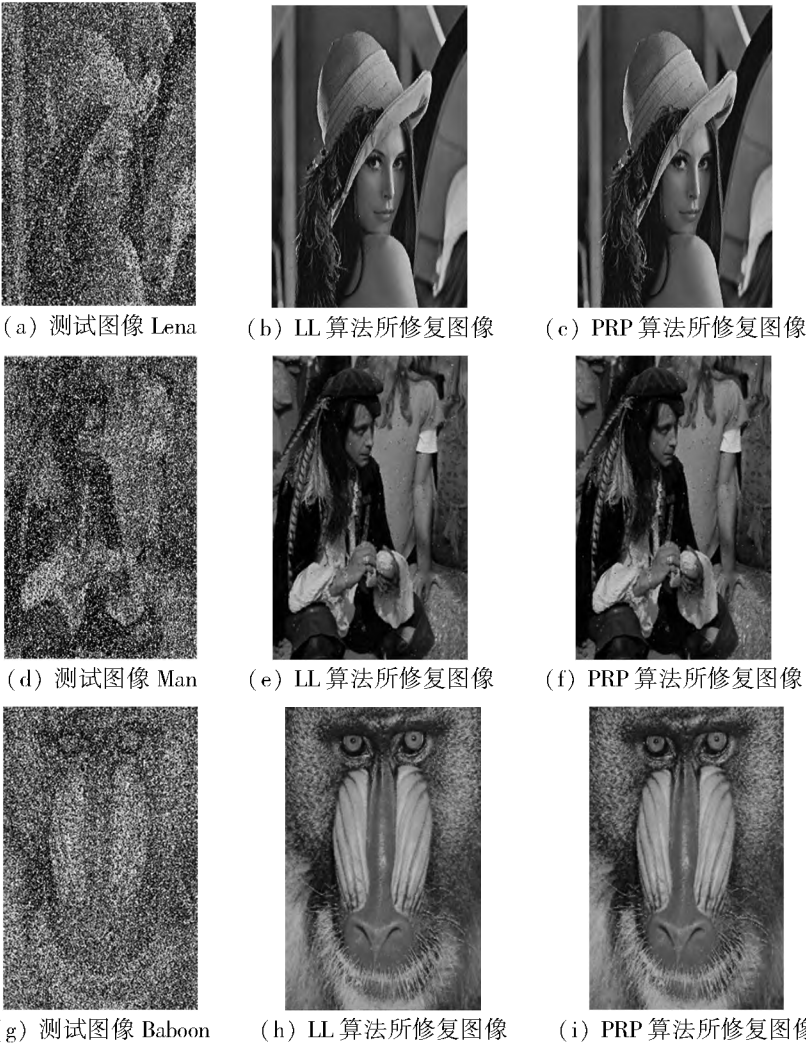


图 2 50 % 噪音测试图像与经 LL 算法, PRP 算法修复图像

Fig.2 50 % noise test image and LL algorithm, PRP algorithm to repair the image

表 1 LL 算法,PRP 算法基于 30 % 噪声修复图像使用 CPU 时间

Tab.1 LL algorithm, PRP algorithm based on 30 % noise image used by the CPU time 单位:s

算法	Lena	Man	Baboon	Total
LL 算法	1. 015 625	4. 281 250	0. 921 875	6. 218 750
PRP 算法	1. 031 250	4. 312 500	1. 359 375	6. 703 125

表 2 LL 算法,PRP 算法基于 50 % 噪声修复图像使用 CPU 时间

Tab. 2 LL algorithm, PRP algorithm based on 50 % noise image used by the CPU time 单位:s

算法	Lena	Man	Baboon	Total
LL 算法	1. 421 875	5. 328 125	1. 359 375	8. 109 375
PRP 算法	1. 625 000	5. 421 875	1. 453 125	8. 500 000

从表 1、表 2、图 1 和图 2 可以得出结论: ①当测试图像含有 50 % 噪声时,比含有 30 % 噪声的图像恢复所需的 CPU 时间较多; ②当噪声率增加时,恢复图像所需要的时间也随之增加。③LL 算法比 PRP 算法所需 CPU 时间更少,更适用于影像复原问题。

3.2 随机两人零和博弈模型实验

零和博弈又被称为游戏理论或零和游戏,是指一项游戏中,游戏者有输有赢,一方所赢正是另一方所输,而游戏的总成绩永远为零。零和博弈之所以广受关注,主要是因为人们发现社会的方方面面都能发现与零和博弈类似的局面,从个人到国家,从政治到经济,似乎无不验证了世界正是一个巨大的零和博弈场。零和博弈是博弈过程的最基本模型,理想的零和博弈对于金融市场有重要意义,在金融市场实际趋势运行中,理想零和博弈的全过程接近于一个半圆。理想零和博弈,从金融趋势的演变角度来看,最终将构成核心因子。混沌经济学研究者一直希望在证券市场寻找主宰世界命运的“混沌因子”,事实上,所有金融市场的“混沌因子”就是这么一个理想零和博弈的半圆。在这一部分中,笔者研究了一类两人零和随机博弈模型,其中通过连续时间马尔可夫过程给出了具有两个随机动力系统 $[z(t),v(t)]$ 的参与人 1 和参与人 2。在每个时间 t ,对所有参与人的收益函数^[18] 定义为 $\psi[t,z(t),v(t),i]$, $i=1,2$ 。参与人 i 通过如下一个满足耦合偏微分方程鞍点的均衡策略,最大化的期望收益函数定义为

$$0 = \psi_i(i) + \min_{\zeta_i} \max_{\mu_i} \{ -\kappa(i) + \frac{1}{2} \zeta_i^2 \delta_i^2 z^2 \psi_{zz}(i) + \frac{1}{2} \eta_i^2 v^2 \psi_{vv}(i) + \rho_i \eta_i \delta_i \zeta_i z v \psi_{zv}(i) + \\ \gamma_i [z \psi_z(i) + v \psi_v(i)] + \mu_i [\psi(j) - \psi(i)] \}, \tag{26}$$

其中, $j \neq i, i, j=1,2$, $\kappa(i)$ 代表运行成本, $\zeta_i, \rho_i, \eta_i, \gamma_i, \delta_i$ 代表参数, ζ_i, μ_i 是控制变量, 后缀代表 ψ 对空间变量 (z,v) 和时间变量 t 的导数。与许多偏微分方程类似, 耦合偏微分方程 (26) 具有终端条件:

$$\psi(T,z(t),v(t),i) = \pi(T,z,v), \tag{27}$$

以及下列边界条件:

$$\left\{ \begin{aligned} \psi(t,z,\infty,i) &= U(t,z,i), \psi(t,\infty,v,i) = W(t,v,i) \\ \psi(t,z,0,i) &= U^0(t,z,i), \psi(t,0,v,i) = W^0(t,v,i) \end{aligned} \right\}, \tag{28}$$

其中, $\pi(T,z,v), U(t,z,i), U^0(t,z,i), W(t,v,i), W^0(t,v,i)$ 是有确定的表达式的函数。通常,控制变量 ζ_i 的域为 $[-1,1]$, 并且 μ_i 有 $\mu_i^- = \min \{ \mu_i \}$ 和 $\mu_i^+ = \max \{ \mu_i \}$, $(i=1,2)$ 的域名 (μ_i^-, μ_i^+) 。时间变量 t 和空间变量 (z,v) 为非负值是合理的。为成功地求解方程 (26), 问题分为两步进行:

状态 1:

$$0 = \psi_i(1) + \min_{\zeta_1} \max_{\mu_1} \{ -\kappa(1) + \frac{1}{2} \zeta_1^2 \delta_1^2 z^2 \psi_{zz}(1) + \frac{1}{2} \eta_1^2 v^2 \psi_{vv}(1) + \rho_1 \eta_1 \delta_1 \zeta_1 z v \psi_{zv}(1) + \\ \gamma_1 [z \psi_z(1) + v \psi_v(1)] + \mu_1 [\psi(2) - \psi(1)] \}, \tag{29}$$

状态 2:

$$0 = \psi_i(2) + \min_{\zeta_2} \max_{\mu_2} \{ -\kappa(2) + \frac{1}{2} \zeta_2^2 \delta_2^2 z^2 \psi_{zz}(2) + \frac{1}{2} \eta_2^2 v^2 \psi_{vv}(2) + \rho_2 \eta_2 \delta_2 \zeta_2 z v \psi_{zv}(2) + \\ \gamma_2 [z \psi_z(1) + v \psi_v(2)] + \mu_2 [\psi(1) - \psi(2)] \}. \tag{30}$$

问题(29)、(30)分别具有 $i=1$ 和 $i=2$ 的边界条件 (28) 和终端条件 (27)。这两个问题都需要控制变量 $\zeta_i, \mu_i (i=1,2)$ 。问题(29)和(30)称为二阶偏微分方程, 其中含有一阶偏微分项 $\psi_t, \psi_v, \psi_z, \psi_{tv}, \psi_{zz}, \psi_{zv}$ 。下面列出有限差分公式^[19], 这些公式将在程序用到:

$$\psi_t(i) \rightarrow \frac{\psi_{j,k}^n(i) - \psi_{j,k}^{n-1}(i)}{\tau}, \tag{31}$$

$$\psi_z(i) \rightarrow \frac{\psi_{j+1,k}^n(i) - \psi_{j,k}^n(i)}{h_1}, \quad (32)$$

$$\psi_{zz}(i) \rightarrow \frac{\psi_{j+2,k}^n(i) - 2\psi_{j+1,k}^n(i) - \psi_{j,k}^n(i)}{h_1^2}, \quad (33)$$

$$\psi_v(i) \rightarrow \frac{\psi_{j,k+1}^n(i) - \psi_{j,k}^n(i)}{h_2}, \quad (34)$$

$$\psi_{vv}(i) \rightarrow \frac{\psi_{j,k+2}^n(i) - 2\psi_{j,k+1}^n(i) - \psi_{j,k}^n(i)}{h_2^2}, \quad (35)$$

$$\psi_{zv}(i) \rightarrow \frac{\psi_{j+1,k+1}^n(i) - 2\psi_{j,k+1}^n(i) - \psi_{j+1,k}^n(i) + \psi_{j,k}^n(i)}{h_2^2}. \quad (36)$$

其中, $\psi_{j,k}^n(i) = \psi(j h_1, k h_2, n\tau, i)$, $\psi_{j+1,k}^n(i) = \psi[(j+1) h_1, k h_2, n\tau, i]$, $\psi_{j+1,k+1}^n(i) = \psi[(j+1) h_1, (k+1) h_2, n\tau, i]$, 其他公式类似。 $n = 1, 2, \dots; j = 1, 2, \dots; k = 1, 2, \dots$; h_1, h_2, τ 分别是相对于变量 z, v, t 的步长。为了便于分析和运算, 设置 $h_1 = h_2 = h$, 关于 z, v, t 的最大网格点数为 N , 如果 $z_0 = v_0$, 条件 $h_1 = h_2 = h$ 能推出 $z_\infty = v_\infty$ 。YUAN 等^[20]通过极小化方程的 RHS, 充分利用一阶条件, 分别得到

$$\mu_1^* = \mu_1^+ \lambda_{|\psi(2) > \psi(1)|} + \mu_1^- \lambda_{|\psi(2) < \psi(1)|}, \quad \mu_2^* = \mu_2^+ \lambda_{|\psi(1) > \psi(2)|} + \mu_2^- \lambda_{|\psi(1) < \psi(2)|}, \quad (37)$$

其中, λ 为指示函数, 定义如下:

$$\lambda_{|x > y|} = \begin{cases} 1, & \text{如果 } x > y \\ 0, & \text{其他情况} \end{cases}, \quad (38)$$

(ζ^*, μ^*) 表示问题(29)、(30)的最佳选择, 一般情况下 ζ^* 控制变量的范围属于 $[-1, 1]$ 是可行的。为方便起见, 假设这两个问题的解是存在的。首先把 $\mu^* = (\mu_1^*, \mu_2^*)$ 代入式(29)、(30), 定义函数:

$$\begin{aligned} \min_{\zeta_1} f_{\text{pde}}(\zeta_1) = \min_{\zeta_1} \{ & -\kappa(1) + \frac{1}{2} \zeta_1^2 \delta_1^2 z^2 \psi_{zz}(1) + \frac{1}{2} \eta_1^2 v^2 \psi_{vv}(1) + \rho_1 \eta_1 \delta_1 \zeta_1 z v \psi_{zv}(1) + \\ & \gamma_1 [z z \psi_z(1) + v \psi_v(1)] + \mu_1^* [\psi(2) - \psi(1)] \}, \end{aligned} \quad (39)$$

和

$$\begin{aligned} \min_{\zeta_2} f_{\text{pde}}(\zeta_2) = \min_{\zeta_2} \{ & -\kappa(2) + \frac{1}{2} \zeta_2^2 \delta_2^2 z^2 \psi_{zz}(2) + \frac{1}{2} \eta_2^2 v^2 \psi_{vv}(2) + \rho_2 \eta_2 \delta_2 \zeta_2 z v \psi_{zv}(2) + \\ & \gamma_2 [z z \psi_z(2) + v \psi_v(2)] + \mu_2^* [\psi(1) - \psi(2)] \}. \end{aligned} \quad (40)$$

本文的目标是通过 LL 算法获得式(39)、(40)的最优(或近似)解 ζ_1 和 ζ_2 。在本次实验中, 成本函数 $\kappa = 0, \mu_1 \in (0.04, 0.14), \mu_2 \in (0.15, 0.24), \rho_1 = \rho_2 = 0.3, \delta_1 = 0.1, \delta_2 = 0.3, \eta_1 = 0.25, \eta_2 = 0.45, \gamma_1 = 0.03, \gamma_2 = 0.01, M = 100$, 终端时间 $T = 1, \tau = 0.02$ 。空间参数 (z, v) 被放置到一个正方形 $[0, z_{\max}] \times [0, v_{\max}]$ 中, 初始对 $(z_0, v_0) = (1, 1)$, $(z_{\max}, v_{\max}) = (20, 20)$ 、 $(30, 30)$ 、 $(100, 100)$ 分别进行求解, 详细结果如图 3 至图 5。

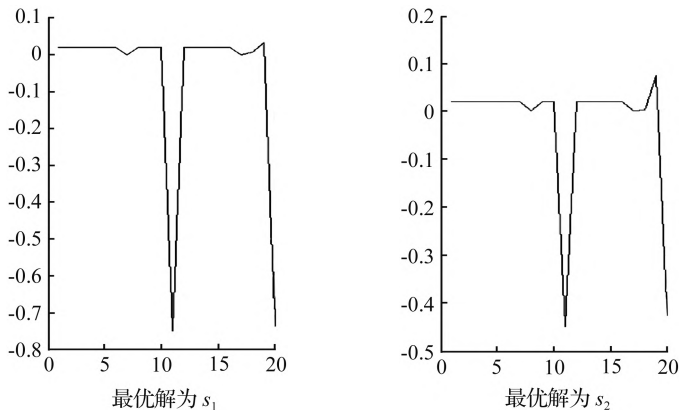


图 3 20 个离散点时的表现 ($t=0.8$)

Fig. 3 Representation at 20 discrete points ($t=0.8$)

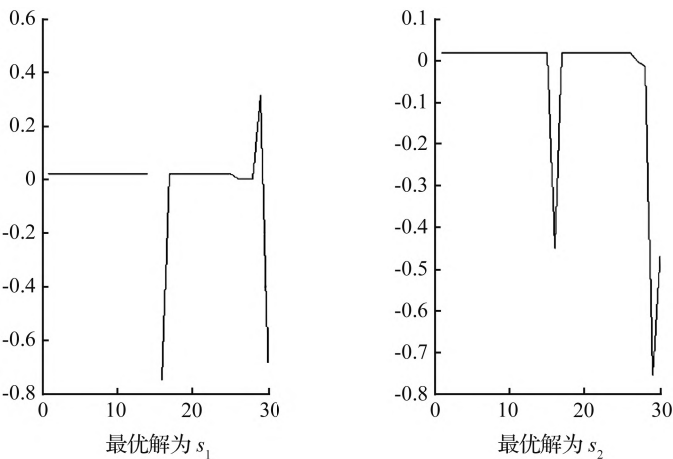


图 4 30 个离散点时的表现($t=0.8$)

Fig. 4 Representation at 30 discrete points ($t=0.8$)

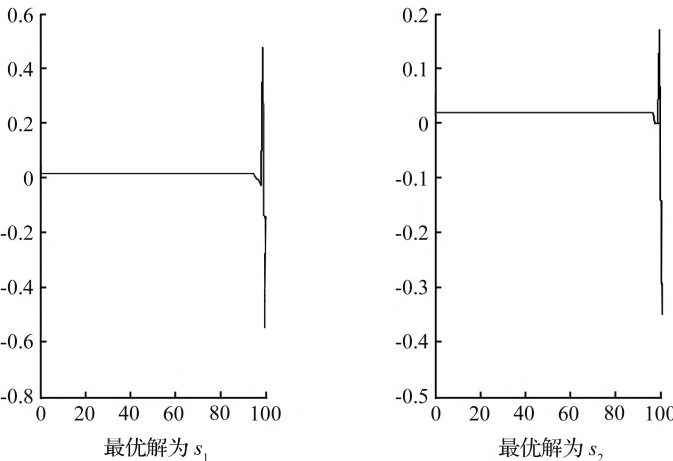


图 5 100 个离散点时的表现($t=0.8$)

Fig. 5 Representation at 100 discrete points ($t=0.8$)

从图 3 至图 5 可以看出,控制变量 $\zeta_1 \in [-1,1], \zeta_2 \in [-1,1]$ 是成立的,同模型规定的变量范围相符,因而可以得出建议的算法对时间变量 $t=0.8$ 而言,至少能成功求解出离散点个数为 20, 30 和 100 的情形,验证了算法的有效性。关于其他时间点 t 以及更多离散点情况,有待进一步验证。

4 总结与展望

对于求解无约束优化问题,笔者根据文献[21]的启发,在求解无约束问题时应用了文献中的方向,验证该算法在能求解非线性方程组问题时,能否求解无约束优化问题,并在一定条件下,证明了算法的充分下降性和全局收敛性等算法性质。同时在数值实验结果中通过对比,确定本文所提算法比 PRP 算法有效。

笔者相信,随着社会的高速发展,无约束最优化的问题会被人们更加重视,会有越来越多的学者投入到无约束问题的研究中,无约束问题的研究将会在 21 世纪取得长足的进步。

参考文献:

[1] WEI Z, YAO S, LIU L. The convergence properties of some new conjugate gradient methods [J]. Applied Mathematics and Computation, 2006,183(2):1341-1350.

- [2] WU Y, ZHANG M, LI Y. Modified three-term liu-storey conjugate gradient method for solving unconstrained optimization problems and image restoration problems[J]. Mathematical Problems in Engineering, 2020,2020:1-20.
- [3] WAZIRI M, AHMED K, SANI' U J. A Dai-Liao conjugate gradient method via modified secant equation for system of nonlinear equations[J]. Arabian Journal of Mathematics, 2020(9):443-457.
- [4] YUAN G, LI T, HU W. A conjugate gradient algorithm for large-scale nonlinear equations and image restoration problems [J]. Applied Numerical Mathematics, 2020,147(Jan.):129-141.
- [5] DONG X. A modified nonlinear Polak-Ribière-Polyak conjugate gradient method with sufficient descent property[J]. Calcolo, 2020,57(3):1-14.
- [6] HU Y, LIU L, WANG Y. Wei-Yao-Liu conjugate gradient algorithm for nonsmooth convex optimization problems[J]. Statistics Optimization & Information Computing, 2020,8(2):403-413.
- [7] YUAN G, MENG Z, LI Y. A modified hestenes and stiefel conjugate gradient algorithm for large-scale nonsmooth minimizations and nonlinear equations [J]. Journal of Optimization Theory and Applied, 2016,168(1):129-152.
- [8] YUAN G, WEI Z, LI G. A modified Polak-Ribière-Polyak conjugate gradient algorithm for nonsmooth convex programs [J]. Computational and Applied Mathematics, 2014,255:86-96.
- [9] YUAN G, ZHANG M. A three-terms Polak-Ribière-Polyak conjugate gradient algorithm for large-scale nonlinear equations [J]. Computational and Applied Mathematics, 2015,286:186-195.
- [10] POLYAK B. The conjugate gradient method in extremal problems [J]. Ussr Computational Mathematics and Mathematical Physics, 1969,9(4):94-112.
- [11] POWELL M. Restart procedures of the conjugate gradient method [J]. Mathematical Programming, 1977,12(1):241-254.
- [12] YUAN G, WEI Z, LU X. Global convergence of BFGS and PRP methods under a modified weak Wolfe-Powell line search [J]. Applied Mathematics Modelling, 2017(47):811-825.
- [13] TOUATI-AHMED D, STOREY C. Efficient hybrid conjugate gradient techniques [J]. Journal of Optimization Theory and Applied, 1990,64(2):379-394.
- [14] Al-BAALI A. Descent property and global convergence of the Fletcher-Reeves method with inexact line search [J]. IMA Journal Numerical Analysis, 1985,5(1),121-124.
- [15] GILBERT J, NOCEDAL J. Global convergence properties of conjugate gradient methods for optimization [J]. SIAM Journal on Optimization, 1992,2(1):21-42.
- [16] HU Y, STOREY C. Global convergence result for conjugate gradient methods [J]. Journal of Optimization Theory and Applications, 1991,71(2):399-405.
- [17] 王博朋, 袁功林, 李春念. 求解非线性方程组的一种修正 WPRP 共轭梯度法[J]. 广西大学学报(自然科学版), 2017,42(5):1967-1973.
- [18] ANDREI N. An unconstrained optimization test function collection [J]. Adv Model Optim, 2008,10(1),147-161.
- [19] RAMACHANDRAN K, TSOKOS C. Stochastic differential games: theory and applications [M]. Paris: Atlantis Press, 2014.
- [20] YUAN G, LI X. A numerical algorithm for the coupled PDEs control problem [J]. Computational Economics, 2019, 53, 697-707.
- [21] YUAN G, LU X, WEI Z. A conjugate gradient method with descent direction for unconstrained optimization [J]. Journal of Computational and Applied Mathematics, 2009,233(2):519-530.