

基于 BFGS 算法的广义 Lagrange 乘子法研究

熊茜, 吴泽忠

(成都信息工程大学应用数学学院, 四川 成都 610225)

摘要: 广义 Lagrange 乘子法是解决约束优化的问题的一种重要方法, 基于 BFGS 算法, 利用 MATLAB 工具, 研究了初始点的选取、罚因子的改变及罚因子修正系数的改变对该算法收敛效果的影响。结果表明: (1) 对于初始点的选取, 应尽量在最优点附近进行取值, 才能有不错的收敛效果。(2) 罚因子过小或过大都对算法求解问题产生困难。如果罚因子太小, 大量的搜索时间将花费在非可行域, 使迭代次数增加。另一方面, 如果罚因子过大, 算法将很难被推进到可行域以内, 导致算法收敛失败。(3) 随着罚因子修正系数的变化, 随时会出现无法收敛的现象, 故该系数的值应在迭代成功率相对较高的分段选取。

关键词: 应用数学; 最优化理论; 约束优化; 广义 Lagrange 乘子法; 罚因子; 修正系数

中图分类号: O224

文献标志码: A

doi: 10.16836/j.cnki.jcui.2020.02.014

0 引言

为了求解等式约束优化问题, Hestenes^[1]提出了 Lagrange 乘子法, 又称 PH 算法。随后, Rockfellar^[2-3]推广到求解不等式约束优化问题, 称为广义 Lagrange 乘子法, 简称 PHR 算法。其基本思想是: 从原问题的拉格朗日函数出发, 再加上适当的罚函数, 得到的函数称为广义 Lagrange 函数, 从而将原问题转化为求解广义拉格朗日函数的一系列无约束优化子问题。

W. C. Davidon^[4]提出拟牛顿法, 该方法在解决非线性优化问题十分有效, 在这种算法提出后, 各个国家的数学家对其进行改进, 提出了各种变形公式及改进算法。Broyden^[5]研究得到 BFGS 算法, BFGS 算法也是目前最有效的拟牛顿校正。本文将在计算广义拉格朗日函数的无约束优化子问题时采用 BFGS 算法。

在工业生产与工程应用中, 广义拉格朗日乘子法得到很好应用, 根据算法定义, 广义拉格朗日乘子法的收敛效果与收敛速度会受到初始点、初始罚因子值及罚因子修正系数值的影响^[6]。如果适当改变这些参数的值, 拉格朗日乘子法的收敛速度和收敛效果会发生怎样的改变? 现就初始点、罚因子及罚因子修正系数分别单独改变, 研究对拉格朗日乘子法收敛速度和收敛效果的影响。

1 预备知识

考虑一般约束优化问题, 即同时带有等式和不等式约束的优化问题:

$$\begin{aligned} & \min f(\mathbf{x}), \mathbf{x} \in \mathbf{R}^n, \\ \text{s. t. } & \begin{cases} h_i(\mathbf{x}) = 0, & i = 1, 2, \dots, l \\ g_i(\mathbf{x}) \geq 0, & i = 1, 2, \dots, m \end{cases} \end{aligned} \quad (1)$$

1.1 广义 Lagrange 乘子法

针对混合约束优化问题(1), 构造广义拉格朗日函数为

$$\begin{aligned} \psi(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \alpha) = & f(\mathbf{x}) - \sum_{i=1}^l \mu_i h_i(\mathbf{x}) + \frac{\alpha}{2} \sum_{i=1}^l h_i^2(\mathbf{x}) + \\ & \frac{1}{2\alpha} \sum_{i=1}^m ([\min\{0, \alpha g_i(\mathbf{x}) - \lambda_i\}]^2 - \lambda_i^2). \end{aligned} \quad (2)$$

乘子迭代公式为

$$\begin{aligned} (\boldsymbol{\mu}_{k+1})_i = & (\boldsymbol{\mu}_k)_i - \alpha h_i(\mathbf{x}_k), i = 1, 2, \dots, l \\ (\boldsymbol{\lambda}_{k+1})_i = & \max\{0, (\boldsymbol{\lambda}_k)_i - \alpha g_i(\mathbf{x}_k)\}, i = 1, 2, \dots, m \end{aligned} \quad (3)$$

令

$$\beta_k = \left(\sum_{i=1}^l h_i^2(\mathbf{x}_k) + \sum_{i=1}^m \left[\min\{g_i(\mathbf{x}_k), \frac{(\boldsymbol{\lambda}_k)_i}{\alpha}\} \right]^2 \right)^{1/2} \quad (4)$$

则有终止准则为

$$\beta_k \leq \varepsilon \quad (5)$$

其中, $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $\boldsymbol{\mu}$ 是等式约束乘子向量, $\boldsymbol{\lambda}$ 是不等式约束乘子向量, α 是罚因子, β_k 是检验终止条件的值。

1.2 BFGS 算法

为了求解无约束优化问题 $\min \psi(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}, \alpha)$, 选择 BFGS 算法。

拟牛顿法的基本思想为: 在基本牛顿法中使用 Hesse 矩阵 $\mathbf{G}_k = \nabla^2 f(\mathbf{x}_k)$ 的某个近似矩阵 \mathbf{B}_k 取代 \mathbf{G}_k , 通过构造 \mathbf{B}_k , 计算得到的 BFGS 秩为 2 的修正公式为

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k + y_k y_k^T}{s_k^T B_k s_k + y_k^T s_k} \quad (6)$$

注: G_k 为 Hesse 矩阵, B_k 为 G_k 的近似矩阵, s_k 为位移, y_k 为梯度差。

1.3 基于 BFGS 算法的广义的拉格朗日乘子法计算步骤

Step 1: 给定初始点 $x_0 \in \mathbf{R}^n$, 等式约束乘子向量 $\mu_1 \in \mathbf{R}^l$, 不等式约束乘子向量 $\lambda_1 \in \mathbf{R}^m$, 终止准则 $0 \leq \varepsilon = 1$, 初始化对称正定矩阵 B_0 , 取为 $G(x_0)$ 或单位矩阵 I_n , 罚因子 $\alpha_1 > 0$, 参数 $\theta \in (0, 1)$, $\eta > 1$, $k = 1$;

Step 2: 利用 BFGS 算法, 以 x_{k-1} 为初始点, 以 Step 3 到 Step 7 作为内循环求解无约束问题

$$\min \psi(x, \mu_k, \lambda_k, \alpha_k)$$

Step 3: 计算 f 在点 x 的梯度值 $g_k = \nabla f(x_k)$ 。若 $\|g_k\| \leq \varepsilon$, 终止程序, 输出 x_k 为 f 极小点。否则计算下一步。

Step 4: 解线性方程组 $B_k d_k = -g_k$, 得到 d_k 。

Step 5: 设 m_k 是下列不等式的最小非负整数的解 m :

$$f(x_k + \beta^m d_k) \leq f(x_k) + \alpha \beta^m g_k^T d_k.$$

令 $\delta_k = \beta^{m_k}$, $x_{k+1} = x_k + \delta_k d_k$ 。

Step 6: 由 BFGS 秩为 2 的修正公式确定 B_{k+1} 。

Step 7: 令 $k = k + 1$, 转 Step 3。得到一般约束的最小值 x_k , 其中 $\psi(x, \mu_k, \lambda_k, \alpha_k)$ 为一般的广义拉格朗日函数。

Step 8: 若 $\beta_k \leq \varepsilon$, 则停止迭代, 得到一般约束优化问题最小值的近似值; 否则执行 Step 9。

Step 9: 更新乘子向量。

不等式约束的乘子迭代式:

$$(\mu_{k+1})_i = (\mu_k)_i - \alpha h_i(x_k), i = 1, 2, \dots, l$$

等式约束的乘子迭代式:

$$(\lambda_{k+1})_i = \max\{0, (\lambda_k)_i - \alpha g_i(x_k)\}, i = 1, 2, \dots, m$$

Step 10: 若 $\beta_k \geq \theta \beta_{k-1}$, 令 $\alpha_{k+1} = \eta \alpha_k$; 否则, $\alpha_{k+1} = \alpha_k$;

Step 11: 令 $k = k + 1$, 转 Step 1。

2 算例实验变量选取分析

数值实验基于 MATLAB(R2016a), 通过对初始点 $x^{(0)}$ 、初始罚因子 α 及罚因子修正系数 η 的选取研究其对混合约束优化问题的影响。数值实验中设置精度为 1×10^{-10} 。研究 3 种不同参数对算法收敛速度的影响, 因此在实验时, 当研究一个参数的影响时, 另外两个参数不变。

2.1 初始点的选取

对于初始点的选取, 考虑罚因子 α 及罚因子修正

系数 η 不变。在罚因子 α 及罚因子修正系数 η 不变的情况下, 通过选取不同的初始点研究算法的内循环迭代次数和外循环迭代次数的变化。

2.2 罚因子的选取

对于初始罚因子 α 的选取, 考虑初始点 $x^{(0)}$ 及罚因子修正系数 η 不变。由于罚因子的作用是对迭代点不在约束范围内的情况进行惩罚, 因此罚因子的选取应大于 0, 且如果罚因子过大, 则有可能造成算法收敛困难或不收敛^[7-8]。为了研究罚因子 α 选取多大时算法收敛速度最佳, 文中选取大于 0 的不同罚因子 α 值进行研究。

2.3 罚因子修正系数的选取

对于罚因子修正系数 η 的选取, 考虑初始点 $x^{(0)}$ 及初始罚因子 α 不变。根据之前的研究可知, 如果罚因子修正系数 η 过小, 罚因子 α 的增加速度会过慢, 可能起不到惩罚的目的, 从而造成算法不收敛^[9]。如果罚因子修正系数 η 过大, 罚因子 α 很快增加到一个大数, 可能造成算法收敛困难或者不收敛。由于罚因子修正系数 η 的作用是令罚因子的值在迭代后增加, 因此罚因子修正系数 η 应大于 1^[10]。综上, 应选取适当大小的罚因子修正系数 η 值进行研究。

2.4 算法内外循环迭代说明

对于广义拉格朗日函数, 使用 BFGS 算法对该函数进行计算, 在计算过程中, 拉格朗日乘子法作为算法的外部, 外部的迭代会对罚因子 α 、罚因子修正系数 η 等进行更新, 并且重新更新拉格朗日函数; BFGS 算法作为算法内部, 内部的迭代是根据选定的初始点计算拉格朗日函数的最优解^[11-12]。将拉格朗日乘子法的迭代次数作为外迭代次数, BFGS 算法的迭代次数作为内迭代次数进行对比研究。

3 算案例

实例 1 考虑二元非线性混合约束的优化问题

$$\begin{aligned} \min f(x) &= (x_1 - 2)^2 + (x_2 - 3)^2 \\ \text{s. t. } &\begin{cases} 2x_1 - x_2 - 1 = 0 \\ (x_1 - 2)^2 - x_2 \geq 0 \end{cases} \end{aligned}$$

(1) 不同初始点对收敛速度的影响

设置精度为 10^{-10} , 修正罚因子判别参数 $\theta = 0.8$, 初始罚因子修正系数 $\eta = 2$, 选取初始罚因子 α 值分别为 10 与 30, 将初始点选择从 $(-200, -200)$ 依次增加到 $(200, 200)$, 计算结果见表 1、表 2。

表 1 罚因子为 10 时不同初始点对迭代次数的影响

初始点编号	初始点 $x^{(0)}$	最优值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
-200	(-200,-200)	5.0000	11	90	(1.0000,1.0000)	1.0883E-11
-199	(-199,-199)	5.0000	11	105	(1.0000,1.0000)	1.0181E-11
...
0	(0,0)	5.0000	11	70	(1.0000,1.0000)	1.0817E-11
1	(1,1)	5.0000	11	1070	(1.0000,1.0000)	9.4392E-12
2	(2,2)	5.0000	11	1061	(1.0000,1.0000)	6.1189E-12
3	(3,3)	5.0000	11	1061	(1.0000,1.0000)	6.1189E-12
4	(4,4)	45.0000	20	2133	(5.0000,9.0000)	3.7444E-11
...
199	(199,199)	45.0000	20	151	(5.0000,9.0000)	3.7296E-11
200	(200,200)	45.0000	20	2134	(5.0000,9.0000)	3.7208E-11

表 2 罚因子为 30 时不同初始点对迭代次数的影响

初始点编号	初始点 $x^{(0)}$	最优值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
-200	(-200,-200)	45.0000	26	10517	(5.0000,9.0000)	4.3738E-11
...
-196	(-196,-196)	45.0000	26	6483	(5.0000,9.0000)	1.4515E-11
-195	(-195,-195)	5.0000	21	4619	(1.0000,1.0000)	4.1118E-11
...
1	(1,1)	5.0000	7	65	(1.0000,1.0000)	9.7740E-11
2	(2,2)	5.0000	7	47	(1.0000,1.0000)	9.7659E-11
3	(3,3)	5.0000	7	92	(5.0000,9.0000)	9.8748E-11
4	(4,4)	45.0000	12	3050	(5.0000,9.0000)	2.2912E-11
...
192	(192,192)	45.0000	26	6983	(50000,9.0000)	4.4685E-11
193	(193,193)	5.0000	21	4862	(1.0000,1.0000)	3.8656E-11
...
200	(200,200)	5.0000	21	5758	(1.0000,1.0000)	4.1398E-11

由表 1、表 2 可知,该优化问题的最优解为(1,1),最小值为 5。当罚因子 $\alpha=10$,在初始点编号 $p \in [4, 200]$ 时,算法陷入局部收敛。当罚因子 $\alpha=30$ 时,在初

始点编号 $p \in [-200, -196]$ 与 $p \in [4, -192]$ 时,算法陷入局部收敛。故将初始点选取范围去除陷入局部收敛的点后再作图,如图 1 所示。

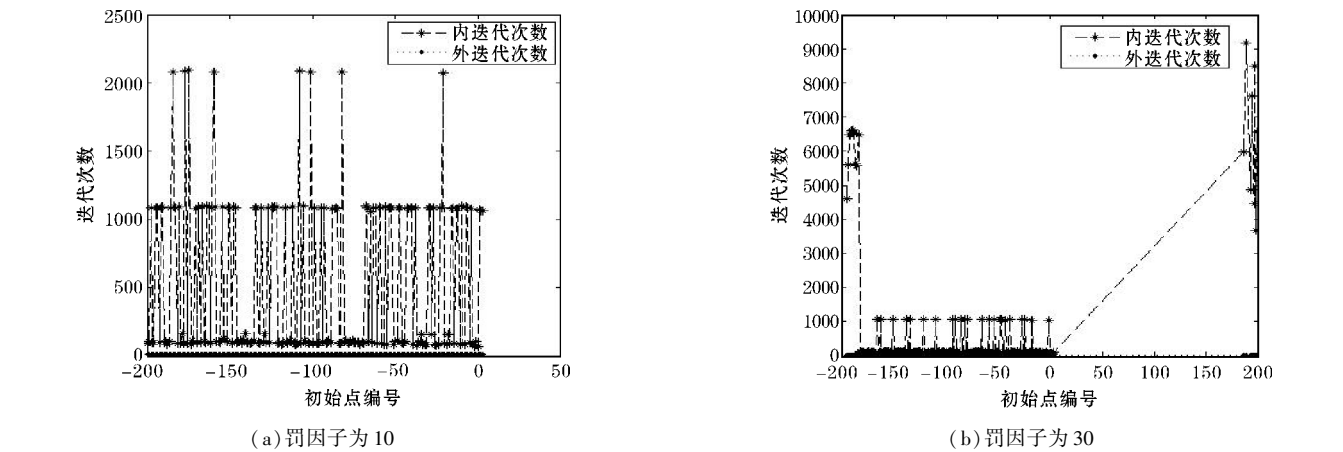


图 1 不同初始点对迭代次数的影响

结果说明与分析:

当罚因子 $\alpha=10$ 时,在初始点编号 $p \in [-200,3]$ 时,内迭代次数集中在 100 次左右与 1080 次左右,其中部分初始点迭代次数在 2100 次左右。对于初始点的选取,应尽量选取小于最优点的值,可以得到不错的收敛效果。

当罚因子 $\alpha=30$ 时,在初始点编号 $p \in [-200,-194]$ 时,内迭代次数集中在 6600 次左右与 5600 次左右;在初始点编号 $p \in [-193,3]$ 时,内迭代次数集中于 100 次左右,部分迭代次数集中在 1070 次左右;在初始点编号 $p \in [193,200]$ 时,内迭代次数在 4400 次以上。对于初始点的选取,应尽量选取小于最优点的

值,但是不能过小,过小的初始点会增加收敛次数。

当选择较大的罚因子时,会出现部分点迭代次数过大。若选取的初始点小于最优点时大概率会很快收敛,若选取的初始点大于最优点时大概率会陷入局部收敛,所以,选取的值应尽量靠近最优点,且选取的点较小。

(2)不同罚因子对收敛速度的影响

设置精度为 10^{-10} ,修正罚因子判别参数 $\theta=0.8$,选取初始罚因子修正系数 $\eta=2$,由对初始点选取的研究可知初始点编号应选取 $p \leq 3$ 最合适,故从中选取 2 个初始点分别为 $(1,1)$ 和 $(-10,-10)$,选取不同的初始罚因子,计算结果见表 3、表 4。

表 3 初始点为(1,1)时不同初始罚因子对收敛速度的影响

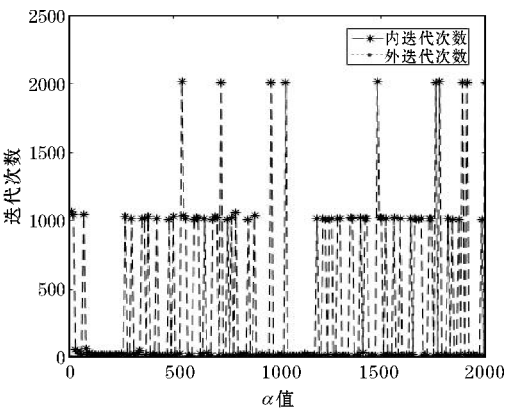
罚因子 α	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
1	5.0000	42	249	(1.0000,1.0000)	9.7532E-11
2	5.0000	22	126	(1.0000,1.0000)	8.6939E-11
3	5.0000	18	119	(1.0000,1.0000)	3.9577E-11
4	5.0000	15	103	(1.0000,1.0000)	8.8131E-11
...
12	5.0000	8	1047	(1.0000,1.0000)	7.1044E-11
13	5.0000	7	65	(1.0000,1.0000)	9.7741E-11
...
100	5.0000	6	41	(1.0000,1.0000)	2.5307E-12
101	5.0000	6	27	(1.0000,1.0000)	1.3000E-12
...
1998	5.0000	4	1013	(1.0000,1.0000)	3.3106E-12
1999	5.0000	4	17	(1.0000,1.0000)	1.6047E-13
2000	5.0000	4	1009	(1.0000,1.0000)	1.0142E-12

表 4 初始点为(-10,-10)时不同初始罚因子对收敛速度的影响

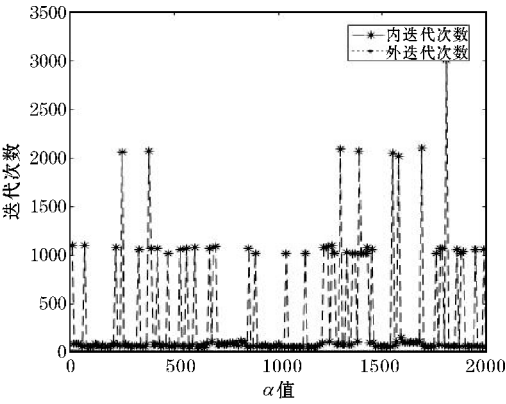
罚因子 α	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
1	5.0000	42	304	(1.0000,1.0000)	9.8912E-12
2	5.0000	22	134	(1.0000,1.0000)	8.6665E-11
3	5.0000	18	132	(1.0000,1.0000)	3.9386E-11
4	5.0000	15	117	(1.0000,1.0000)	8.8131E-11
...
10	5.0000	11	1097	(1.0000,1.0000)	9.8040E-12
11	5.0000	8	84	(1.0000,1.0000)	7.1082E-11
...
100	5.0000	6	61	(1.0000,1.0000)	2.5464E-12
101	5.0000	6	55	(1.0000,1.0000)	1.4587E-12
...
1998	5.0000	4	61	(1.0000,1.0000)	1.6926E-13
1999	5.0000	4	1053	(1.0000,1.0000)	1.3892E-13
2000	5.0000	4	58	(1.0000,1.0000)	1.4921E-13

由表 3、表 4 可知,初始点为 $(1,1), (-10,-10)$ 时,对不同罚因子的选择,算法均能达到全局收敛,故

根据不同的罚因子作图,如图 2 所示。



(a) 初始点为(1,1)



(b) 初始点为(-10,-10)

图2 不同初始罚因子对收敛速度的影响

结果说明与分析:

当初始点为(1,1)时,在罚因子 $\alpha \in [2,6]$ 时,内迭代次数集中在 100 次左右,在罚因子 $\alpha \in [7,30]$ 时,内迭代次数集中在 50 次左右,罚因子 $\alpha \in [31,200]$ 时,内迭代次数集中在 30 次左右,其中部分初始点迭代次数在 1050 次左右,随着罚因子的增加,内迭代次数减少,出现了很多不稳定的点,使内迭代次数很高,故在该初始点时,罚因子的取值范围在 $[7,30]$ 最好。

当初始点为 $(-10,-10)$ 时,在罚因子 $\alpha \in [2,5]$ 时,内迭代次数集中在 130 次左右,在罚因子 $\alpha \in [6,174]$ 时,内迭代次数集中在 80 次左右,罚因子 $\alpha \in [175,200]$ 时,内迭代次数集中在 50 次左右,其中部分初始点迭代次数在 1070 次左右,并且有少量点迭代次数在 2000 次以上,随着罚因子的增加,内迭代次数减少,出现的不稳定点更多,说明用较小的罚因子来避免不稳定的迭代点。

罚因子较小时,内迭代次数相对高一点,罚因子较大时,总体内迭代次数有小幅降低。但是出现了不稳定的点,故罚因子的最优取值为 $[7,30]$,若选取罚因子 $\alpha \in [2,5]$,收敛速度可接受。

(3) 罚因子修正系数 η 对收敛速度的影响

设置精度为 10^{-10} ,修正罚因子判别参数 $\theta=0.8$,选取初始罚因子 $\alpha=2$,初始点依然选择(1,1), $(-10,-10)$ 两个点。选取罚因子修正系数 η 从 2 依次增加到 10×10^5 ,计算结果见表 5、表 6。

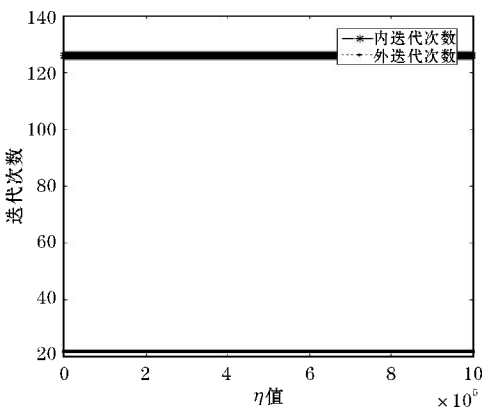
表 5 初始点为(1,1)时不同罚因子修正系数对收敛速度的影响

修正系数 η	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
2	5.0000	22	126	(1.0000,1.0000)	8.6938E-11
3	5.0000	22	126	(1.0000,1.0000)	8.6938E-11
4	5.0000	22	126	(1.0000,1.0000)	8.6938E-11
5	5.0000	22	126	(1.0000,1.0000)	8.6938E-11
...
999999	5.0000	22	126	(1.0000,1.0000)	8.6938E-11
1000000	5.0000	22	126	(1.0000,1.0000)	8.6938E-11

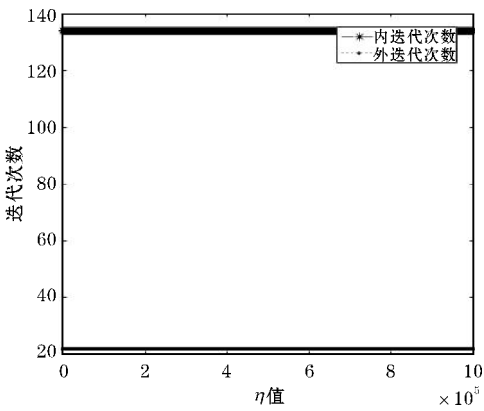
表 6 初始点为(-10,-10)时不同罚因子修正系数对收敛速度的影响

修正系数 η	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
2	5.0000	22	134	(1.0000,1.0000)	8.6664E-11
3	5.0000	22	134	(1.0000,1.0000)	8.6664E-11
4	5.0000	22	134	(1.0000,1.0000)	8.6664E-11
5	5.0000	22	134	(1.0000,1.0000)	8.6664E-11
...
999999	5.0000	22	134	(1.0000,1.0000)	8.6664E-11
1000000	5.0000	22	134	(1.0000,1.0000)	8.6664E-11

由表 5、表 6 可知,在初始点为(1,1),(-10,-10)时,对不同的罚因子修正系数,算法的内外迭代次数均相同,并且均能收敛到全局最优点,故根据不同罚因子修正系数作图,如图 3 所示。



(a) 初始点为(1,1)



(b) 初始点为(-10,-10)

图 3 不同罚因子修正系数对收敛速度的影响

结果说明与分析:
当初始点为(1,1)时,无论罚因子修正系数 η 如何变化,得到的最优解为 5,内迭代次数为 126 次,外迭代次数为 22 次。

当初始点为(-10,-10)时,无论罚因子修正系数 η 如何变化,得到的最优解为 5,内迭代次数为 134 次,外迭代次数为 22 次。

在罚因子修正系数 η 变化后,内外迭代次数均不发生变化。针对此情况,通过对算法方案的分析,实验选取的初始点并不一定在约束范围内,但是罚因子的修正系数改变后,内外迭代次数没有改变。这里有两种可能性,第一种是终止条件 $\beta_k < \theta \beta_{k-1}$ 一直成立,故罚因子在外循环中不改变;第二种是在第一次外循环中,无论初始点在哪里,迭代点始终落在约束条件内,导致约束条件失去约束能力,从而不影响迭代次数。故实验中,罚因子修正系数 η 不对实验结果产生影响。

实例 2 三元非线性混合约束问题

$$\min f(x) = -3x_1^2 - x_2^2 - 2x_3^2$$

s. t.
$$\begin{cases} x_1^2 + x_2^2 + x_3^2 = 3 \\ x_2 \geq x_1 \\ x_1 \geq 0 \end{cases}$$

(1) 不同初始点对收敛速度的影响

设置精度为 10^{-10} ,修正罚因子判别参数 $\theta=0.8$,初始罚因子修正系数 $\eta=2$,选取初始罚因子 α 值分别为 10 与 30,将初始点选择从(-200,-200,-200)依次增加到(200,200,200),计算结果见表 7、表 8。

表 7 罚因子为 10 时不同初始点对迭代次数的影响

初始点 编号	初始点 $x^{(0)}$	最小值 f	迭代次数 k	内迭代次 数 ink	计算结果 x	终止条件 β
-200	(-200,-200,-200)	NAN	1	1000	(NAN,NAN,NAN)	NAN
...
-30	(-30,-30,-30)	NAN	1	1000	(NAN,NAN,NAN)	NAN
-29	(-29,-29,-29)	-6.0000	15	193	(0.1536E-10,0.3883E-10,-1.7321)	2.8170E-11
-28	(-28,-28,-28)	-6.0000	15	202	(0.1446E-10,0.3839E-10,-1.7321)	2.7987E-11
...
-1	(-1,-1,-1)	-6.0000	15	184	(0.1726E-10,0.4075E-10,-1.7321)	2.9171E-11
0	(0,0,0)	-5.1379	3	1031	(0.1034E-10,0.13896E-10,0)	6.9131E-12
1	(1,1,1)	-6.0000	15	328	(0.1446E-10,0.3839E-10,1.7321)	2.7981E-11
...
28	(28,28,28)	-6.0000	15	205	(0.1446E-10,0.3839E-10,1.7321)	2.7987E-11
29	(29,29,29)	-6.0000	15	212	(0.1726E-10,0.4075E-10,1.7321)	2.9171E-11
30	(30,30,30)	NAN	1	1000	(NAN,NAN,NAN)	NAN
...
200	(200,200,200)	NAN	1	1000	(NAN,NAN,NAN)	NAN

表 8 罚因子为 30 时不同初始点对迭代次数的影响

初始点 编号	初始点 $x^{(0)}$	最小值 f	迭代次数 k	内迭代 次数 ink	计算结果 x	终止条件 β
-200	(-200,-200,-200)	NAN	1	1000	(NAN,NAN,NAN)	NAN
...
-18	(-18,-18,-18)	NAN	1	1000	(NAN,NAN,NAN)	NAN
-17	(-17,-17,-17)	-6.0000	9	119	(0.2051E-10,0.4867E-10,-1.7321)	3.4842E-11
-16	(-16,-16,-16)	-6.0000	9	114	(0.1978E-10,0.4727E-10,-1.7321)	3.3867E-11
...
-1	(-1,-1,-1)	-6.0000	9	127	(0.1976E-10,0.4726E-10,-1.7321)	3.3870E-11
0	(0,0,0)	-4.8071	3	79	(0.0950E-10,0.1447E-10,0)	1.2496E-12
1	(1,1,1)	-6.0000	9	132	(0.1692E-10,0.4570E-10,1.7321)	3.3388E-11
...
16	(16,16,16)	-6.0000	9	114	(0.1978E-10,0.4728E-10,1.7321)	3.3875E-11
17	(17,17,17)	-6.0000	9	116	(0.1978E-10,0.4728E-10,1.7321)	3.3877E-11
18	(18,18,18)	NAN	1	1000	(NAN,NAN,NAN)	NAN
...
200	(200,200,200)	NAN	1	1000	(NAN,NAN,NAN)	NAN

由表 7、表 8 可知,该优化问题的最优解为 $(0,0,\sqrt{3})$ 和 $(0,0,-\sqrt{3})$,最小值为-6。当罚因子 $\alpha=10$,初始点编号 $p \in [-200,-30]$ 与 $p \in [30,200]$ 时,算法迭代次数超过了最大限制,无法收敛到最小点。初始点编号 $p=0$,函数陷入局部收敛;当罚因子 $\alpha=30$,初始

点编号 $p \in [-200,-18]$ 与 $p \in [18,200]$ 时,算法迭代次数超过了最大限制,无法收敛到最小点,初始点编号 $p=0$,函数陷入局部收敛。故去掉初始点中不能计算的点后再作图,如图 4 所示。

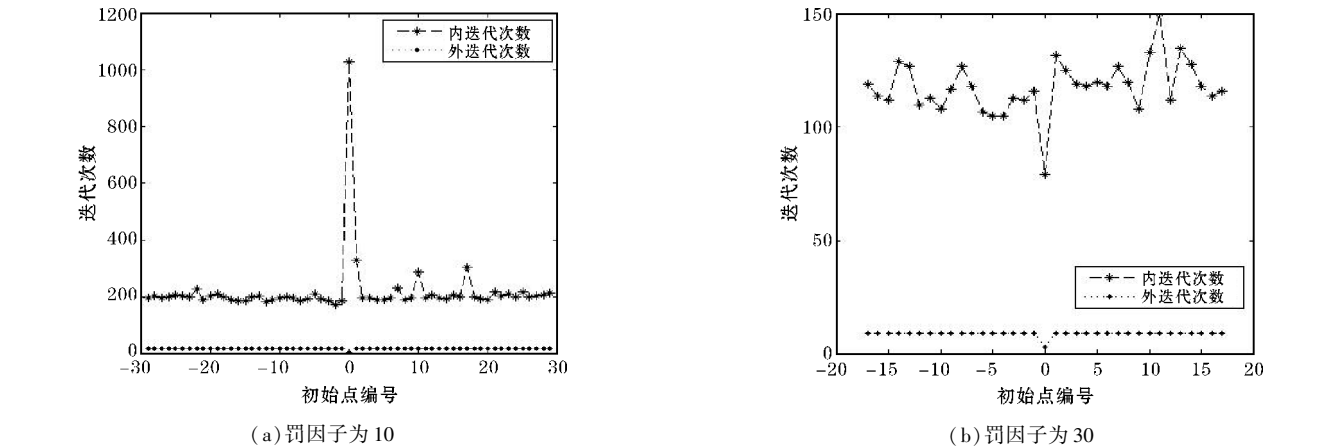


图 4 不同初始点对迭代次数的影响

结果说明与分析:
当罚因子 $\alpha=10$ 时,在初始点编号 $p \in [-29,29]$ 时,内迭代次数集中在 200 次左右,其中少量初始点迭代次数在 300 次左右,仅仅在初始点为 $(0,0,0)$ 时,内迭代次数超过 1000 次,并且陷入局部收敛。对于初始点的选取,应尽可能靠近最优点,在初始点附近的迭代次数相对集中,但是在离初始点很近的 $(0,0,0)$ 处陷入局部收敛,说明原点作为一个特殊点,不能带入计算。
当罚因子 $\alpha=30$ 时,在初始点编号 $p \in [-17,17]$ 时,内迭代次数集中在 115 次左右,最高为 152 次,仅仅在初始点为 $(0,0,0)$ 时,内迭代次数为 79 次,并且

陷入了局部收敛,与罚因子 $\alpha=10$ 得出同样的结论。
当选择较大的罚因子时,内迭代次数和外迭代次数都有所减少。而初始点的选取不能距离最优点太远,应尽量靠近最优点,以达到最佳收敛速度。
(2) 不同罚因子对收敛速度的影响
设置精度为 10^{-10} ,修正罚因子判别参数 $\theta=0.8$,选取初始罚因子修正系数 $\eta=2$,由对初始点选取的研究可知初始点编号应选取 $p \in [-17,17]$ 最合适,故从中选取 2 个初始点分别为 $(1,1,1)$, $(5,5,5)$,将初始罚因子 α 选择从 1 依次增加到 200,计算结果见表 9、表 10。

表 9 初始点为(1,1,1)时不同初始罚因子对迭代次数的影响

罚因子 α	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
1	-6.0000	33	1564	(-0.4021E-10,0.4075E-10,1.7321)	8.2209E-11
2	-6.0000	28	314	(-0.4021E-10,0.4075E-10,1.7321)	9.0403E-11
3	-6.0000	37	357	(0.6603E-10,0.4822E-10,1.7321)	6.8587E-11
...
39	-6.0000	8	117	(0.2657E-10,0.6950E-10,1.7321)	2.6568E-11
40	-6.0000	8	856	(0.2211E-10,0.5794E-10,-1.7321)	2.2108E-11
41	-6.0000	8	114	(0.1835E-10,0.4802E-10,-1.7321)	1.8350E-11
...
111	-6.0000	6	124	(0.1129E-10,0.3006E-10,1.7321)	2.1882E-11
112	-6.0000	6	126	(0.1071E-10,0.2849E-10,1.7321)	2.0762E-11
113	NAN	1	1000	(NAN,NAN,NAN)	NAN
...
200	NAN	1	1000	(NAN,NAN,NAN)	NAN

表 10 初始点为(5,5,5)时不同初始罚因子对迭代次数的影响

罚因子 α	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
1	-6.0000	34	561	(-0.2503E-10,0.9887E-10,1.7321)	5.1443E-11
2	-6.0000	28	346	(-0.4021E-10,0.4075E-10,1.7321)	9.0403E-11
3	-6.0000	37	357	(0.6603E-10,0.4822E-10,1.7321)	6.8588E-11
...
69	-6.0000	7	101	(0.8225E-10,0.2164E-10,1.7321)	1.5741E-11
70	NAN	1	1000	(NAN,NAN,NAN)	NAN
71	-6.0000	7	131	(0.6823E-10,0.1782E-10,1.7321)	1.2947E-11
...
106	-6.0000	6	123	(0.1505E-10,0.3954E-10,1.7321)	2.8750E-11
107	NAN	1	1000	(NAN,NAN,NAN)	NAN
...
200	NAN	1	1000	(NAN,NAN,NAN)	NAN

由表 9、表 10 可知,当初始点为(1,1,1)时,当 $\alpha \in [113, 200]$ 时,算法迭代次数超过了最大限制,无法收敛到最小点。当初始点为(5,5,5)时,当 $\alpha \in [107, 200]$ 时,算法迭代次数超过了最大限制,无法收敛到最小点。故对罚因子的选择,去除无法计算的值作图,如图 5 所示。

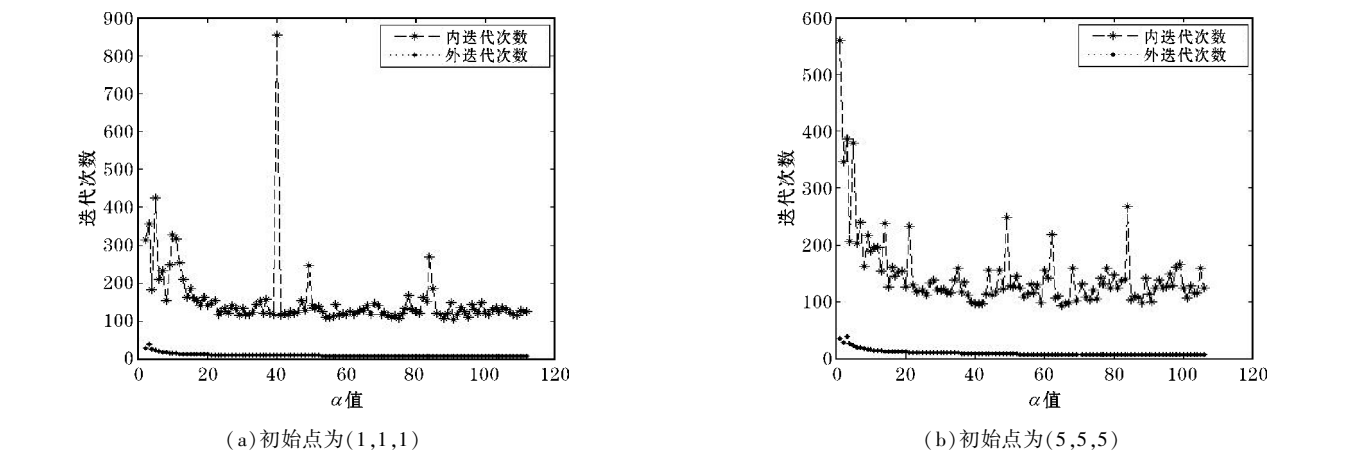


图 5 不同初始罚因子对迭代次数的影响

结果说明与分析:

当初始点为(1,1,1)时, $\alpha \in [113,200]$ 时,算法内迭代次数超过了最大限制,无法收敛到最小点; $\alpha \in [2,17]$ 时,内迭代次数在 150~450 次波动; $\alpha \in [18,112]$ 时,内迭代次数集中在 120 次左右,其中仅有一个初始点迭代次数超过 800 次。可知当初始罚因子选取过小时,由于惩罚力度不够,使算法迭代次数增加,而初始罚因子过大,则增加了问题的复杂度,导致程序无法运行,故在该初始点时,罚因子取 $\alpha \in [18,112]$ 最有效。

当初始点为(5,5,5)时, $\alpha \in [107,200]$ 时,算法迭

代次数超过了最大限制,无法收敛到最小点; $\alpha \in [2,17]$ 时,内迭代次数集中在 180~550 次波动; $\alpha \in [18,106]$ 时,内迭代次数集中在 130 次左右,其中部分初始点迭代次数超过 200 次。该结果得到与初始点为(1,1,1)一样的结论。罚因子取 $\alpha \in [18,106]$ 最有效。

(3) 罚因子修正系数 η 对收敛速度的影响

设置精度为 10^{-10} ,修正罚因子判别参数 $\theta=0.8$,选取初始罚因子 $\alpha=2$,与之前的选择思路一致,初始点选择(1,1,1),(5,5,5)两个点。选取罚因子修正系数 η 从 2 依次增加到 200,计算结果见表 11 与表 12。

表 11 初始点为(1,1,1)时不同罚因子修正系数对迭代次数的影响

修正系数 η	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
2	-6.0000	28	314	(-0.4021E-10,0.4075E-10,1.7321)	9.0403E-11
3	-6.0000	22	262	(0.3086E-10,0.8055E-10,1.7321)	5.8508E-11
4	-6.0000	19	253	(0.3127E-10,0.8143E-10,1.7321)	5.9108E-11
5	-6.0000	17	309	(0.3809E-10,0.9915E-10,1.7321)	7.1964E-11
...
16	-6.0000	12	199	(0.6414E-10,0.1678E-10,1.7321)	1.2194E-11
17	NAN	4	1048	(NAN,NAN,NAN)	NAN
...
200	NAN	4	1048	(NAN,NAN,NAN)	NAN

表 12 初始点为(5,5,5)时不同罚因子修正系数对迭代次数的影响

修正系数 η	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
2	-6.0000	28	411	(-0.4021E-10,0.4075e-10,1.7321)	9.0403E-11
3	-6.0000	22	358	(0.3086E-10,0.8055e-10,1.7321)	5.8508E-11
4	-6.0000	19	357	(0.3059E-10,0.8106e-10,1.7321)	5.9055E-11
5	-6.0000	17	382	(0.3742E-10,0.9830e-10,1.7321)	7.1946E-11
...
16	-6.0000	12	295	(0.6414E-10,0.1678E-10,1.7321)	1.2194E-11
17	NAN	4	1048	(NAN,NAN,NAN)	NAN
...
200	NAN	4	1048	(NAN,NAN,NAN)	NAN

由表 11、表 12 可知,当初始点为(1,1,1)时,初始点编号 $p \in [17,200]$ 时,算法迭代次数超过了最大限制,无法收敛到最小点;当初始点为(5,5,5)时,初始

点编号 $p \in [17,200]$ 时,算法迭代次数超过了最大限制,无法收敛到最小点。故将初始点选取去除陷入局部收敛的点作图,如图 6 所示。

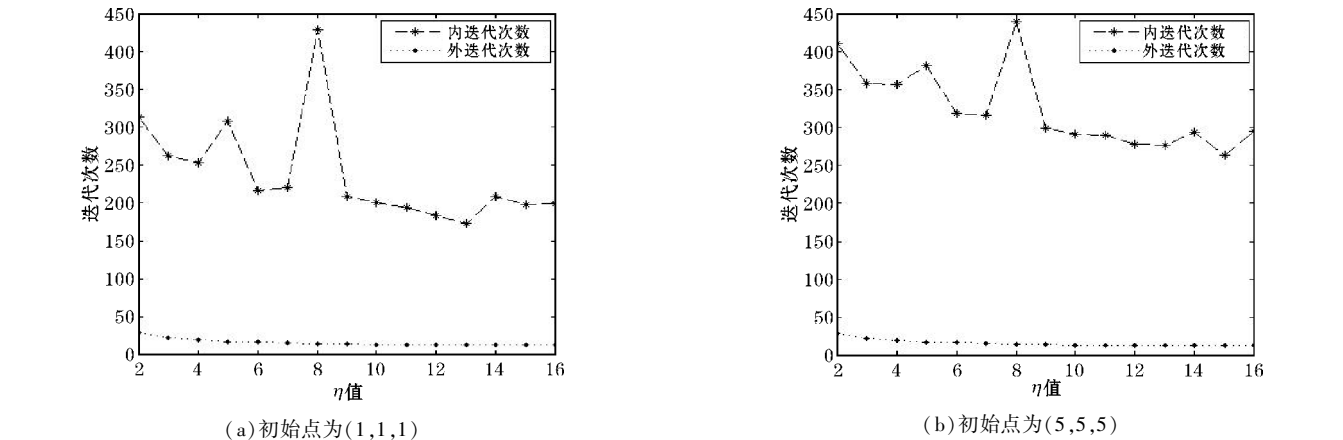


图 6 不同罚因子修正系数对迭代次数的影响

当初始点为(1,1,1)时,罚因子修正系数 $\eta \in [2, 9]$ 时,内迭代次数在 200 次到 450 次之间震荡,罚因子修正系数 $\eta \in [10, 16]$ 时,内迭代次数稳定在 200 次左右。可知罚因子修正系数取值不应过大,否则算法不能得到最小值,并且从图 6 可以看出有小幅的下降趋势,但是曲线还是相对平缓,故本例中罚因子修正系数的最优取值范围为[2,9]。

当初始点为(5,5,5)时,罚因子修正系数 $\eta \in [2, 9]$ 时,内迭代次数在 280 次到 450 次之间震荡,罚因子修正系数 $\eta \in [10, 16]$ 时,内迭代次数稳定在 290 次左右。该初始点得到的曲线与初始点为(1,1,1)时大致

相同,故得到同样的结论,最优取值范围为[2,9]。

实例 3 考虑含有指数函数的二元非线性混合约束的优化问题

$$\begin{aligned} \min \quad & 3e^{x_1}-2x_1x_2+2x_2^2 \\ \text{s. t.} \quad & \begin{cases} x_1^2+x_2^2=9 \\ 3x_1+2x_2\leq 6 \end{cases} \end{aligned}$$

(1)不同初始点对收敛速度的影响

设置精度为 10^{-10} ,修正罚因子判别参数 $\theta=0.8$,初始罚因子修正系数 $\eta=2$,选取初始罚因子 α 值分别为 10 与 30,将初始点选择从(−200,−200)依次增加到(200,200),计算结果见表 13 与表 14。

表 13 罚因子为 10 时不同初始点对迭代次数的影响

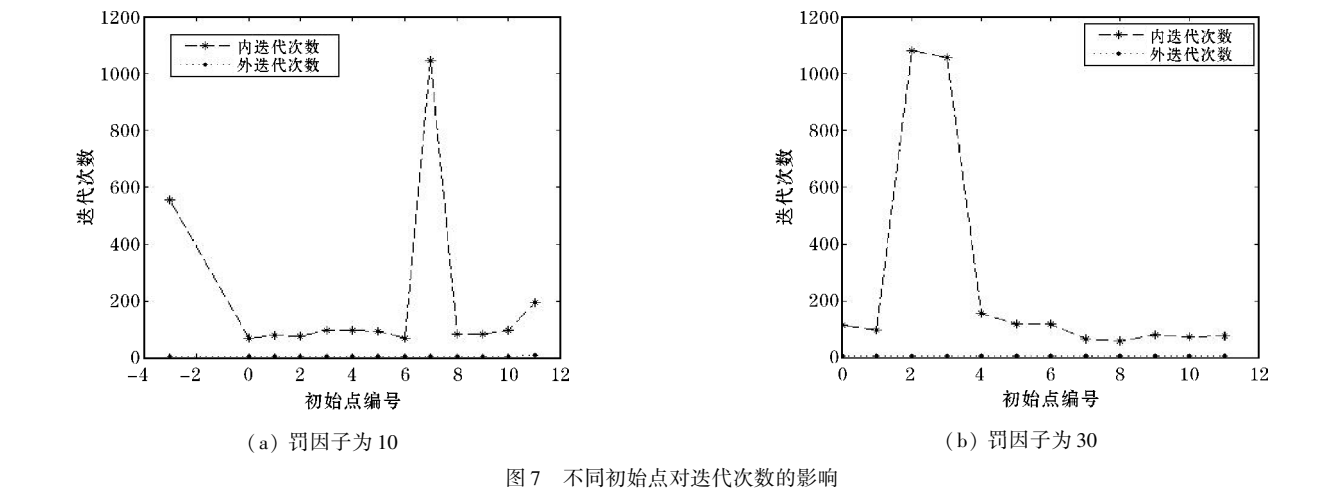
初始点编号	初始点 $x^{(0)}$	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
−200	(−200,−200)	NAN	1	1000	(NAN,NAN,NAN)	NAN
...
−4	(−4,−4)	NAN	1	1000	(NAN,NAN,NAN)	NAN
−3	(−3,−3)	21.0000	6	556	(0.4545E−14,3.0000)	6.7842E−12
−2	(−2,−2)	21.0000	1	1000	(NAN,NAN,NAN)	NAN
−1	(−1,−1)	21.0000	1	1000	(NAN,NAN,NAN)	NAN
0	(0,0)	21.0000	6	67	(0.5183E−14,3.0000)	6.7687E−12
1	(1,1)	21.0000	6	78	(0.6189E−14,3.0000)	6.7630E−12
...
11	(11,11)	21.0000	6	196	(0.8782E−14,3.0000)	2.9068E−12
12	(12,12)	NAN	1	1000	(NAN,NAN,NAN)	NAN
...
200	(200,200)	NAN	1	1000	(NAN,NAN,NAN)	NAN

表 14 罚因子为 30 时不同初始点对迭代次数的影响

初始点编号	初始点 $x^{(0)}$	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
−200	(−200,−200)	NAN	1	1000	(NAN,NAN)	NAN
...
−1	(−1,−1)	NAN	1	1000	(NAN,NAN)	NAN
0	(0,0)	21.0000	5	115	(0.1440E−13,3.0000)	4.1690E−12
1	(1,1)	21.0000	5	98	(0.1440E−13,3.0000)	4.1693E−12
2	(2,2)	21.0000	5	1081	(0.1439E−13,3.0000)	4.1668E−12
...
9	(9,9)	21.0000	5	79	(0.1440E−13,3.0000)	4.1690E−12
10	(10,10)	21.0000	5	72	(0.1439E−13,3.0000)	4.1683E−12
11	(11,11)	21.0000	5	74	(0.1440E−13,3.0000)	4.1698E−12
12	(12,12)	NAN	1	1000	(NAN,NAN)	NAN
...
200	(200,200)	NAN	1	1000	(NAN,NAN)	NAN

由表 13、14 可知,当罚因子 $\sigma=10$ 时,初始点编号 $p \in [-200,-4]$ 或 $p \in [12,200]$ 时,算法迭代次数超过了最大限制,无法收敛到最小点;当罚因子 $\sigma=30$ 时,初始

点编号 $p \in [-200,-1]$ 且 $p \in [12,200]$ 时,算法迭代次数超过了最大限制,无法收敛到最小点。故将初始点选取范围去除陷入局部收敛的点作图,如图 7 所示。



结果说明与分析:
当罚因子 $\alpha=10$ 时,初始点编号 $p \in [-3,11]$ 时,内迭代次数集中在 70 ~ 100 次,其中只有 3 个点的数值不在这个范围内,并且最反常的点迭代次数超过 1000 次。可知该例对初始点的要求较高,可能是受优化函数复杂度的影响,故该例最优点的初始点编号为 $[0,6]$ 。
当罚因子 $\alpha=30$ 时,初始点编号 $p \in [0,11]$ 时,内迭代次数集中在 70 ~ 160 次,其中只有 2 个点的数值不在这个范围内,并且这两个点迭代次数超过了 1000

次。在可行域内除去反常点速度大致相同,故该例选择的初始点应大于最优点,最优范围为 $[4,11]$ 。
(2) 不同罚因子对收敛速度的影响
设置精度为 10^{-10} ,修正罚因子判别参数 $\theta=0.8$,选取初始罚因子修正系数 $\eta=2$,由对初始点选取的研究可知初始点编号应选取 $p \in [0,11]$ 最合适,故从中选取 2 个初始点分别为 $(1,1), (-10,-10)$,将初始罚因子选择从 2 依次增加到 2000,计算结果见表 15、表 16。

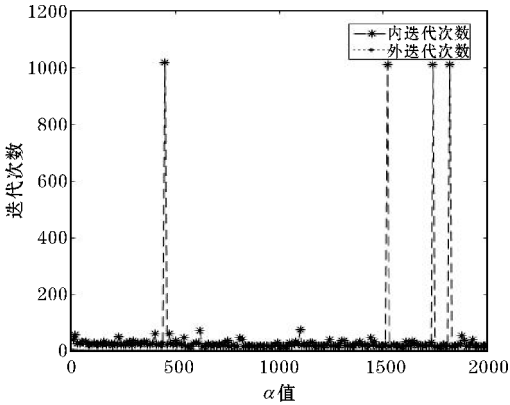
表 15 初始点为(1,1)时不同罚因子对迭代次数的影响

罚因子 α	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
1	21.0000	18	104	(7.8709E-11,3.0000)	2.9764E-11
2	21.0000	11	59	(-0.63209E-11,3.0000)	2.1304E-11
3	21.0000	9	50	(4.2675E-11,3.0000)	1.4038E-11
4	21.0000	8	50	(-0.12898E-11,3.0000)	3.9869E-11
...
10	21.0000	7	44	(6.1239E-11,3.0000)	2.0145E-12
11	21.0000	6	61	(-5.5829E-11,3.0000)	1.8210E-12
...
450	21.0000	4	1016	(-2.3666E-14,3.0000)	7.6512E-14
451	21.0000	4	25	(-1.7997E-14,3.0000)	6.8459E-14
...
1999	21.0000	3	21	(7.1880E-11,3.0000)	2.1006E-12
2000	21.0000	3	21	(7.0773E-11,3.0000)	2.0683E-12

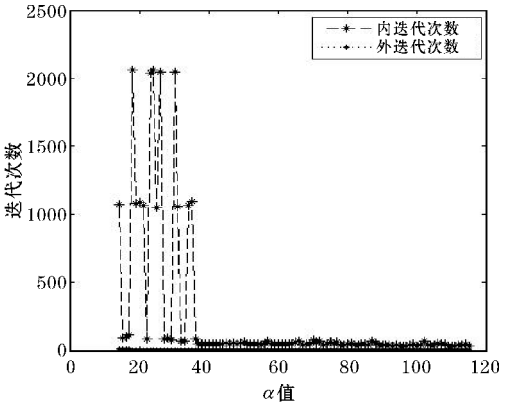
表 16 初始点为(-10,-10)时不同罚因子对迭代次数的影响

罚因子 α	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
1	21.0000	18	127	(0.7860E-12,3.0000)	3.0525E-11
2	50.1977	10	99	(0.2494E-12,3.0000)	1.2969E-11
3	52.5016	7	93	(0.2594E-12,3.0000)	8.9038E-11
4	53.6803	6	78	(0.2642E-12,3.0000)	9.5923E-14
...
13	55.9964	6	3019	(0.2734E-12,3.0000)	3.1049E-13
14	21.0000	8	1071	(0.2734E-13,3.0000)	2.4362E-12
15	21.0000	8	88	(0.2734E-13,3.0000)	1.6171E-13
...
115	21.0000	4	35	(-0.5781E-12,3.0000)	1.7455E-11
116	NAN	1	1000	(NAN,NAN)	NAN
...
2000	NAN	1	1000	(NAN,NAN)	NAN

由表 15、表 16 可知,在初始点为(1,1)时,对不同罚因子的选择,算法均能达到全局收敛,在初始点(-10,-10)时,去掉没有收敛到最小值的点作图,如图 8 所示。



(a) 初始点为(1,1)



(b) 初始点为(-10,-10)

图 8 不同罚因子对迭代次数的影响

结果说明与分析:

当初始点为(1,1)时,罚因子 $\alpha \in [1,2000]$ 时,大部分内迭代次数集中在 50 次以下,部分点的迭代次数在 100 ~ 150 次,有 4 个点迭代次数超过 1000 次。罚因子对收敛速度的影响较小,只是在罚因子较小时,迭代次数稍微多一点,故最优取值范围为[4,1587]。

当初始点为(-10,-10)时,罚因子 $\alpha \in [14,37]$ 时,内迭代次数在 100 次、1070 次、2100 次之间震荡,罚因子 $\alpha \in [38,115]$ 时,内迭代次数集中在 50 次左右。罚

因子较小时迭代次数波动较大,但是随着罚因子增加,迭代次数相对比较稳定,并且与之前例子一样,罚因子取值不能过大,故最优取值范围为[38,115]。

(3) 罚因子修正系数 η 对收敛速度的影响

设置精度为 10^{-10} ,修正罚因子判别参数 $\theta=0.8$,选取初始罚因子 $\alpha=2$,初始点选择(1,1),(9,9)两个点。由于罚因子修正系数 η 需要大于 1,故选取罚因子修正系数 η 从 2 依次增加到 10×10^5 ,计算结果见表 17 与 18。

表 17 初始点为(1,1)时不同罚因子修正系数对收敛速度的影响

修正系数 η	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
2	21.0000	11	59	(-6.3209E-12,3.0000)	2.1304E-11
3	21.0000	11	59	(-6.3209E-12,3.0000)	2.1304E-11
4	21.0000	11	59	(-6.3209E-12,3.0000)	2.1304E-11
...
999999	21.0000	11	59	(-6.3209E-12,3.0000)	2.1304E-11
1000000	21.0000	11	59	(-6.3209E-12,3.0000)	2.1304E-11

表 18 初始点为(9,9)时不同罚因子修正系数对收敛速度的影响

修正系数 η	最小值 f	迭代次数 k	内迭代次数 ink	计算结果 x	终止条件 β
2	21.0000	7	2040	(0.1749E-11,3.0000)	5.7989E-11
3	21.0000	7	2042	(0.3425E-12,3.0000)	1.1382E-11
4	21.0000	7	2042	(0.1084E-12,3.0000)	3.5911E-12
5	21.0000	7	2039	(0.4592E-13,3.0000)	1.4987E-12
6	NAN	10	4126	(NAN,NAN)	NAN
7	21.0000	7	2041	(0.1236E-13,3.0000)	3.9571E-13
...
15	21.0000	6	2038	(-0.8487E-12,3.0000)	2.7600E-11
16	NAN	7	4105	(NAN,NAN)	NAN
...
29	NAN	7	5045	(NAN,NAN)	NAN
30	NAN	7	4174	(NAN,NAN)	NAN
31	21.0000	6	2201	(-0.9615E-13,3.0000)	3.1327E-12
32	21.0000	6	2201	(-0.8738E-13,3.0000)	2.8446E-12
...
90	21.0000	6	2048	(-0.3928E-14,3.0000)	1.2977E-13
91	NAN	5	4019	(NAN,NAN)	NAN
...
1000000	NAN	5	4019	(NAN,NAN)	NAN

由表 17、表 18 可知,在初始点为(1,1)时,对不同罚因子修正系数的选择,算法的内外迭代次数均相同,并且均能收敛到全局最优点,在初始点为(9,9)时,去掉没有收敛到最小值的点作图,如图 9 所示。

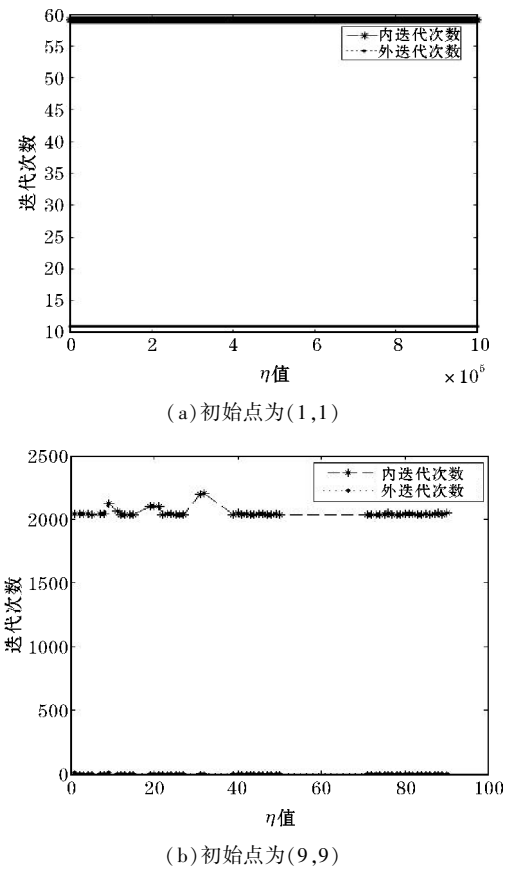


图 9 不同罚因子修正系数对收敛速度的影响

结果说明与分析:

当初始点为(1,1)时,无论罚因子修正系数 η 如何变化,得到的最优解为 21,内迭代次数为 46 次,外迭代次数为 6 次。出现了与之前同样的情况,说明在这个初始点下,罚因子没有进行惩罚,原因与例 1 一样,故罚因子修正系数可任取。

当初始点为(9,9)时,随着罚因子修正系数 η 的变化,在迭代中不间断有超出迭代次数,无法收敛到最小值的点,但是从能够迭代到最小值的点来看,迭代次数维持在 2000 ~ 2050 次。仅有部分点的迭代次数超过这个范围,但是迭代次数不超过 2500 次。罚因子修正系数的变化在部分问题中也是相当重要的。随着罚因子修正系数的变化,随时会出现无法收敛的现象,故该系数的取值应在迭代成功率相对较高的分段,并且由于不同优化问题的随机性,最优取值范围在[2, 15]。同时,发现初始点的选取对罚因子修正系数的选取影响较大。

4 结束语

对于广义拉格朗日乘子法求解最优值问题,研究了 3 种不同的混合约束优化问题,对初始点、初始罚因子和罚因子修正系数取值进行对比分析。从实例中不难看出,初始点的选取对算法收敛速度和收敛效果影响较大,如果初始点的选取距离最优点过远,可能导致算法陷入局部收敛或直接导致算法失败,因此初始点的取值应尽量在最优点附近。对于罚因子的选取,由

实例1可知罚因子可能在部分优化问题中对函数收敛速度不影响,这是由于迭代点在约束条件的可行域内,惩罚没有产生效果。但在实例2与实例3中,罚因子过小会增加迭代次数,过大会导致算法收敛失败,故罚因子的最优取值范围建议在 $[5, 15]$ 。对于罚因子修正系数的选取,在实例1中发现,罚因子修正系数不影响收敛效果,因此可以任取;在实例2和实例3中,如果罚因子修正系数过小可能导致迭代速度较慢,收敛效果不佳。当适当增加修正系数的值时,收敛速度有所增加,而罚因子修正系数过大时,算法收敛失败,无法收敛到最小值,故罚因子的最优取值范围建议在 $[2, 15]$ 。数值实验中,初始点的选取具有一定的主观性,且初始点分量为整数,在今后的研究中可以考虑选取更加客观的初始点,并且可以尝试选取分量为分数或无理数的初始点。

参考文献:

- [1] Hestenes M R. Multiplier and gradient methods [J]. Journal of Optimization Theory and Applications, 1969, 4(5): 303–320.
- [2] Rockafellar R T. The multiplier method of Hestenes and Powell applied to convex programming [J]. Journal of Optimization Theory and Applications, 1973, 12(6): 555–562.
- [3] Rockafellar R T. A dual approach to solving nonlinear programming problems by unconstrained optimization [J]. Mathematical Programming, 1973, 5(1): 354–373.
- [4] Davidon W C. Optimally conditioned optimization algorithms without line searches [J]. Mathematical Programming, 1975, 9(1): 1–30.
- [5] Broyden C G. The convergence of a class of double-rank minimization algorithms [J]. J. Institute of Mathematics and Its Applications, 1970, 6(1): 76–90.
- [6] Williams T M. Practical Methods of Optimization. Unconstrained Optimization [J]. Journal of the Operational Research Society, 1981, 32(5): 417.
- [7] 刘兴高, 胡云卿. 应用最优化方法及 MATLAB 实现 [M]. 北京: 科学出版社, 2014.
- [8] Fiacco A V, McCormick, Garth P. Nonlinear Programming: Sequential Unconstrained Minimization Techniques [J]. Philadelphia: SIAM, 1990.
- [9] 宋菲, 吴泽忠. 外罚函数法与广义 Lagrange 乘子法的比较研究 [J]. 成都信息工程大学学报, 2017(6).
- [10] Andreani R, Martínez, J. M, Santos L T, et al. On the behaviour of constrained optimization methods when Lagrange multipliers do not exist [J]. Optimization Methods and Software, 2014, 29(3): 646–657.
- [11] 陈宝林. 最优化理论与算法 [M]. 2 版. 北京: 清华大学出版社, 2005.
- [12] 袁亚湘, 孙文瑜. 最优化理论与方法 [M]. 上海: 科学出版社, 2001.

Research on Generalized Lagrange Multiplier Method based on BFGS Algorithm

XIONG Qian, WU Zezhong

(College of Applied Mathematics, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: The generalized Lagrange multiplier method is an important method to solve the problem of constrained optimization. Based on BFGS algorithm, this paper uses MATLAB tool to study the influence of initial point selection, penalty factor change and penalty factor change correction coefficient. The results show that: (1) For the selection of the initial point, the value should be selected near the best point, and the convergence effect can be good. (2) The penalty factor too small or too large will make the algorithm difficult to solve. If the penalty factor is too small, a large amount of search time will cost in the non-feasible domain, which will increase the number of iterations. On the other hand, if the penalty factor is too large, the algorithm will be difficult to be pushed into the feasible domain, which will cause the algorithm fail to converge. (3) With the change of the penalty factor correction coefficient, there is a phenomenon of fail to converge at any time. Therefore, the value of the coefficient is supposed to select during the segment of iteration success rate that is relatively high.

Keywords: applied mathematics; optimization theory; constrained optimization; generalized Lagrange multiplier method; penalty factor; correction factor