文章编号：2096-1618(2019)02-0209-07

基于 Armijo 搜索步长的几种共轭梯度法的分析对比

黄 飞， 吴泽忠
（成都信息工程大学应用数学学院，四川 成都 610225）

摘要：共轭梯度法是解决无约束优化问题的一种重要方法，使用不精确的 Armijo 搜索步长的方法，利用 MATLAB 工具对 FR 共轭梯度法、PRP 共轭梯度法、HS 共轭梯度法 3 种方式的收敛效果进行对比。结果表明：在低次函数里使用 FR 共轭梯度法效果较好，在高次函数里使用 PRP 共轭梯度法或 HS 共轭梯度法的收敛效果较好，并且在函数波动较大时，初值的选择应尽量靠近收敛点，才能有不错的收敛效果。

关 键 词：应用数学；最优化理论；FR 共轭梯度法；PRP 共轭梯度法；HS 共轭梯度法；无约束最优化；Armijo 搜索

中图分类号：O221.2 **文献标志码**：A

doi：10.16836/j.cnki.jcuit.2019.02.0017

0 引言

共轭梯度法期初是由 Hesteness 和 Stiefel 在 20 世纪中叶，为了求解线性方程组[1]而提出来的，随后，在各国科学家的努力下，将这种方法推广到求解无约束最优化问题[2]，共轭梯度法逐渐成为一种重要的最优化方法，共轭梯度法是一种介于最速下降法与牛顿法之间的一种求解无约束最优化的方法，这种方法仅需要目标函数以及梯度函数值[3]，利用已知点处的梯度构造一组共轭向量，然后利用这组方向基于 Armijo 搜索方法[4]进行搜索，便可以求出目标函数的极小值。并且具有收敛较快的优点。由于共轭梯度法有上述优点，共轭梯度法在信号处理、遥感技术、航空航天、大型工业等领域都有较为重要的作用，所以共轭梯度法的研究也成了很多学者热衷的研究方向，针对不同的共轭梯度法，取不同的初值以及各种类型的函数，来分析各种共轭梯度法的优缺点，以解决在实际问题中，求解无约束优化问题的时候，能选择适当的共轭梯度法进行求解，达到提高计算效率的目的。

1 预备知识

因为共轭梯度法可以用来计算一个函数的极小点，所以考虑如下形式方程的最小值：

$$\min_{x\in R^n} f(x) \tag{1}$$

$f(x)$ 是一个连续可微的函数，其共轭梯度法的迭代公式为

$$x^{(k+1)} = x^{(k)} + \lambda_k d^{(k)} \tag{2}$$

$$d^{(k)} = \begin{cases} -g_k, & k=1 \\ -g_k + \beta_k d^{(k-1)}, & k \geq 2 \end{cases} \tag{3}$$

其中，$d^{(k)}$ 是搜索方向；g_k 为 $x^{(k)}$ 点处的梯度；β_k 为计算因子，在不同的共轭梯度法中，计算因子各不相同；λ_k 为步长因子。

在这个迭代公式中，$x^{(k)}$ 会沿着搜索方向 $d^{(k)}$ 搜索，并且希望 $x^{(k)}$ 在这个搜索方向上能搜索到最小值，所以，步长因子 λ_k 满足下面的条件：

$$\lambda_k = \arg\min_{\lambda\geq 0}\{f(x^{(k)} + \lambda d_k)\} \tag{4}$$

经过各国数学家的努力，共轭梯度法的发展也越来越深入，出现了多种不同方法的共轭梯度法，例如 FR 共轭梯度法、PRP 共轭梯度法、HS 共轭梯度法、DY 共轭梯度法、LS 共轭梯度法、CD 共轭梯度法等[2,5-9]，文中对前 3 种常见的共轭梯度法进行对比。

1.1 FR 共轭梯度法

共轭梯度法最初由 Hesteness 和 Stiefel 于 1952 年提出，最初是为了求解线性方程组，后来 Fletcher 等首次提出了一种非线性的共轭梯度法，也就是现在最常见的 FR 共轭梯度法，FR 法中的因子 β_k 表达式为

$$\beta_k = \frac{\| g_{k+1} \|^2}{\| g_k \|^2},\ k \geq 1, g_k \neq 0 \tag{5}$$

对于 FR 方法的收敛性而言，如果选择这种步长规则，可能在计算过程中不能产生很好的数值效果，因为在这种迭代的过程中，Powell 在精确的线性搜索基

收稿日期：2018-09-13
基金项目：国家自然科学基金资助项目（71672013）；四川省软件科学研究计划资助项目（2014ZR0016）；四川省社科重点研究基地资助项目（Xq14B06）

础上分析得出,如果出现了一个较小的步长,那么可能导致迭代过程中出现很多非常小的步长,影响迭代的次数[10]。

1.2　PRP 共轭梯度法

1969 年,PRP 共轭梯度法首次由 Polyak[5] 提出,PRP 法在实际使用时是数值表现很好的一种共轭梯度法之一,PRP 法中的因子 β_k 表达式为

$$\beta_k = \frac{g_k^{\mathrm{T}}(g_k - g_{k-1})}{\| g_{k-1} \|^2},\ k \geq 1, g_k \neq 0 \tag{6}$$

PRP 法有效地避免了 FP 法中可能出现连续小步长的不足,但是对于 PRP 法是否比 FP 法收敛更快,不同的学者持不同的态度,有人认为 PRP 比 FP 法更好,而有人也认为两种算法各有各自的优点[3]。

1.3　HS 共轭梯度法

HS 共轭梯度法由 Hestenes 和 Stiefel 提出[6],HS 法有着和 PRP 算法类似的性质,并且数值效果也较为不错,HS 法中的因子 β_k 表达式为

$$\beta_k = \frac{g_k^{\mathrm{T}}(g_k - g_{k-1})}{d_{k-1}^{\mathrm{T}}(g_k - g_{k-1})},\quad k \geq 1, g_k \neq 0 \tag{7}$$

HS 算法不依赖于任何步长规则,对于不同的步长规则,都有 $d_{k+1}^{\mathrm{T}}(g_{k+1} - g_k) = 0$, 由于 HS 算法有很好的性质,很多学者对其进行了后续的修改,逐步对该算法进行了更好的改进。

1.4　Armijo 搜索步长

在共轭梯度法中,使用精确一维搜索步长的方法时,对于有最小值的二次凸函数来说,使用精确的一维搜索的共轭梯度法具有二次终止性,也就是只需要迭代两次就能将一个二次凸函数收敛到最小值,但是对于稍加复杂的多维高次函数,这种搜索方法带来大量的计算,需要付出较大的努力[11]。所以,文中采取了利用了一种不精确的 Armijo 条件来搜索步长,该方法也为近年来研究较多的一个方法[12-15],其中这个步长计算的终止条件为

$$f(x^{(j)} + \lambda_j d^{(j)}) - f(x^{(j)}) \leq \delta \lambda_j g_j^{\mathrm{T}} d^{(j)} \tag{8}$$

$$\lambda_j = \sigma^m \tag{9}$$

其中 $x^{(j)}$ 表示初始点 x_0 第 j 次迭代后得到的值,λ_j 表示第 j 次迭代的步长, $d^{(j)}$ 表示第 j 次迭代的搜索方向, $\delta \in (0,1)$ 为常数, $\sigma \in (0,1)$ 为常数,且只有 $m \in N^+$ 为未知数,所以步长 λ_j 由最小的正整数 m 值确定。

1.5　共轭梯度法算法步骤

使用共轭梯度法求解 $\min f(x)$,共轭梯度法的计算步骤如下:

Step1:给定初始点 $x^{(1)}$,设置允许的误差为 $\varepsilon>0$,令 $y^{(1)}=x^{(1)}$,$d^{(1)}=-\nabla f(y^{(1)})$,$k=j=1$;

Step2:如果 $\| \nabla f(y^{(j)}) \| < \varepsilon$,则停止计算;否则,使用 Armijo 搜索步长,计算

$$f(x^{(j)} + \lambda_j d^{(j)}) - f(x^{(j)}) \leq \delta \lambda_j g_j^{\mathrm{T}} d^{(j)} \tag{10}$$

$$\lambda_j = \sigma^m \tag{11}$$

根据取值范围,取 $\delta = 0.4$, $\sigma = 0.6$, $g_j = \nabla f(x^{(j)})$

令 $y^{(j+1)} = y^{(j)} + \lambda_j d^{(j)}$;

Step3:如果 $j < n$ (n 为 $f(x)$ 中未知数的个数),则进行 Step4;否则,执行 Step5;

Step4:令 $d^{(j+1)} = -\nabla f(y^{(j+1)}) + \beta_j d^{(j)}$, 其中不同的共轭梯度法,计算因子 β_j 不同, β_j 的计算公式上文已经给出。令 $j = j + 1$, 转到 Step2;

Step5:令 $x^{(k+1)}=y^{(n+1)}$,$y^{(1)}=x^{(k+1)}$,$d^{(1)}=-\nabla f(y^{(1)})$,令 $j=1$,$k=k+1$,转到 Step2.

2　算法案例

此次计算实验是基于 MATLAB(R2014a),通过带入不同的初值 x_0, 以及这 3 种共轭梯度法,对不同的函数进行迭代求解最小值,并得到计算结果 x ,迭代次数 k ,函数的最小值 val ,希望得到的最小值尽可能的接近于函数的最小值,并且若精度太小会使迭代步数太大,设置精度为 10^{-10},若函数在此精度下收敛步数太大或者函数过于复杂,可以将精度适当增加到 10^{-6}。由于采用的是不精确的 Armijo 条件来搜索步长,所以有可能出现迭代结果一直达不到精度要求,所以需要设置迭代次数上限为 50000 次,从而比较不同的初值与不同的共轭梯度法的类型对函数收敛速度的影响,有助于分析每种共轭梯度法的优缺点。

2.1　考虑二次函数无约束优化问题

实例 1　$\min f(x) = x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{2}x_3^2$

设置精度为 10^{-10},对于三元函数,选取 3 个相同的值作为初始值,再选取 3 个距离较远的值作为初始值,以便对不同共轭梯度法的收敛速度进行分析,所以选取了 2 个初始点分别为(20,20,20),(20,50,80),计算得到的结果见表 1、表 2。

表 1　实例 1 初始值为(20,20,20)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 x ($\times 10^{-10}$)	最小值
FR 法	(20,20,20)	18	(-0.0637,-0.1066,-0.1066)	1.5419×10^{-22}
PRP 法	(20,20,20)	25	(0.2392,-0.1632,-0.1632)	8.3886×10^{-23}
HS 法	(20,20,20)	29	(0.2711,-0.1399,-0.1399)	9.3069×10^{-22}

表 2　实例 1 初始值为(20,50,80)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 x($\times 10^{-10}$)	最小值
FR 法	(20,50,80)	20	(-0.0641, 0.3515, 0.5674)	6.3132×10^{-23}
PRP 法	(20,50,80)	21	(-0.1014,-0.0326,-0.0521)	1.2178×10^{-22}
HS 法	(20,50,80)	28	(-0.0668,-0.2183,-0.3493)	8.9273×10^{-22}

比较分析:该函数是一个结构相对简单三元二次函数,由表 1、表 2 中的迭代次数可以看出,初值对于这 3 种算法迭代次数的影响较小,但是又有一定影响,在初始值相同的情况下,FR 法的收敛速度最快,PRP 法的收敛速度较快,HS 法的收敛速度最慢。由函数可以很容易得到函数的最小值点为(0,0,0),所以可以看出(20,50,80)到原点的距离比(20,20,20)远,但是在 PRP 法与 HS 法中,初值较远的点反而收敛速度更快,得出 FR 收敛速度最快,PRP 法收敛速度较快,HS 法收敛速度较慢,并且初值对收敛速度有一定的影响,但是影响比较小。

实例 2　$\min f(x) = x_1^2 - x_1x_2 + x_2^2 + 2x_1 + 4x_2$

设置精度为 10^{-10},对于二元函数,选取 2 个相同的值作为初始值,再选取 2 个距离较远的值作为初始值,以便对不同共轭梯度法的收敛速度进行分析,所以选取的 2 个初始点分别为(20,20),(20,50),计算结果见表 3、表 4。

表 3　实例 2 初始值为(20,20)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 x	最小值
FR 法	(20,20)	53	(-2.6667,-3.3333)	-9.3333
PRP 法	(20,20)	104	(-2.6667,-3.3333)	-9.3333
HS 法	(20,20)	40	(-2.6667,-3.3333)	-9.3333

表 4　实例 2 初始值为(20,50)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 x	最小值
FR 法	(20,50)	174	(-2.6667,-3.3333)	-9.3333
PRP 法	(20,50)	69	(-2.6667,-3.3333)	-9.3333
HS 法	(20,50)	36	(-2.6667,-3.3333)	-9.3333

比较分析:从表 3、表 4 可以看出,FR 在初始值为(20,20)时迭代次数为 53,在初始值(20,50)时迭代次数为 174,相差比较明显,所以初始值对 FR 法的影响较大,PRP 法在不同初始值下迭代次数分别是 104,69,相差也是比较明显,但是差异比 FR 法较小,HS 法在不同初始值下迭代次数分别是 40,36,相差相对较小,在初始值不同的时候相对稳定。并且在 PRP 法与 HS 法中,初始值较远的(20,50)迭代到最小值的速度更快,说明不是越靠近计算结果的初始值,迭代次数越快,可以得到在这个函数中,HS 法的收敛速度最快,FR 法与 PRP 法收敛速度根据初始值的不同,收敛速度变化较大,无法比较收敛速度。

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

设置精度为 10^{-10},对于三元函数,同样选择(20,20,20),(20,50,80)两个点作为初始值,计算得到的结果见表 5、表 6。

表5 实例3初始值为(20,20,20)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 $x(\times 10^{-10})$	最小值
FR 法	(20,20,20)	14	(-0.2333,-0.2333,-0.2333)	1.6327×10^{-21}
PRP 法	(20,20,20)	19	(-0.0075,-0.0075,-0.0075)	1.6973×10^{-24}
HS 法	(20,20,20)	28	(0.1105, 0.1105, 0.1105)	3.6653×10^{-22}

表6 实例3初始值为(20,50,80)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 $x(\times 10^{-10})$	最小值
FR 法	(20,50,80)	26	(0.1749, 0.1445, 0.1142)	7.0046×10^{-22}
PRP 法	(20,50,80)	33	(0.0364, 0.0318, 0.0273)	3.2067×10^{-23}
HS 法	(20,50,80)	38	(0.0709, 0.1175, 0.1640)	5.8737×10^{-22}

比较分析:从表5、表6可以看出,初始值为(20,20,20)时,3种共轭梯度法的迭代次数分别为14,19,28,当初始值为(20,50,80)时,迭代次数分别为26,33,38,表明不同的初始点取值对迭代次数的影响较大,对不同的共轭梯度法迭代次数进行对比,发现在相同的初始值下,FR法收敛速度最快,PRP法收敛速度较快,HS法收敛速度最慢。

通过前面3个不同的二次函数的实例,在实例1与实例3中,FR法收敛速度最快,PRP法收敛次数较快,HS法收敛速度最慢,但是在实例2中,HS法收敛速度是最快的,FR法和PRP法受初值影响较大,无法判断收敛速度快慢,由资料得知,在FR法中,若迭代过程中出现较小的步长,极有可能在接下来的计算中连续出现非常多的小步长[10],这样便增加了迭代次数,造成该算法的迭代效率不佳,故在实例2中,FR法的迭代过程中极有可能出现了非常多的小步长,总的来说,在低次函数中,FR法出现小步长的概率比较小。所以在低次函数中,使用FR法较好。根据文献[3],学者认为PRP法与HS法效率相差无几,甚至PRP法应该优于HS法。根据以上的数值实验,可以看出时而PRP法优于HS法,时而HS法优于PRP法,所以PRP法与HS法的优劣还需进一步计算。

2.2 考虑含有三角函数的二元二次函数无约束优化问题

实例4 $\min f(x)=x_1^2+x_2^2-100\cos(2\pi x_1)-100\cos(2\pi x_2)+20$

对于该函数,同样设置精度为10^{-10},结合以上实例考虑的因素,应选取2个相同的值作为初始值,并选择两个距离较远的值作为初始值,并且由于该函数在0点附近的波动较大,故选取2个初始值分别为(0.2,0.2),(0.2,0.4),计算得到的结果见表7、表8。

表7 实例4初始值为(0.2,0.2)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 $x(\times 10^{-8})$	最小值
FR 法	(0.2,0.2)	8	(-0.3005,-0.3005)	-180
PRP 法	(0.2,0.2)	8	(-0.4118,-0.4118)	-180
HS 法	(0.2,0.2)	8	(-0.7471,-0.7471)	-180

表8 实例4初始值为(0.2,0.4)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 $x(\times 10^{-8})$	最小值
FR 法	(0.2,0.4)	8	(-1.5930,-1.7750)	-180
PRP 法	(0.2,0.4)	10	(-0.1850,-0.4617)	-180
HS 法	(0.2,0.4)	12	(-0.1142,-0.2883)	-180

比较分析:由于该函数的波动较大,所以对初始值的选取有一定的要求,经过测试,初始值选取较大的数值都不能通过该共轭梯度法收敛到最小值,故证明了初始值对震荡较大的函数影响较大,并且由表7、表8可知,3种共轭梯度法在不同的初始值下迭代次数都比较接近,并且FR法小优于PRP法,PRP法小优与HS法。

由以上4个实例可以看出,在二次或低次函数中,

FR 法的收敛速度比较好,PRP 法与 HS 法收敛速度接近,并且在多次计算中,PRP 法要小优于 HS 法。

2.3　考虑高次函数无约束优化问题

实例 5　$\min f(x)(x_2-2x_1)^2+(x_2-2)^4$

该函数为二元四次函数,由于该函数的次数较高,在试验时发现很难达到精度要求,所以为了更好地对 3 种共轭梯度法进行对比分析,将精度调整到 10^{-6},再选择初始点分别为(20,20),(20,50),使用共轭梯度法进行迭代计算,得到的结果见表 9、表 10。

表 9　实例 5 初始值为(20,20)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 x	最小值
FR 法	(20,20)	6116	(0.9970,1.9940)	1.2712×10^{-9}
PRP 法	(20,20)	4706	(1.0033,2.0065)	1.8272×10^{-9}
HS 法	(20,20)	2916	(1.0032,2.0063)	1.6084×10^{-9}

表 10　实例 5 初始值为(20,50)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 x	最小值
FR 法	(20,50)	6107	(1.0030,2.0060)	1.2702×10^{-9}
PRP 法	(20,50)	4196	(1.0033,2.0065)	1.8098×10^{-9}
HS 法	(20,50)	2910	(1.0032,2.0063)	1.6094×10^{-9}

分析比较:从表 9、表 10 可以看出,不同的初始值对同种共轭梯度法的影响较小,FR 法的迭代次数均为 6000 多次,PRP 法迭代次数均为 4000 多次,HS 法迭代次数均为 2000 多次,那么可以非常明显地看出,HS 法收敛速度最快,PRP 法收敛速度较快,FR 法收敛速度较慢。

实例 6　$\min f(x)=\dfrac{1}{3}x_1^6-2.1x_1^4-4x_2^4+x_1x_2+4x_1^2-4x_2^2$

该函数为二元六次函数,与例 5 同样将精度调整到 10^{-6},由于在初始点离收敛点较大时,不能在规定的迭代次数下迭代到符合精度要求的最小值,所以初值分别选择了(1,1)和(1,5)两个值,使用共轭梯度法进行计算,得到的结果见表 11、表 12。

表 11　实例 6 初始值为(1,1)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 x	最小值
FR 法	(1,1)	19	(-0.0898, 0.7227)	-1.0316
PRP 法	(1,1)	13	(-0.0898, 0.7227)	-1.0316
HS 法	(1,1)	13	(-0.0898, 0.7227)	-1.0316

表 12　实例 6 初始值为(1,5)计算结果

计算方法	初始值 x_0	迭代次数 k	计算结果 x	最小值
FR 法	(1,5)	21	(-0.0898, 0.7227)	-1.0316
PRP 法	(1,5)	14	(-0.0898, 0.7227)	-1.0316
HS 法	(1,5)	17	(-0.0898, 0.7227)	-1.0316

分析比较,从表 11、表 12 可以看出,对于相同的初始点,PRP 法收敛速度最快,HS 法收敛速度较快,FR 法收敛速度一般。当初始值不同时,FR 法与 PRP 法的迭代次数变化不大,HS 法有较小的变化。

实例 5、6 中,FR 法的收敛速度都是最慢的,说明在高次函数下,FR 法容易出现较小的步长,从而造成收敛次数增加,在实例 5 中,HS 法优于 PRP 法,但是在实例 6 中,PRP 法优于 HS 法,所以可以认为这两种方法无法分出优劣。从初始值不同的角度来说,在高次函数下,不同的初始值对函数的迭代次数影响较小。

3　结束语

采用基于 Armijo 条件的共轭梯度法来解决求无约束优化中最小值的问题,讨论了 3 种不同的共轭梯度法对于不同函数的收敛速度,并给出了对于不同种

类的函数使用共轭梯度法的建议。由于 Armijo 条件搜索步长是一种不精确的搜索步长的方法,所以没有精确搜索步长的共轭梯度法的一些特征,比如对于二次凸函数的二次终止性,也就是对于二次凸函数,精确搜索步长的共轭梯度法可以很快地得到最优解,但是,精确搜索步长的共轭梯度法在处理较为复杂的函数时,造成了大量的计算,以至于无法得出最后的最优解,为了减少计算量,选择了 Armijo 条件作为步长搜索的工具,计算出很多精确搜索步长的共轭梯度法没法计算的函数,对此,为了减少迭代次数这一目的,应在函数为二次凸函数时,采用精确搜索步长的共轭梯度法求解,在函数较为复杂的情况下,使用文中的基于 Armijo 条件的共轭梯度以及给出的建议进行求解,以达到减少迭代次数这一目的。

通过以上的计算以及分析,发现在二次函数中,使用 FR 法的收敛速度一般都比较快,PRP 法与 HS 法在不同的函数中各有优势。文中测试的 4 个二次函数中,只在一次测试中是 FR 法收敛速度最慢,是由于原函数的梯度较大,造成出现了较小的步长,所以出现了 FR 法收敛缓慢的问题,同时对比不同的初始点对共轭梯度法的影响,不同的初始点对共轭梯度法迭代次数的影响也较大。在含有三角函数的二次函数里,也就是函数变成了相对复杂的振荡函数,计算发现很多初始点不能在设置的迭代步数里迭代出最终的答案,故只能在初始值很接近于最小计算结果的时候,才能很快得出计算结果。最后,分别在四次函数和六次函数下对 3 种函数进行对比,发现在高次函数下,不同的初始值对迭代次数的影响较小,PRP 法与 HS 法的收敛速度都比 FR 法快,但是依然无法比较 PRP 与 HS 法的收敛速度的快慢。综上所述,在解决无约束优化问题时,低次函数求最小值使用 FR 法较快,高次函数求最小值用 PRP 法或 HS 法较快,在震荡幅度较大的函数中与低次函数中,初值的选择也十分重要,应尽量在接近最小值的地方选取初始值。

参考文献:

[1] Hestenes M R. Iterative methods for solving linear equations[J]. Journal of Optimization Theory & Applications,1973,11(4):323-334.

[2] Fletcher R,Reeves C M. Function minimization by conjugate gradients[J]. Computer Journal,1964,7(2):149-154.

[3] 陈宝林. 最优化理论与算法[M]. 2 版. 北京:清华大学出版社,2005.

[4] Dai Y H. Conjugate Gradient Methods with Armijo-type Line Searches[J]. Acta Mathematicae Applicatae Sinica,2002,18(1):123-131.

[5] Polyak B T. The conjugate gradient method in extremal problems[J]. Ussr Computational Mathematics & Mathematical Physics,1969,9(4):94-112.

[6] Hestenes M R, Stiefel E. Methods of Conjugate Gradients for Solving Linear Systems[J]. J. res. nat. bur. stand,1952,49(6):409-436.

[7] Fletcher R. Practical Methods of Optimization, Vol Ⅰ: Unconstrained Optimization[M]. New York: John Wiley & Sons Press,1987,10-30.

[8] Liu Y, Storey C. Efficient generalized conjugate gradient algorithms, Part 1: Theory[J]. Journal of Optimization Theory and Applications, 1991, 69(1):129-137.

[9] Dai Y H, Yuan Y. A nonlinear conjugate gradient method with a strong global convergence property[J]. SIAM Journal on Optimization,2000,10(1):177-182.

[10] Powell M J D. Restart procedures for the conjugate gradient method[J]. Mathematical Programming,1977,12(1):241-254.

[11] 黄元元. 求解无约束优化问题及非线性方程组的共轭梯度法[D]. 西安:西安电子科技大学,2014.

[12] 连淑君,徐衍聪. Armijo 搜索下共轭梯度算法的全局收敛性[J]. 曲阜师范大学学报(自然科学版),2000,26(2):26-29.

[13] Li Y H, Riofrio C A, Cevher V. A General Convergence Result for Mirror Descent with Armijo Line Search[J]. 2018.

[14] Zhu M, Gao T, Zhang B, et al. An Efficient Elman Neural Networks Based on Improved Conjugate Gradient Method with Generalized Armijo Search[C]. International Conference on Intelligent Computing. Springer, Cham,2018:1-7.

[15] Zhang B, Wang J, Wu S, et al. Fully Complex-Valued Wirtinger Conjugate Neural Networks with Generalized Armijo Search[C]. International Conference on Intelligent Computing. Springer, Cham,2018:123-133.

Analysis and Comparison of Several Conjugate Gradient Methods based on Armijo Search Step Length

HUANG Fei,　WU Zezhong

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

Abstract: Conjugate gradient method is an important method to solve the problem of unconstrained optimization. In this paper, the method of inaccurate Armijo-type line search is used, the convergence effects of FR conjugate gradient method, PRP conjugate gradient method and HS conjugate gradient method were compared using MATLAB tools. The results show that the FR conjugate gradient method is better in subharmonic function better convergence using PRP conjugate gradient method or HS conjugate gradient method in higher order functionshen the function fluctuates greatly, the initial value should be chosen as close as possible to the convergence point in order to have a good convergence effect.

Keywords: applied mathematics; optimization theory; FR conjugate gradient method; PRP conjugate gradient method; HS conjugate gradient method; unconstrained optimizations; Armijo-type line search