第8卷第7期
1997年7月

软 件 学 报
JOURNAL OF SOFTWARE

Vol. 8, No. 7
Jul. 1997

非线性最小二乘全局解的混合计算智能算法*

赵明旺

(武汉冶金科技大学自动化系　武汉　430081)

摘要　通过在遗传算法中嵌入牛顿算子,并定义适当的适应度和数据结构,从而得到可结合遗传算法和牛顿法两者长处,既有较快收敛性,又能以较大概率求得非线性最小二乘全局解的混合计算智能算法.数值结果表明了该方法显著优于遗传算法和牛顿法.

关键词　计算智能,遗传算法,牛顿法,非线性最小二乘问题.

中图法分类号　TP18

非线性最小二乘(NLLS)问题是非线性方程和非线性优化领域中重要问题,它在系统工程、控制工程、统计学等学科有着重大应用背景.如何快速有效地求解该问题是计算数学界和应用科学领域倍受注目的问题.目前,该问题的主要数值方法有:传统的非线性优化算法(如最速下降法、牛顿法、共轭梯度法等)、Gauss-Newton 方法[1]、Dennis 改进法[2]以及锥模型算法[3]等.众所周知,这些算法都是局部极值算法,得到的仅是局部解而非全局解.对 NLLS 问题,目前还缺乏有效的全局解算法,实际应用中只能通过在多个初始点上使用传统数值优化方法来尽可能地求取全局解,但该种处理方法求得全局解的概率不高,可靠性低.建立以尽大可能(概率)求解全局解的算法仍是一个重要的问题.

近年来,模拟生物进化中"物竞天择,适者生存"原则的计算智能(computational intelligent)方法——遗传算法(GA)以其能以较大概率求得全局最优解,具有较强鲁棒性、适应性和并行性等特点吸引了众多领域中的研究人员,并在函数优化、模式识别、图象处理、人工智能得到广泛应用.[4,5]但对于连续函数的优化问题,GA 由于收敛性相对较慢,编码长度对精度影响大等因素,与常用数值优化方法相比并不具有优势.由于 NLLS 问题可视为一种特殊的非线性函数优化问题,因此 GA 也可以推广到 NLLS 全局解求解.

本文利用 GA 中杂交(crossover)算子和变异(mutation)算子在全变量空间,以较大概率搜索全局极值的特点,以及在解点附近牛顿法收敛快、计算精度高的特点,研究具有杂交算子、变异算子和 NLLS 算子 3 种基本运算,适用于求解 NLLS 全局解的混合计算智能算

* 本文研究得到武汉市科委"晨光计划"资助.作者赵明旺,1964 年生,博士,教授,主要研究领域为人工智能,计算智能,智能控制.

本文通讯联系人:赵明旺,武汉 430081,武汉冶金科技大学自动化系

本文 1996-08-26 收到修改稿

法(HCIA).

1 问题描述

考虑如下有限空间的 NLLS 问题的全局解求解

$$\min_{\substack{a_i \leqslant x_i \leqslant b_i \\ i=\overline{1,n}}} f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{r}^{\tau}(\boldsymbol{x})\boldsymbol{r}(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{r}(\boldsymbol{x})\|^2 \tag{1}$$

其中 n 为变量向量 $\boldsymbol{x}$ 的维数,$\boldsymbol{r}(\boldsymbol{x})=[r_1(\boldsymbol{x}) \quad r_2(\boldsymbol{x}) \quad \cdots \quad r_p(\boldsymbol{x})]^{\tau}$. 对 NLLS 全局解求解,下面先分别讨论基于牛顿法和 GA 的求解算法.

1.1 牛顿法

牛顿法是具有二阶局部收敛性的数值优化算法[6],其特点是收敛较快,但需要求解函数的二阶导数矩阵(Hessian 矩阵). 牛顿法求解式(1)所示的 NLLS 问题的计算过程如下:

Begin
  选定初始迭代点
  Repeat
    计算 $f(x_k)$, $\nabla f(x_k)=\left(\frac{\mathrm{d}\boldsymbol{r}(x_k)}{\mathrm{d}x}\right)^{\tau} r(x_k)$, 以及
    $\nabla^2 f(x_k)=\left(\frac{\mathrm{d}\boldsymbol{r}(x_k)}{\mathrm{d}x}\right)^{\tau}\frac{\mathrm{d}\boldsymbol{r}(x_k)}{\mathrm{d}x}+\left[\frac{\mathrm{d}}{\mathrm{d}x_{1,k}}\left(\frac{\mathrm{d}r(x_k)}{\mathrm{d}x}\right)^{\tau} r(x_k) \quad \cdots \quad \frac{\mathrm{d}}{\mathrm{d}x_{n,k}}\left(\frac{\mathrm{d}r(x_k)}{\mathrm{d}x}\right)^{\tau} \boldsymbol{r}(x_k)\right]$
    if (Hessian 矩阵 $\nabla^2 f(x_k)$ 可逆)
      搜索方向 $p_k=[\nabla^2 f(x_k)]^{-1}\nabla f(x_k)$
    else
      搜索方向 $p_k=\nabla f(x_k)$
    用黄金分割法作线性搜索 x_{k+1}=linear-search$(x_k, -p_k)$
  Until (结束迭代逻辑条件满足)
End

1.2 遗传算法(GA)

如引言中所述,GA 对于 NLLS 全局解,与数值优化方法相比并不具有优势. 即便如此,它的杂交算子、变异算子和选择算子在求取全局解中的效能对求取全局解还是具有指导作用的. GA 求解式(1)所示的 NLLS 全局解的计算过程如下:

Begin
  随机产生初始父代群体,计算其适应度并对每个个体进行编码
  Repeat
    选择群体中两个个体以杂交运算概率 P_c 进行杂交运算
    选择群体中个体以变异运算概率 P_m 进行变异运算
    对每个个体进行解码并计算其适应度
    选择子代群体(适者繁殖,不适者被淘汰)
  Until (结束迭代逻辑条件满足)
End

GA 应用于求解式(1)所示的 NLLS 全局解的困难之处在于: 1) 编码方法和编码长度. 2) 如何控制群体向某个或某几个个体集中的速度. 若该速度快,则求得全局极值的概率将变小;若过慢,则 GA 将退化为 Monte-Carlo 完全随机试验算法. 3) 适应度的确定.

2 混合计算智能算法(HCIA)

针对牛顿法和 GA 在求解 NLLS 全局解上的不足,综合各自的优势,本文提出如下

HCIA.

```
Begin
  随机产生初始父代群体并计算其适应度
  Repeat
    选择群体中两个个体以概率 Pc 进行编码，杂交，解码运算，将父代和子代都加入子代群体
    对子代群体中每个个体以概率 Pm 进行编码，变异，解码运算，将父代和子代都加入子代群体
    对子代群体中每个个体以概率 Pn 进行牛顿算子搜索，将子代取代父代加入子代群体
    对每个个体计算适应度
    以既定的群体规模，选择下一次繁殖的子代群体(适者繁殖，不适者被淘汰)
  Until (结束迭代逻辑条件满足)
End
```

对上述 HCIA，说明如下：

2.1 适应度

由于进行杂交运算和子代选择时所需的概率与适应度成正比，因此适应度对 GA 和 HCIA 有全局性的影响，是算法成功求得全局解的关键之一. 本文的适应度 $g(x)$ 定义为

$$g(x) = f_{\max} - f(x) + k_1(f_{\max} - f_{\min}) \tag{2}$$

其中 $f_{\max}$ 和 $f_{\min}$ 分别为当前群体中的最大和最小函数值；k_1 为控制参数.

由上式可知，在每次繁殖的群体中最大与最小适应度之比为 $(1+k_1)/k_1$. 若 k_1 过大，该比值趋近于 1，即当前群体中最大函数值的个体和最小函数值的个体有较接近的适应度和选择概率，则杂交算子和变异算子的作用将退化至随机搜索. 若 k_1 过小，该比值趋近于∞，将使得函数值较大的个体的适应度和选择概率过小，失去繁殖的机会，从而导致群体快速向某个个体或某几个个体集中，使得求得全局解的概率下降. 一般情况下，k_1 的取值在 0.01～0.1之间较适宜，相应的个体间的最大和最小适应度之比为 11～101 之间. 由于每代的 $f_{\max}$ 和 $f_{\min}$ 随群体不同而变化，使得 HCIA 在确定适应度和选择概率上具有自适应性和鲁棒性.

2.2 群体的数据结构和编码方法

GA 的主要数据结构是数字位串，一般常用的是二进制. 当 GA 应用于 NLLS 问题时，位串长度和编码方法对计算精度具有决定性影响，并直接影响全局解的求解.

本文的 HCIA 的数据结构采用混合式数据结构，在群体中每个个体采用实型数的数据结构，只有在个体被确定进行杂交运算和变异运算时才进行编码，运算结束后产生子代的要进行解码才能加入新的子代群体. 采用混合数据结构的优点是可以避免编码的有限字长对精度的影响，充分发挥牛顿算子精度高以及能局部细致地搜索解的特性. HCIA 在杂交运算和变异运算中个体的编码采用各变量 $x_i(i=1,2,\cdots,n)$ 的二进制位串以顺序排列，其中变量 x_i 的长度为 N 的二进制位串 $\bar{x}_i$ 的编码算法为

$$\bar{x}_i = \mathrm{int}\left(\frac{x_i - a_i}{b_i - a_i}(2^N - 1)\right) \tag{3}$$

2.3 牛顿算子

为了加速 GA 在求解 NLLS 全局解的收敛性，发挥传统数值优化算法的局部收敛性好以及在计算速度与精度上的优势，本文的 HCIA 中嵌入了牛顿算子. 该算子主要进行的是函数优化牛顿法的线性搜索运算. 在每次繁殖中产生的子代，都要以概率 P_n 判断是否需要进行线性搜索运算. 本文的牛顿算子所进行的具体运算为：

```
Repeat
  对群体中第 i 个体产生[0,1]间随机数. 若该随机数大于既定牛顿算子概率 P_n 则进行下述运算
  计算 f(x_k), ∇f(x_k) = (dr(x_k)/dx)^τ r(x_k), 以及
  ∇²f(x_k) = (dr(x_k)/dx)^τ dr(x_k)/dx + [ d/dx_{1,k} (dr(x_k)/dx)^τ r(x_k) ... d/dx_{n,k} (dr(x_k)/dx)^τ r(x_k) ]
  if  (Hessian 矩阵 ∇²f(x_k) 可逆)
      搜索方向 p_k = [∇²f(x_k)]^{-1} ∇f(x_k)
  else
      搜索方向 p_k = ∇f(x_k)
  用黄金分割法作线性搜索 x_{k+1} = linear-search(x_k, -p_k)
  将产生的子代取代父代加入新的子代群体
Until(搜索完整个群体)
```

其中 $\nabla^2 f(x_k)=\left(\frac{\mathrm{d}r(x_k)}{\mathrm{d}x}\right)^\tau\frac{\mathrm{d}r(x_k)}{\mathrm{d}x}+\left[\frac{\mathrm{d}}{\mathrm{d}x_{1,k}}\left(\frac{\mathrm{d}r(x_k)}{\mathrm{d}x}\right)^\tau r(x_k)\ \cdots\ \frac{\mathrm{d}}{\mathrm{d}x_{n,k}}\left(\frac{\mathrm{d}r(x_k)}{\mathrm{d}x}\right)^\tau r(x_k)\right]$

可以这么说,本文 HCIA 中的遗传算子——杂交算子、变异算子和选择算子的作用是宏观搜索,处理的是大范围搜索问题,而牛顿算子中的线性搜索过程的作用是极值局部搜索,即微观搜索,处理的是小范围搜索问题和搜索加速问题.牛顿算子概率 P_n 的取值应能保证在繁殖(迭代)过程中,群体的每个个体都有机会得到一定机会牛顿算子的线性搜索运算.因此,确定牛顿算子概率 P_n 的大小需考虑的因素为:(1)所求解问题的牛顿法线性搜索迭代的收敛性.若其线性搜索迭代收敛较快,则相应的 P_n 可取小一些.否则,则取大一些.(2)预定的繁殖代数(迭代次数).若预定的繁殖代数较大,则相应的 P_n 可取小一些.否则,则取大一些.

2.4 选择算子,选择概率和控制个体间的距离

GA 与传统优化算法比较,其之所以能以较大概率求得全局解,是因为其算法思想中体现了两个基本原则:一为适者生存,二为选择的相对随机性.而选择算子、选择概率和控制个体间的距离是体现这些原则的最重要的因素.HCIA 对这些因素的考虑是:

(1) 杂交运算和变异运算后,子代和父代同时都进入子代群体;而由于牛顿算子是一种能保证迭代产生的序列,是函数单调下降的良好的局部极值算法,因此其子代较好地继承了父代的优良品质,故可取代父代进入子代群体,而不需要父代和子代都进入子代群体.

(2) 在选择下次繁殖的子代群体时,每次以候选群体中个体的概率随机地选择两个个体,其中适应度大的个体进入子代群体,直至产生完所有繁殖所需的下代群体为止.这样既可保证有相对的适者生存,又使选择具有一定的随机性.

(3) 为避免在繁殖中群体向某个或有限几个个体快速集中,降低求得全局解的概率,在选择算子运算中还必须计算群体中各个体间的距离(常用的有 2-范数,∞-范数,Hamming 距离等),相邻较近的两个个体必须剔出其一.由于本文的 HCIA 中的牛顿算子能局部地、细致地进行优化搜索,故在控制群体中个体间的距离时,可考虑让个体之间的相对距离较大(一般相对误差可在 1% 以上),使能在更大的空间范围内,以较大概率搜索全局解.

3 数值计算和结束语

本节考虑如下在有限区间的 NLLS 问题的全局解求解.

例 1:

$$f(x)=\|[x_1-10x_2\quad \sqrt{5}(x_3-x_4)\quad (x_2-2x_3)^2\quad \sqrt{10}(x_1-x_4)^2\quad 7\sin(kx_1)\quad 7\sin(kx_2)\quad 7\sin(kx_3)\quad 7\sin(kx_4)]^\tau\|^2\quad -1\leqslant x_i\leqslant 1 \qquad (4)$$

例 1 当 $k=0$ 时是函数全局优化问题和 NLLS 问题的著名算例 Powell 函数[7]，许多文献都采用它对求解全局优化的效率作说明. Powell 函数的全局解为 $f(\boldsymbol{x}^*)=f(0,0,0,0)=0$，且在 $\boldsymbol{x}^*$ 处的 Hessian 矩阵奇异，许多算法对此将产生麻烦. 例 1 当 k 为非零时相当于在 Powell 函数密集地增加许多局部解点（非全局解点），大大增加求解全局解的难度. 传统数值优化算法和 NLLS 求解算法对此更是显得不足，求得全局解的效率极低.

本文利用 HCIA 进行了大量计算机数值模拟计算. 在模拟计算中，先产生[－1,1]间的 1500 组 4 维随机向量，每 30 组作为初始繁殖群体，共进行 50 次计算. 考虑到 GA 和本文的 HCIA 编码长度的因素，对 $k=0$ 时若 $f(\boldsymbol{x})<10^{-6}$，对 k 为非零时若 $f(\boldsymbol{x})<10^{-5}$，则认为搜索到相应的解，算法成功. 计算结果如表 1 所示.

表 1　计算结果

		牛顿法	混合计算智能算法($P_c=.9, P_m=.1$)							
			$P_n=.0$	.025	.05	.075	.1	.15	.2	.3
情况 1 $k=0$	搜索到解的次数	50	19	50	50	50	50	50	50	50
	搜索到解的概率	1.0	0.38	1.0	1.0	1.0	1.0	1.0	1.0	1.0
情况 2 $k=8$	搜索到解的次数	3	16	25	34	42	46	50	50	50
	搜索到解的概率	0.06	0.32	0.5	0.68	0.84	0.92	1.0	1.0	1.0
情况 3 $k=12$	搜索到解的次数	1	13	23	26	35	41	47	50	50
	搜索到解的概率	0.02	0.26	0.46	0.52	0.7	0.82	0.94	1.0	1.0
繁殖代数		2000	10000	2000	2000	2000	2000	2000	2000	2000

对上述计算结果有如下说明和分析：

(1)当 $P_c=P_m=0, P_n=1$，HCIA 相当于在 30 个初始迭代点同时进行搜索的牛顿算法，表 1 的牛顿法求解即按此方式进行. 由求解结果可知，当 $k=0$ 时，其求解迭代过程都能收敛于其解，而当 k 为非零时，由于函数 $f(\boldsymbol{x})$ 非单调下降，存在多个局部解点，经常使得迭代过程不能收敛于其解而失败，导致求解效率低.

(2)当 $P_n=0$ 时，HCIA 退化为 GA，上述算例的 GA 求解即按此方式进行. 对与 Powell 函数类似的单调性较好的函数，在指定的精度范围和有限的繁殖代数内，GA 求得解的概率低于牛顿法且需耗费更多的计算时间（代价）；对与修改 Powell 函数类似的单调性较差的多峰函数，求得解的概率则高于牛顿法.

(3)本文的结合 GA 和牛顿法特长的 HCIA 能充分发挥遗传算子和变异算子大范围搜索全局解的特点，以及牛顿算子在局部细致地搜索的特点，对与修改 Powell 函数类似的单调性较差的多峰函数，求得解的概率则显著高于牛顿法和 GA；对与 Powell 类似的单调性较好的函数，求得解的概率与牛顿法相当. 选择概率 P_n 的大小还直接影响求解效率. 一般来说，对存在多个极值点的函数，P_n 可取得大一些；对极值点数少、单调性较好的函数，P_n 可取得小一些. 图 1 和图 2 分别是当 k 为 8 和 12 时，不同的 P_n 取值下，求得解的概率与 P_n 的关系曲线图.

由上述计算结果可知，本文的基于 GA 和牛顿法的 HCIA 在指定空间求得全局解的概率显著地大于牛顿法和 GA，是一种可行的求解 NLLS 全局解的数值算法.

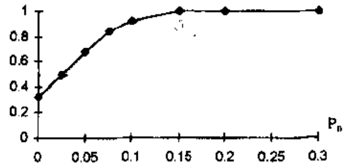

图1 当$k=8$求得解的概率与P_n的关系曲线

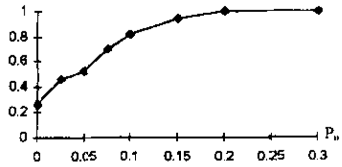

图2 当$k=12$求得解的概率与P_n的关系曲线

参考文献

1 Dannis J E, Schnabel R B. Numerical method for unconstrained optimization and nonlinear equations. Prentice-Hall, Englewood Cliffs, N.J., 1983.

2 Dannis J E, Gay D M, Welsch R E. An adaptive nonlinear least-squares algorithm. TOMS, 1981, **7**: 348～368

3 韩乔明，盛松柏. 解非线性最小二乘问题的锥模型算法的局部收敛性. 高等学校计算数学学报，1996, **18**(1): 77～86.

4 Goldberg D E. Genetic algorithms in search, optimization, and machine reading. MA: Addison-Wesley, 1989.

5 姚新，陈国良，徐惠敏等. 进化算法研究进展. 计算机学报，1995, **18**(9): 694～706.

6 席少霖，赵风治. 最优化计算方法. 上海：上海科学技术出版社，1983.

7 Wolfe M A. Numerical methods for unconstrained optimization. VNR, 1978.

A HYBRID COMPUTATIONAL-INTELLIGENT ALGORITHM FOR THE NONLINEAR LEAST-SQUARES GLOBAL SOLUTION

ZHAO Mingwang

(*Department of Automation Wuhan Yejin University of Science & Technology* *Wuhan* 430081)

Abstract Through embedding a Newtonian operator into the genetic algorithm and defining a proper fitness and a numerical structure, a hybrid computational-intelligent algorithm for the global solution of the nonlinear least-squares problem, combined the advances of both of the genetic algorithm and the Newtonian algorithm, is got with the faster convergence and the greater probability for the global solution. The numerical results show that the method is distinctly superior to the genetic algorithm and the Newtonian algorithm.

Key words Computational intelligent, genetic algorithm, Newtonian algorithm, nonlinear least-squares problem.

Class number TP18