****

**毕 业 论 文**

|  |
| --- |
| **改进遗传算法在非线性方程组中的求解** |
|  |

**涂清韶**

**201330320420**

|  |  |
| --- | --- |
| 指导教师 | **李康顺 教授** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 学院名称 |  | **数学与信息学院** | 专业名称 |  | **计算机科学与技术** |
| 论文提交日期 |  | 2017年4月 | 论文答辩日期 |  | 2017年5月 |

答辩委员会主席 \_\_\_\_\_\_\_\_\_\_\_\_

评 阅 人 \_\_\_\_\_\_\_\_\_\_\_\_

摘 要

遗传算法是在生物学的自然选择学说中关于遗传进化机制理论里得到启示的思想，是一种具有高度自适应性、本质并行性、通用性强的随机迭代优化搜索算法。众所周知，求解非线性方程组所采用的传统方法，例如牛顿法、单纯形算法等、都是通过迭代来运算得出结果，但毋庸置疑的是这些方法都存在运算量大、求解速度慢、求解精度误差大等问题。本文通过利用求解目标函数的极值的思想，并结合遗传算法的特点，设置适当的选择算子、杂交算子、变异算子等参数，对非线性方程组进行转化为目标函数，并最终求得目标函数极值相应的解得出相应非线性方程组的解。通过选择典型的非线性方程组进行验证，相应的实验结果表明，该方法显然比传统的求解非线性方程组的方法出色，其还具有运算速度快、求解精度高、通用性好的特点。

关键词：遗传算法 非线性方程组 目标函数

目 录

[1 引言 1](#_Toc293939657)

[1.1 研究背景、目的与意义 1](#_Toc293939658)

[1.2 本论文的研究内容和结构安排 1](#_Toc293939660)

[2 非线性方程组分析 2](#_非线性方程组分析)

[2.1 非线性方程组问题描述 2](#_非线性方程组问题描述)

[2.2 传统求解非线性方程组方法分析 4](#_Toc293939661)

[2.3 本章小结 4](#_Toc293939661)

[3 遗传算法的分析 2](#_遗传算法的分析)

[3.1 演化计算 2](#_演化计算)

[3.2 遗传算法 4](#_遗传算法)

[3.3 本章小结 4](#_本章小结)

[5 总结与展望 33](#_Toc293939690)

[5.1 本论文总结 33](#_Toc293939691)

[5.2 今后的展望 33](#_Toc293939692)

[致 谢 34](#_Toc293939693)

[参 考 文 献 35](#_Toc293939694)

[英 文 摘 要 38](#_Toc293939695)

# 引言

## 研究背景、目的与意义

随着经济的快速发展和人们生活水平的不断提高，对科学技术的研究成果运用在生活中的要求也在不断的更新，导致当前科学技术进入了多学科互相交叉、互相渗透、互相影响的时代。生物学科与计算机学科、理学学科的碰撞、交叉、渗透和相互促进是其中一个典型例子，计算机学科的高速发展为理学学科、生物学科提供了理论定理验证和实现的可能，运用计算学科验证无法用人工计算得到的结论是近代科学技术发展的一个显著特点。最近几年来，遗传算法的研究、改进、优化等热议正体现了科学发展的这一特点和趋势。

非线性方程组求解是实际工程领域的一个重要问题 ,在数值天气预报、石油地质勘探、计算生物化学、控制领域和轨道设计等方面均有较强的应用背景。非线性问题亦是现代数学的主要研究课题之一，这不仅是由于科学技术发展的需要，而且是由于计算机科学与技术的高速发展提供了解决这类问题的可能，求解非线性方程组的方法很多， 如牛顿法、 单纯形算法、 同伦延拓法、 区间迭代法等。但这些传统的非线性方程组解法存在运算量大、求解速度慢、大多数的数值求解依赖于迭代初始点的选择等缺点。

由于方程组求解问题可以转化为目标函数求极值问题以及遗传算法在函数优化的大量的运用，我们可以应用遗传算法求解非线性方程组，通过将非线性方程组的问题转换为求解目标函数的极值问题而得到其相应的解。该方法具有求解速度快，精度高、通用性广的特点，从而优化了传统求解非线性方程组的缺陷，在现实中具有重大的进步意义。

综上，研究如何运用遗传算法求解非线性方程组将具有十分重要的现实意义，亦是计算机科学对数学领域相互影响、相互促进发展的结晶。

## 本论文的研究内容和结构安排

本论文研究的内容是对非线性方程组的分析，对遗传算法的介绍，进而运用遗传算法进行求解非线性方程组。本论文分为五章，第一章是引言，介绍了求解非线性方程组的研究背景、运用遗传算法求解的重要意义。第二章是对非线性方程组的分析以及传统的求解算法的简单概括。第三章是对遗传算法的原理和相关概念及其应用的分析和描述。第四章是基于改进的遗传算法对非线性方程组求解。第五章是结论，对本论文实现的改进遗传算法求解非线性方程组进行一个总结，分析其优点以及有待改进和完善的地方。

# 非线性方程组问题的分析

## 非线性方程组问题描述

非线性方程组是含有N个未知数，由M个方程组组成的实际函数。其一般的表示形式为：

(1)

其中含有n个变量X = 和m个方程。

求解此方程组可转化为求解下面一个极值优化问题：

(2)

(2)式中A为解区间的上限，B为解区间的下限()，则当最小值为0时，对应的X即为方程组(1)的解。当最小值不为0时，方程组(1)无解。那么便把非线性方程组求解问题转化为求目标函数极值的函数优化问题。

## 传统求解非线性方程组方法分析

## 本章小结

# 遗传算法的分析

## 演化计算

大自然是我们解决各种问题时获得灵感的源泉。将生物界所提供的答案应用于实际问题求解已被证明是一个成功的方法，并且已形成一个专门的科学分支——仿生学(bionics)。自然界所提供的答案是经过漫长的自适应过程——演化过程而获得的结果。除了演化过程的最终结果，我们也可以利用这一过程本身去解决一些较为复杂的问题。这样，我们不必非常明确地描述问题的全部特征，只需要根据自然法则来产生新的更好解。

演化计算正是基于这种思想而发展起来的一种通用的问题求解方法。它采用简单的编码技术来表示各种复杂的结构，并通过对一组编码表示进行简单的遗传操作和优胜劣汰的自然选择来指导学习和确定搜索的方向。

由于它采用种群的方式组织搜索，这使得它可以同时搜索解空间内的多个区域。而且用种群组织搜索的方式使得演化算法特别适合大规模并行。在赋予演化计算自组织、自适应、自学习等特征的同时，优胜劣汰的自然选择和简单的遗传操作使演化计算具有不受其搜索空间限制性条件(如可微、连续、单峰等)的约束及不需要其它辅助信息(如导数)的特点。这些崭新的特点使得演化算法不仅能获得较高的效率而且具有简单、易于操作和通用的特性，而这些特性正是演化计算越来越受到人们青睐的主要原因之一。

演化计算诞生在20世纪六七十年代，但在这些年代未受到普遍的重视，亦没有广泛的推广使用。主要原因如下：1. 当时这些方法本身还不够成熟；2. 这些方法需要较大的计算量，而当时的计算机还不够普及而且速度也跟不上要求，这样便限制了它们的应用；3. 当时基于符号处理的人工智能方法正处于其顶峰状态，使得人们难以认识到其它方法的有效性及适应性。

到了80年代，人们越来越清楚地意识到传统人工智能方法的局限性，并且由于计算机速度的提高以及并行计算机的普及，已使得演化计算对机器速度的要求已不再是制约其发展的因素。随着演化计算的不断发展及其在一些应用领域内取得的成功，目前演化计算在机器学习、过程控制、经济预测、工程优化等领域都取较大得了成功，也引起了包括数学、物理学、化学、生物学、计算机科学、社会科学、经济学以及工程应用等领域科学家们得极大兴趣。演化计算目前有四大分支：

* 遗传算法(Genetic Algorithm，简称GA)
* 演化规划(Evolutionary Programming，简称EP)
* 演化策略(Evolution Strategy，简称ES)
* 遗传程序设计(Genetic Programming，简称GP)

本论文研究并使用到得就是其分支之一：遗传算法。

## 遗传算法

### 遗传算法由来介绍

随着科学技术的不断发展，把计算机科学与进化论结合起来的尝试开始于20世纪50年代末，但由于缺乏一种通用的编码方案，使得人们只能依赖变异而不是杂交来产生新的基因结构，故而收效甚微。到20世纪60年代中期，美国Michigan大学的John Holland提出了位串编码技术，这种编码既适合于变异又适合杂交操作，并且他强调将杂交作为主要的遗传操作。随后，Holland将该算法用于自然和人工系统的自适应行为的研究之中，并于1975年出版其开创性的著作《Adaptation in Natural and Artificial Systems》。后来，Holland与他的学生们将该算法加以推广并应用到优化及机器学习等问题之中，因此遗传算法便诞生了。

### 遗传算法思想概述

遗传算法是一种借鉴大自然选择规律(适者生存、不适者被淘汰的遗传机制)演变而来的具有高度自适应性、本质并行性、通用性强的随机迭代优化搜索算法。它是依据一类初始种群在不断生存中进行基因变化，持续产生最优秀个体，通过最优个体之间的选择、交叉、变异产生更优秀的基因，不断地迭代繁殖、演化以找到最适应环境的个体，自下而上的优化最终形成最优种群的过程而形成的关键技术。

### 遗传算法相关术语

* 种群 (Population) ：生物的进化以群体的形式进行，这样的一个群体称为种群。
* 种群规模 (Population size)：种群内个体数目的总和称为种群规模。一般而言，种群规模在整个遗传算法的演变过程中是固定不变的。
* 个体 (Individual) ：组成种群的单个生物，在遗传算法中通常表述为染色体，因遗传操作主要发生在染色体上。
* 基因 ( Gene ) ：个体生物的遗传因子。
* 染色体 ( Chromosome ) ：包含一组的基因，遗传算法中最重要的依托，因为选择、杂交、变异、复制等遗传操作都发生在染色体上，从而形成最优个体。
* 适应度 (Fitness) ：指个体在一定环境下的生存适应程度。对环境适应度高的、优秀的个体参与繁殖的机会比较多，后代就会越来越多。适应度低的个体参与繁殖的机会比较少，后代就会越来越少。
* 编码 (Coding) : 根据相关的实际问题，将问题的解空间转化为相应的基因空间，即运用字符串表示解空间，这是使用遗传算法之前需第一步进行的过程，编码的目的是为了有效地执行遗传操作。选择何种编码会对算法的性能、 效率等产生很大影响。常见的基因编码方案有：二进制编码、实数编码、格林编码等。
* 适应度函数 (Fitness Function) : 用于评价某个染色体的适应度，常用表示。即通过相应的函数值来区分优秀的个体，并让优秀的个体进行下一步的遗传，适应度函数的确定对遗传算法的收敛速度有一定的关系。
* 遗传与变异：新个体会遗传父母双方各一部分的基因，同时有一定的概率发生基因变异。
* 遗传算子 (Genetic operators) : 即在遗传、繁殖、演变过程的过程中进行的操作，与自然选择学说中的基因多样化原因如初一则，其包括复制操作、选择操作、杂交操作和基因变异操作。
* 选择 (Selection) : 是指从种群中以随机的机会选择一些具有优秀基因的染色体来产生下一代。一种常用的选择策略是 “比例选择”，也就是个体被选中的概率与其适应度函数值成正比。
* 复制 (Reproduction) : 指从原本的种群中以随机的概率复制一些个体到新生的种群中进行繁殖和遗传。
* 杂交 (Crossover) : 是指对两个相互配对的染色体依据交叉概率按某种方式相互交换其部分基因，从而形成两个新的个体。杂交运算在遗传算法中起关键作用，是产生新个体的主要方法。
* 变异 (Mutation) : 是指染色体上的一个位置上或者几个位置上的基因发生变化，从而形成新个体的过程，产生的新个体可能是适应度高或者劣质的个体，这里需通过运用适应度函数判断。

### 遗传算法的相关描述

遗传算法的基本结构可用如下伪代码所表示：

{

随机初始化种群

计算中个体的适应度函数值；

While(不满足终止条件) do

{

用经过遗传操作形成新的种群；

计算中个体的适应度函数值，t = t + 1;

}

}

遗传算法可以形式化描述为：

*GA*＝(*P*(0), *N*, *l*, *s*, *g*, *p*, *f*, *t*)

其中*P*(0)＝(*a*1(0), *a*2(0),…,*aN*(0)) ∈ *IN*，表示初始种群；

*I* = *Bl* = {0,1}*l*表示长度为*l*的二进制串全体，称为位串空间；

*N*表示种群中含有个体的个数；

*l*表示二进制串的长度；

*s* : *IN* → *IN*表示选择策略；

*g*表示遗传算子，通常包括繁殖算子*Or* : *I* → *I*、杂交算子 *Oc* : *I*×*I* → *I*×*I*和变异算子*Om* : *I* → *I*

*p*表示遗传算子的操作概率，包括繁殖概率 *pr* 、杂交概率 *pc* 和变异概率 *pm*；

*f* : *I* → *R*+ 是适应函数；

*t* : *IN* → {0, 1}是终止准则。

### 简单遗传算法介绍

Holland 等人提出的遗传算法常被称为简单遗传算法(简记SGA)，SGA对操作对象运用的是二进制编码方案，种群内的每个个体都对应于实际问题的一个解。

该算法以初代种群为起点，通过计算每个个体相应的适应函数值，然后依据每个个体适应函数值在整个种群的适应函数值比例，从中选择优秀的个体，使用杂交和变异操作衍生新一代优秀的种群，通过这个不断演化的过程，最终形成最优种群亦或者达到满足的期望终止条件。

SGA的算法流程图x所示：

开始

初始化群体*P*(0), *t*←0

计算*P*(t)中每个个体的适应值

满足终止条件?

输出结果

结束

*t*=*t*+1

*k*←0

Y

N

*k*≥*N*?

根据个体的适应函数值比例选择两父体

N

*round*[0,1] ≤ *pc* ?

对两父体进行杂交操作

将后代串插入到*P*(*t*+1)

将两父体变异后插入到*P*(*t*+1)

*k*←*k*+2

Y

N

图x SGA算法流程图

SGA的基本执行步骤可用如下文字表述：

1. 随机产生初始种群，个体数目一定，每个个体用二进制编码表示染色体基因；
2. 依据适应度函数，计算当前种群每个个体的适应函数值，并与终止算法的界限做出对比，若满足，则把最优个体及其相应的适应函数值输出，并且终止算法。否则，则继续第三步的过程，直到满足相应的界限条件为止；
3. 根据每个个体的适应函数值比例选择两个个体，并且对其进行下一步的繁殖操作，在这一步，适应函数值低的个体将可能被淘汰出局；
4. 依据随机生成[0,1]区间的值和杂交概率进行对比，若满足小于等于则对第三步选择的两个个体进行杂交，产生新的个体；否则进行第五步操作；
5. 依据一定的变异概率, 对第三步选择的两个个体进行变异，产生新的个体；
6. 如果满足通过上一代种群经过演化而来的个体总数与上一代种群的个体总数一致，便进行第二步操作，否则转到第三步操作。

### 遗传算法的评价及其应用

遗传算法与传统的搜索算法相比较，主要有以下优势：

* 1. 遗传算法把决策变量的编码作为运算对象。传统的搜索算法往往直接运用决策变量实际值本身，而遗传算法处理决策变量是通过采用某种编码形式，使得我们可以借鉴自然选择学说中的染色体和基因的概念，可以模仿自然界生物的遗传和进化机制，也使得我们对实际问题的操作中能更简单、方便的使用遗传算子。
  2. 遗传算法不是盲目穷举搜索的算法，本身带有启发式，不是运用确定性规则的思想来寻求问题答案的解，其使用的是概率搜索技术，通过遗传算子中杂交、变异的不确定性以及随机转移规则进行搜索。
  3. 遗传算法直接以目标函数和相应的适应度函数值为搜索信息，不像传统搜索算法，需要一些其它辅助的信息，如运用一次、二次导数、矩阵等方法。况且适应度函数不受连续、可微等条件的约束，适用范围很广。
  4. 遗传算法是从一个群体即多个点而不是从一个点开始搜索，具有并行性处理问题的能力，这是它能以较大的概率找到整体最优解的主要原因之一；相应的它会产生的多个可能的解，如何选择最终的解则需要我们使用者进行决定。
  5. 遗传算法具有通用性强、容易实现的特质。一旦有了一个遗传算法的程序，如果想解决一个新的问题，只需针对新的问题重新进行基因编码就行；如果编码方法也相同，那只需要改变一下适应度函数就可以了，无需修改算法的其它部分。

正如每个硬币都有两面的性质，遗传算法在带来一定优势的情况下，也会有它缺陷的地方存在，其主要缺点如下：

1. 局部搜索能力不强,很容易过快的陷入局部最优解而使得算法的速度减慢甚至停止。经过人们的大量实验表明，遗传算法存在过早向目标函数的真实解收敛，这种问题被称为遗传算法的“早熟”现象。其中有些是由于目标函数的特性造成的，另外一些则是由于算法设计不当。
2. 太过依赖适应度函数，一旦没有确切的适应度函数，整个算法便失去方向。

尽管遗传算法存在一定的缺陷但经过一段时间的发展，人们已经找到优化遗传算法的方法，常见的改进遗传算法有：与模拟退火算法(SA)的结合使用；在选择策略上有竞争选择、按序挑选等方案；在编码方案上提出实数编码、互换编码、树形编码等方案。这使得遗传算法的发展不但没有停滞不前，反而在生物技术和生物学、函数优化以及组合优化、人工生命、计算智能、机器学习等领域起到至关重要的作用。

本论文采用的也是改进后的遗传算法对问题的求解，详细的介绍分析将在第四章描述。

## 本章小结

本章主要从生物进化论得到的演化计算开始，一步步阐述遗传算法的细节，主要对遗传算法的概念、思想、特点及其应用进行详细的说明，为第四章节改进的遗传算法打下坚实的理论概念基础。

# 改进遗传算法求解非线性方程组