

# 模拟退火算法的动力系统模型及收敛性分析

李元香 项正龙 夏界宁

(武汉大学计算机学院 武汉 430072)

**摘要** 模拟退火算法是经典的拟物类自然计算方法,其算法设计及应用研究取得了丰硕的成果,模拟退火策略也广泛地融入到现代群智能演化算法的研究之中.早期的性能分析和收敛性分析等理论研究主要是基于随机过程中的马尔科夫链理论,获得了依概率意义的收敛性定理.由于物理和数学已经积淀了深厚的理论基础和丰富的分析工具,可以用来进行随机启发式算法的理论分析和设计.该文试图运用动力系统理论分析模拟退火算法的运行机理和收敛性,将算法搜索最优解的过程比拟为质点作弹性运动,算法运行过程中函数值的变化就是质点在作简谐振动或阻尼振动,建立其常微分方程动力系统模型.运用常微分方程的定性理论对该动力系统模型进行求解和分析,证明了模拟退火算法前、中期的局部收敛性和后期的全局收敛性,对其运行机理给出了合理的理论解释.同时,基于建立的动力系统模型,分析了算法衰减因子与收敛速度的关系,得到了模拟退火算法收敛速度的估计.在此基础上,提出了一个模拟退火回火算法的改进策略,一个简单易行的回火时刻判据,当弹性系数趋于很小的值时,即可以当作回火时刻.选取几个典型的测试问题,运用基本的模拟退火算法进行实验验证.首先,实验表明数值收敛曲线与理论分析的收敛性结论相吻合;其次,实验验证了收敛速度随退火温度变化的理论分析与数值实验相吻合;同时实验也验证了提出的回火时刻判据的有效性.最后,理论与实验分析表明该文建立的动力系统模型适合描述模拟退火算法.

**关键词** 模拟退火算法;动力系统;收敛性分析;弹性系数;弹性势能;演化计算

中图法分类号 TP311 DOI号 10.11897/SP.J.1016.2019.01161

## Dynamical System Models and Convergence Analysis for Simulated Annealing Algorithm

LI Yuan-Xiang XIANG Zheng-Long XIA Jie-Ning

(School of Computer Science, Wuhan University, Wuhan 430072)

**Abstract** Simulated annealing algorithm is a classical nature-inspired computational method of imitating physics, which has achieved tremendous results in algorithm design and application researches. Simulated annealing strategy is also widely integrated into researches of modern swarm intelligence evolutionary algorithms. Early theoretical researches for the algorithm on performance analysis and convergence analysis were mainly based on the Markov chain theory in the random process, and obtained the convergence theorem in probability. However, it is difficult to analyze the actual searching behaviors of the algorithm using the Markov chain theory. Also, it's hard to get some improvement strategies that are useful for the algorithm. The main attempt is to analyze the operating mechanism and the convergence of the simulated annealing algorithm by using dynamical system theory in this paper. Due to mathematics and physics have a solid theoretical foundations and a variety of analytical methods, they can be used to analyzing and modeling this class of random heuristic algorithms theoretically, so that directing designation and improvement

of algorithms based on problem characteristics. Therefore, by comparing the searching process of the algorithm to the elastic motion of a particle, the change of the function value during the running of the simulated annealing algorithm can be regarded as the particle doing simple harmonic vibration or damping vibration. So, the dynamical system models of the ordinary differential equation are established for the algorithm according to the principle of elastic mechanics and the algorithm running mechanism. And then, the models are solved and analyzed by using the stability theory of ordinary differential equations. The local convergence of the simulated annealing algorithm in the early-and-mid stage and the global convergence in the late stage are proved, and a rational theoretical explanation of its operating mechanism is stated. Hereafter, based on the dynamical system models, the relationship between the damping coefficient of the model equation and the convergence speed of the algorithm is derived. Furthermore, the convergence speed of the algorithm is inferred, which is associated with the annealing temperature. Based on the theoretical analyses, in response to the drawbacks of the tempering mechanism which usually relied on the experience before, a simple and easy tempering time criterion is proposed in order to improve global searching performance of the simulated annealing algorithm. That is, when the elastic coefficient tends to a small value, it can be used as a tempering point for tempering strategy. Experiments applying the primitive simulated annealing algorithm are implemented in order to verify the theoretical results on several typical test problems. Firstly, these experiments show that the numerical convergence lines coincide with the convergence results of theoretical analysis. Secondly, experiments verify the changing tendency of the convergence speed along with annealing temperature in theory is consistent with the actual numerical experiment. As the annealing temperature decreases, the damping coefficient increases and the attenuation factor decreases, and then the convergence speed increases. Thirdly, the experiments also show the validity of the proposed tempering moment criterion on the benchmark problem. In the end, theoretical and experimental analysis demonstrates that the dynamical system models established in this paper are suitable for describing the simulated annealing algorithm.

**Keywords** simulated annealing algorithm; dynamical system; convergence analysis; elasticity coefficient; elastic potential energy; evolutionary computation

## 1 引 言

模拟退火算法是一个经典的拟物类自然计算方法,它从改进局部搜索算法出发,模拟金属热浴之后的缓慢降温(退火)达到能量最低态(即常温固态)的过程,试图搜索优化问题的全局最优解<sup>[1-2]</sup>. 20 世纪 80 年代以来该算法获得广泛的研究和应用,取得了丰富的成果<sup>[3-7]</sup>. 而由于它属于单点启发式搜索算法,此后,被一类群智能启发式算法所取代,但模拟退火策略常常融入这些群智能算法的设计和应用之中,依然焕发着新的活力<sup>[8-12]</sup>.

在算法的理论分析上,传统的模拟退火算法收敛性分析是将算法的搜索寻优看成随机过程,运用随机过程理论建立算法运行过程的马尔科夫

(Markov)链模型,得到概率意义下的收敛性和时间复杂性<sup>[13-16]</sup>. 文献[17]建立了一种等态等价关系与强/弱态偏序关系模型,用于分析进化算法在收敛性上的等价性与可比性,基于吸收态马尔科夫性,满足等态关系的进化算法具有等价的收敛性,在收敛性意义上实现了进化算法的等价类划分.

而在演化算法的时间复杂度分析上,近年也有许多方法出现. 文献[18]将转移分析方法引入演化算法时间复杂度分析,通过定义距离函数来计算算法每代向最优解前进的平均距离从而获得算法的平均首次到达时间. 文献[19]提出了基于吸收马尔科夫链的演化算法计算复杂度分析框架,将演化算法的时间复杂度度量表达为 Markov 链理论中的首次到达时间. 文献[20]通过研究两个马尔科夫链之间进行对齐映射,提出一种通用的演化算法计算复杂度

分析方法,可称为交替分析法(switch analysis). 总体而言,直到目前对自然与演化类算法的理论分析主要是以随机过程为理论工具,分析的首要任务就是建立算法的马尔科夫链模型,得到概率意义下的收敛性、复杂性度量和收敛速度.

近些年来,又兴起了一种新的自然计算的算法设计与分析方法,就是运用动力系统和热力学与统计力学的理论建立群体演化的动力学模型,基于模型分析群体演化的动力学特征和多样性保持策略. 文献[21]仿照模拟退火算法的退火机制,提出了基于热动力学选择规则的遗传算法,以统计热力学中的自由能极小定律为依据进行种群的选择和替换,实现种群系统的能量和熵之间的动态竞争机制,从而自适应调整算法探测能力(搜索有潜力的解)和开采能力(搜索最优解)之间的均衡. 文献[22]利用自由能极小原理和熵增原理来实现演化群体的寻优机制,建立了基于分子动力学的动力学模型,在此模型的基础上提出了一种近似最优的群体选择和替换规则. 文献[23]利用少量先验知识即可以对数据进行分析,建立了一个多时间尺度动力系统模型,进而提出了一类求解符号建模问题的混合演化算法.

上述思想方法的核心是基于数学和物理学的理论建立动力系统模型,作为算法设计和问题求解的基础,依据问题特征进行模型的具体化和算法设计,这类方法有较好的普适性,能够推广到一般自然计算算法的设计之中. 另一方面,由于数学和物理学有着非常厚实的理论和丰富的方法,这类思想方法还可进一步用于对自然计算算法进行理论分析,本文即是受此启发,运用动力系统理论对模拟退火算法进行建模和理论分析. 将算法搜索寻优过程中目标函数值的变化类比于一个质点与一根弹簧相连而做弹性运动,算法运行过程中函数值的变化就是质点在作简谐振动或阻尼振动,正好适合用常微分方程动力系统描述,于是建立其动力系统模型,然后以模型为基础进行求解、机理分析和收敛性分析,给出模拟退火算法运行机理的理论解释,证明其收敛性定理. 最后理论结果还要通过实验验证,以表明数值实验结果与理论分析结论的吻合性,动力系统模型对于模拟退火算法的适用性.

本文的目的是希望以简单而经典的模拟退火算法为典型例子和出发点,探讨自然计算的数学物理基础和算法设计与分析的动力系统方法. 建立的模型和理论分析结果又可应用于算法的设计之中,以期催生出一类新的算法,发展自然计算的动力系统

方法. 本文第 2 节首先介绍模拟退火算法及其背景知识,然后在此基础上建立算法的常微分方程动力系统模型;第 3 节基于上述模型,运用常微分方程动力系统理论分析模拟退火算法的运行机理和收敛性;第 4 节基于模拟退火算法的 Metropolis 准则,导出弹性系数和弹性势能的表达式,然后结合典型的测试问题对理论结论进行实验验证;第 5 部分是结论与展望.

## 2 动力系统模型

### 2.1 模拟退火算法

金属固体热浴(即溶化)后,分子运动的自由度变大,要再次回到稳定的固态,即能量最小态,必须缓慢降温,称为徐徐退火,否则,快速降温将使金属溶液冷却后处于一个亚稳定的固态. 在某恒定温度  $T$  下模拟金属达到热平衡的过程, Metropolis 提出重要性采样法:用统计物理描述方法,设金属系统当前处于微观状态  $i$ ,其能量为  $E_i$ ,然后分子运动使系统状态变为  $j$ ,能量为  $E_j$ ;如果  $E_j < E_i$ ,则新状态  $j$  作为“重要”状态,如果  $E_j > E_i$ ,则状态  $j$  是否为“重要”状态,要依状态  $j$  和  $i$  的 Boltzmann 分布之比,  $r = \exp(E_i - E_j / k_B T)$  ( $k_B$  称为 Boltzmann 常数)决定. 此时,  $r < 1$ ,随机产生一个  $[0, 1)$  区间的随机数  $\xi$ ,若  $r > \xi$ ,则接受新状态  $j$  为“重要”状态,否则舍弃,当前状态仍为  $i$ . 此即著名的 Metropolis 准则,它模拟的是金属系统分子热运动过程,“重要”状态保持系统热运动的潜在驱动力,而不至于过早地陷于亚稳定状态.

模拟退火算法从改进局部搜索法出发,将优化问题的目标函数比拟为金属系统的能量,于是顺理成章地引入 Metropolis 准则作为新解接受与否的判据. 设有优化问题  $\min_{i \in S} f(i)$ ,其中  $S$  是目标函数  $f(i)$  可行解的集合,若从当前解  $i$  产生了新解  $j$ ,则  $j$  是否接受为当前解由以下转移概率决定:

$$P_i(i \rightarrow j) = \begin{cases} 1, & \text{若 } f(j) < f(i) \\ \exp\left(\frac{f(i) - f(j)}{T}\right), & \text{否则} \end{cases} \quad (1)$$

其中,  $T$  称为控制参数,模拟的是金属系统中的温度. 通常在初始时  $T$  取比较大值  $T_0$  (与热浴温度相对应),然后缓慢递减(相应于徐徐降温)到  $T_k$  ( $k = 1, 2, \dots$ ),在每个  $T_k$  值还要执行  $L_k$  次 Metropolis 准则,再过渡到下一个控制参数  $T_{k+1}$ ,即充分模拟分子热运动,探测具有搜索潜力的区域. 由以上描述给

出模拟退火算法(Simulated Annealing Algorithm, SAA)如下.

### 算法 1. 模拟退火算法(SAA).

输入:  $i_0, T_0, L_0$

输出:  $f(i)$

INITIALIZE ( $i_0, T_0, L_0$ )

BEGIN

$k := 0; i := i_0; L_k := L_0;$

REPEAT

FOR  $n := 1$  to  $L_k$  do

BEGIN

GENERATE ( $j$  from  $i$ );

IF  $f(j) \leq f(i)$  THEN  $i := j$

ELSE IF

$\exp\left(\frac{f(i) - f(j)}{T_k}\right) > \text{random}[0, 1)$

THEN  $i := j$

END;

$k := k + 1;$

CALCULATE ( $T_k$ );

CALCULATE ( $L_k$ );

Until stop criterion

END;

模拟退火算法与局部搜索法的本质区别是它在搜索过程中既接受“好”解,也接受“坏”解,从而有可能跳出局部最优的“陷阱”去搜索更“好”的解.算法中的  $L_k$  ( $k=0, 1, 2, \dots$ ),按照早期的习惯称呼和随机过程理论分析方法,被称为第  $k$  个马尔科夫链的长度.

## 2.2 动力系统描述

模拟退火算法是一个单点启发式搜索过程,如重点关注目标函数值的变化,搜索过程可以看成是一个质点在运动,其运动的空间就是目标函数的取值范围,即一个区间.在控制参数  $T_k$  时,内层循环搜索寻优的过程就是质点在目标函数值区间来回运动的过程,这又可类比为质点与一根假想的弹簧相连,受弹力作用而产生的往复运动.当内层循环到  $L_k$  次结束时,就认为是获得了一个稳定解,弹簧静止,质点也静止.然后更新控制参数为  $T_{k+1}$ ,循环次数更新为  $L_{k+1}$ ,开始新的循环.弹簧也随之更新为另一根,质点开始新一轮的弹性运动.如此周而复始,直到满足停止准则.

基于以上阐述,设控制参数为  $T_k$  时,马尔科夫链的长度为  $L_k$ ,当前解  $i$  的目标函数值  $x = f(i)$ ,表示运动质点的位置坐标,又设此时与质点相连的弹簧的弹性系数为  $\mu_k$ ,质点的质量取为 1,则质点所受的弹力为

$$F_k = -\mu_k x \quad (2)$$

负号表示弹力总是与运动方向相反,于是由牛顿第二定律有动力学方程

$$\frac{d^2 x}{dt^2} + \mu_k x = 0 \quad (3)$$

不失一般性,这里可以假设质点的平衡位置为 0,即目标函数的最小值  $\min_{i \in S} f(i) = 0$ . 方程(3)是一个二阶线性常微分方程,可以解析求解,其解描述的是弹簧作简谐振动<sup>[24]</sup>(见 3.1 节). 对应于 SAA,如果  $L_k \rightarrow \infty$ ,则充分长时间后,搜索将会在某局部极小值附近来回跳动.

导出方程(3)时,没有分析 Metropolis 准则的细节,所以只得到了一个粗略的动力系统方程.为了获得更精细的方程,运用式(1)的转移概率,令

$$r_k = \exp\left(\frac{f(i) - f(j)}{T_k}\right) \quad (4)$$

在同一个控制参数  $T_k$  下,假设  $r_k$  看成一个近似不变的常数.考察 SAA 中内层循环相继的两个搜索步:从第  $n-1$  步到第  $n$  步和从第  $n$  步到第  $n+1$  步.设第  $n-1$  步时,当前目标函数值为  $x_{n-1}$ ,产生的新解为  $x_n$ ,由 Metropolis 准则,第  $n$  步时当前解取值  $x_n$  的概率为  $r_k$ ,取值  $x_{n-1}$  的概率为  $(1-r_k)$ ,于是第  $n$  步当前解的期望值是  $r_k x_n + (1-r_k) x_{n-1}$ . 同理,第  $n+1$  步时当前解的期望值为  $r_k x_{n+1} + (1-r_k) x_n$ . 再者,SAA 中相邻两次循环可看成离散的时间间隔  $\Delta t$ ,也称时间步长.从而方程(3)中第一项对时间的二阶导数运用二阶差商离散化为

$$\Delta = \frac{x_{n+1} - 2x_n + x_{n-1}}{(\Delta t)^2} \quad (5)$$

将  $x_{n+1}$  和  $x_n$  用期望值代替得

$$\Delta = \frac{r_k(x_{n+1} - 2x_n + x_{n-1}) + (1-r_k)(x_n - x_{n-1})}{(\Delta t)^2} \quad (6)$$

再将式(6)右端的差商转换成导数表示,有

$$\Delta = r_k \frac{d^2 x}{dt^2} + \frac{1-r_k}{\Delta t} \frac{dx}{dt} \quad (7)$$

式(7)替换式(3)左端的第一项,又推得一个新的动力学方程

$$r_k \frac{d^2 x}{dt^2} + \frac{1-r_k}{\Delta t} \frac{dx}{dt} + \mu_k x = 0 \quad (8)$$

必须说明的是,我们这里只是形式上的推导,从数学的角度看并不严格,目的是要获得一个更精细的描述模拟退火算法的动力系统模型.方程(8)中还遗留了一个离散时间步长  $\Delta t$ ,它实际上与 SAA 中马尔科夫链的长度  $L_k$  密切相关,如果循环次数  $L_k$  与

一个时间单位  $\tau$  相对应, 则  $\Delta t = \tau/L_k$ , 代入方程(8) 则有

$$\frac{d^2 x}{dt^2} + \eta_k \frac{dx}{dt} + \frac{\mu_k}{r_k} x = 0 \quad (9)$$

其中,  $\eta_k = \frac{1-r_k}{r_k} \frac{L_k}{\tau}$  在弹性力学中称为阻尼系数, 质点运动速度  $dx/dt$  越大, 产生的阻滞力就越大。

联系方程(4)、方程(3)和方程(9)可以大致解释 SAA 的运行过程, 在算法运行初期,  $T_k$  很大,  $r_k$  几乎为 1, 内层循环每步产生的新解几乎被接受, 因此搜索范围大, 而此时方程(9)中第二项的系数  $\eta_k$  几乎为 0, 即变为方程(3)描述的无阻力弹性运动, 即简谐振动。算法运行到后期  $T_k$  缓慢递减,  $0 < r_k < 1$ , 此时, 阻尼系数  $\eta_k$  不为 0, 对弹簧的振动产生阻滞力, 即为方程(9)描述的阻尼振动, 这表明此时算法搜索的范围受限, 以寻优为主。

### 3 动力系统分析

#### 3.1 常微分方程稳定性理论简介

设有初值问题

$$\frac{dx}{dt} = f(t, x), x(t_0) = x_0, x \in G \subset \mathbb{R}^n \quad (10)$$

存在唯一解  $\varphi(t) = x(t, t_0, x_0), t \in (-\infty, +\infty)$ , 若还有  $f(t, x)$  满足  $f(t, 0) = 0$ , 即  $x(t) = 0$  是方程(10)的解, 则称  $x(t) = 0$  为初值问题(10)的零解。需要说明的是, 若问题(10)存在唯一的非零解  $\varphi(t)$ , 则作变化  $y(t) = x(t) - \varphi(t)$  可以得到一个关于  $y(t)$  的初值问题, 而  $y(t)$  具有唯一的零解, 因此, 关于零解稳定性的讨论具有普适意义<sup>[24-25]</sup>。为简便起见, 也不失一般性, 取  $n=1$ , 即  $x$  是一个数值函数而非向量值函数。

**定义 1.** 任意给定的  $\epsilon > 0$ , 都能找到正数  $\delta$  使得当  $|x_0| < \delta$  及  $t > t_0$  时方程(10)的解  $x(t) = x(t, t_0, x_0)$  满足  $|x(t, t_0, x_0)| < \epsilon$ , 则称方程(10)的零解是稳定的。

这个定义表述的是对初始值的稳定性, 即对原初值  $x_0 = 0$  的小于  $\delta$  的扰动而产生的另一个解  $x(t, t_0, x_0)$  与零解也仅差一个任意小的  $\epsilon$ 。

**定义 2.** 是包含零点的开区域, 若对所有  $x_0 \in U$  和任意给定的  $\epsilon > 0$ , 都能找到正数  $T = T(\epsilon, t_0, x_0)$  使得当  $t > t_0 + T$  时方程(10)的解满足  $|x(t, t_0, x_0)| < \epsilon$ , 则称方程(10)的零解是吸引的, 称  $U$  是方程(10)零解的吸引域。

**定义 3.** 方程(10)的零解是稳定的, 吸引的, 则称方程(10)的零解是渐近稳定的; 若方程(10)零解的吸引域是整个求解域  $G$ , 则称方程(10)的零解是全局渐近稳定的。

实际上, 上述三个定义都可以用极限的方式表述, 定义 1 说的是从  $x_0 = 0$  的邻近点出发, 当  $t \rightarrow \infty$  时, 有  $x(t)$  收敛于 0; 定义 2 表明从区域  $U$  内任一点出发都有  $\lim_{t \rightarrow \infty} x(t) = 0$ ; 定义 3 给出了更强的稳定性和收敛性。这几个定义为我们运用动力系统理论分析模拟退火算法的收敛性指明了方向。

#### 3.2 局部收敛性

先求解 2.2 节的动力学方程(3), 若将  $\mu_k$  近似地看成常数, 这是一个二阶线性常微分方程, 其特征方程为

$$\lambda^2 + \mu_k = 0 \quad (11)$$

有两个虚特征根  $\lambda_{1,2} = \pm i \sqrt{\mu_k}$ ,  $i$  为虚数单位。于是方程(3)的通解为简谐振动解

$$x(t) = c_1 \cos(\sqrt{\mu_k} t) + c_2 \sin(\sqrt{\mu_k} t) \quad (12)$$

其中,  $c_1, c_2$  为待定常数, 由初始条件确定, 但这里并不需要其精确值。由式(12)可得  $|x(t)| \leq |c_1| + |c_2|$ , 于是在控制参数  $T_k$  时, 若  $L_k \rightarrow \infty$ , 则  $x(t)$  有若干收敛的子序列, 即  $L_k$  充分长后,  $x(t)$  将在某平衡值  $x_k$  附近振荡。而当  $t \rightarrow \infty$  时, 模拟退火算法收敛于某局部极值, 具体证明如下。

考查外层循环的相邻两次循环, 设控制参数  $T_{k-1}$  时得到的局部极值为  $x_{k-1}$ , 它即是  $T_k$  时内层循环的起点。弹力为有势力, 它做的功只与起止位置有关, 而与运动轨迹无关, 由式(2)得

$$W_k = - \int_{x_{k-1}}^{x_k} \mu_k x dx = \frac{1}{2} \mu_k (x_{k-1}^2 - x_k^2) \quad (13)$$

令  $V_{k-1} = \mu_k x_{k-1}^2 / 2, V_k = \mu_k x_k^2 / 2$ , 则可看成在控制参数  $T_k$  时, 质点运动的开始和终止势能。势能做功转换为质点运动的动能, 故在  $L_k$  次运动后, 可以认为质点处于与参数  $T_k$  相应的一个极小势能状态, 即  $V_{k-1} \geq V_k$ 。于是  $x_{k-1} \geq x_k, \{x_k\}_0^\infty$  是一个单调下降序列, 又由于已假设  $x_k \geq 0$ , 故  $\{x_k\}_0^\infty$  收敛, 即  $\lim_{k \rightarrow \infty} x_k = x^*$ 。但是, 这里不能保证  $x^* = 0$ , 因为方程(3)只是零解稳定, 而非全局稳定, 其极限值依赖于初始值的选取。因此, 只能得到局部收敛的结论, 局部收敛性表明方程(3)刻画的是算法的前期和中期行为, 偏离零解较远的初始值则不能保证收敛到最小值零。

#### 3.3 全局收敛性

要获得全局收敛性, 须运用更精细的动力系统

方程(9)进行分析,如前所述,它刻画了算法的后期行为.为表述简洁令  $\xi_k = \eta_k/2$ ,  $\omega_k^2 = \mu_k/r_k$ ,它们在控制参数  $T_k$ 时可近似看成常数,方程(9)也是一个线性二阶常微分方程,其特征方程为

$$\lambda^2 + 2\xi_k\lambda + \omega_k^2 = 0 \quad (14)$$

特征根为  $\lambda_{1,2} = -\xi_k \pm \sqrt{\xi_k^2 - \omega_k^2}$ ,不论判别式  $D_k = \xi_k^2 - \omega_k^2$ 的符号为正、为零或者为负,根的实部总是负的.用  $-\lambda_1, -\lambda_2$ 表示  $D_k > 0$ 时的根,  $-\lambda$ 表示  $D_k = 0$ 时的根,  $-\alpha \pm \beta i$ 表示  $D_k < 0$ 时的根,此处  $\lambda_1, \lambda_2, \lambda, \alpha$ 都是正数.于是得到三种情形下的通解:

$$(1) x(t) = c_1 e^{-\lambda_1 t} + c_2 e^{-\lambda_2 t}, D_k > 0;$$

$$(2) x(t) = (c_1 t + c_2) e^{-\lambda t}, D_k = 0;$$

$$(3) x(t) = e^{-\alpha t} (c_1 \cos \beta t + c_2 \sin \beta t), D_k < 0;$$

其中的待定常数  $c_1, c_2$ 由初始条件确定.

由以上通解看出,不论哪种情形,从哪个初始点出发,随着时间(循环次数)的增加终将趋于平衡位置,且  $|x(t)|$ 可以任意小,即方程(9)的零解在求解域内是吸引的、全局渐近稳定的.具体而言,在控制参数为  $T_k$ 时,内层循环  $L_k$ 次后,质点处于  $x_k$ ,可能不为零,但由全局稳定性,对  $\forall \varepsilon > 0$ ,只要循环次数(马尔科夫链长度)  $L_k$ 足够大,总可以满足  $|x_k| < \varepsilon$ .另一方面,与前面同理,考查外层循环的相邻循环可知,序列  $\{x_k\}_0^\infty$ 同样是单调下降的,于是当  $k \rightarrow \infty$ 时有  $\lim_{k \rightarrow \infty} x_k = 0$ ,至此我们基于动力学方程(9)证明了模拟退火算法的全局收敛性.

### 3.4 收敛速度分析

从刻画全局收敛性的动力系统方程(9)的三种情形的通解可以看出,  $x(t)$ 的表达式中均有依时间(搜索步数)的衰减因子  $e^{-\alpha t}$ ,  $\sigma = \lambda_1, \lambda_2, \lambda, \alpha$ ,这些参数与一个共同的参数  $\xi_k$ 密切相关,事实上,  $0 < \lambda_1 < \xi_k, \lambda_2 > \xi_k, \lambda = \xi_k, \alpha = \xi_k$ .

为确定起见也不失一般性,仅以第三种形式的解进行分析.首先看到其中的括号部分有上界,即  $|c_1 \cos \beta t + c_2 \sin \beta t| \leq |c_1| + |c_2|$ ,因而衰减因子  $e^{-\alpha t}$ 决定  $x(t)$ 趋于零的快慢.为了得到收敛速度估计,令  $\gamma_k = e^{-\alpha t}$ ,并对其取对数和求时间导数得,

$$\frac{d}{dt}(\ln \gamma_k) = -\alpha \quad (15)$$

由于  $\gamma_k$ 与  $\ln \gamma_k$ 递减的方向和趋势是一致的,因此,式中的  $\alpha$ 可以看成是在控制参数  $T_k$ 时衰减因子  $\gamma_k$ 的衰减速度,也即  $x_k$ 趋于零时收敛速度的估计值.严格地讲,要求  $\gamma_k$ 的衰减速度应对  $\gamma_k$ 求时间导数,但为了简洁,并体现模型方程(9)的本质特征,便

以式(15)作为收敛速度的估计.由2.2节和3.3节可知,  $\alpha = \eta_k/2$ ,  $\eta_k$ 是方程(9)中的阻尼系数,这表明阻尼系数决定着  $x(t)$ 衰减的快慢,也即决定着模拟退火算法的收敛速度.

考察相邻的外层循环,两个控制参数分别为  $T_k$ 和  $T_{k+1}$ ,它们决定着相应内层循环的衰减因子和收敛速度.再看2.2节中式(4)的转移概率  $r_k$ ,其中  $j(f(j) > f(i))$ 为新产生的差解.由于模拟退火算法实际运行时都是从一个较大的“温度”值  $T_0$ 缓慢下降,因此有  $T_k > T_{k+1}$ .又在通常情况下  $T_k$ 的数量级高于  $\Delta f = |f(j) - f(i)|$ 的数量级(因为  $T_k$ 的数量级较低时,模拟退火算法实际上就变成了局部搜索法),因此,  $r_k$ 的数值大小主要由  $T_k$ 决定.于是,由以上分析可知通常有  $r_{k+1} \leq r_k$ ,结合阻尼系数  $\eta_k$ 的表达式有  $\eta_{k+1} \geq \eta_k$ .由此随着外层循环数  $k$ 的增加算法的收敛速度在加快,即  $\alpha_{k+1} \geq \alpha_k$ ,后面的实验也验证了这一点.

这里的讨论表明,算法后期的全局收敛是逐渐加快的.但也必须说明这只是理论分析的结论,实际运行中由于受到问题复杂性和随机因素的影响,有时会陷入局部收敛.因此,在模拟退火算法的研究中一个重要的改进就是回火策略,以跳出局部收敛.后面的实验验证中,结合理论分析给出了一个回火时刻的判据.

## 4 弹性系数与实验分析

### 4.1 弹性系数与弹性势能

由式(13)可知,弹性势能的减少就是弹力在对质点做功,即转化为质点运动的动能,此为机械能守恒原理.下面基于此原理导出弹性系数的一个理论估计公式.

在控制参数  $T_k$ 时,考察内层相邻两次循环,若目标函数值从当前解  $x_{n-1}$ 接受新解变为  $x_n$ ,这就相当于质点在弹簧弹力的作用下产生位移  $\Delta x_n$ ,则有动能  $h_n = \frac{1}{2} \left( \frac{\Delta x_n}{\Delta t} \right)^2$ .在这个  $\Delta t$ 时间内弹力所做的功为  $w_n = \mu_k (\Delta x_n)^2$ ,由机械能守恒原理应有  $w_n = h_n$ ,于是  $\mu_k = 1/2(\Delta t)^2$ .又根据Metropolis准则,第  $n$ 步接受新解  $x_n$ 的概率为  $r_k$ ,不接受新解的概率为  $(1-r_k)$ ,此时弹力未做功,由此有弹性系数期望值

$$\mu_k = \frac{r_k}{2(\Delta t)^2} \quad (16)$$

根据式(16),还可以用  $L_k$ 次循环的平均值作为

在控制参数  $T_k$  时的弹性系数, 即

$$\bar{\mu}_k = \frac{1}{L_k} \sum_{n=1}^{L_k} \mu_k \quad (17)$$

由式(17)给出的弹性系数公式, 可以得到在控制参数  $T_k$  时的弹性势能为

$$\bar{V}_k = \frac{1}{2} \bar{\mu}_k x_k^2 \quad (18)$$

这里还需说明一下离散时间步长  $\Delta t$ , 在 2.2 节给出了  $\Delta t = \tau/L_k$ , 其中  $\tau$  为与  $L_k$  个循环步相应的一个时间单位. 在统计物理和分子运动论中  $\tau$  称为弛豫时间, 指的是一个物理系统从偏离平衡态再回到平衡态(称为系统的弛豫过程)所需的时间. 类比过来, 模拟退火算法在每个控制参数  $T_k$  时进行的  $L_k$  次循环就是模拟这种弛豫过程, 获得在控制参数  $T_k$  时的平衡解.

显然, 在真实的物理世界中弛豫时间  $\tau$  与具体的物理系统相关, 而在模拟退火算法中就只与具体求解的问题相关. 从算法的求解效率考虑, 结合前面 3.3 节和 3.4 节的全局收敛性证明与收敛速度分析, 在控制参数  $T_k$  下, 给出合适的弛豫时间  $\tau_k$ , 进而估计恰当的马尔科夫链长度  $L_k$ , 可提升实际求解时的全局收敛性和收敛速度. 但如何合理选取  $\tau_k$  和  $L_k$  与问题密切相关, 这是本文的后续工作, 将另文结合具体问题的求解进行讨论, 将理论分析的结果运用于指导和改进算法的设计, 提高求解效率.

本文为突出收敛性分析的主题, 而又无损问题的本质, 在理论分析中只认定  $\tau_k$  和  $L_k$  为常数即可, 在下面实验验证中取  $\tau = L_k$ , 即  $\Delta t = 1$ .

## 4.2 收敛性验证

前面的推导和阐述仅仅是理论上的, 提出的动力系统模型和理论分析结论是否适用于模拟退火算法, 必须进行算法的实验验证. 实验验证的目的其一是看实验结果与理论分析的结论是否吻合, 以表明理论模型、理论分析方法和结论的适用性, 其二也希望通过实验获得进一步改进和推广的思路, 包括对理论模型和理论方法的改进, 以及对算法的改进.

### 4.2.1 测试函数与参数设置

本小节选取五个典型的测试问题进行实验验证, 其中前四个为数值优化问题. 前两个分别是经典测试函数 Rosenbrock 函数和 Rastrigin 函数, 记为  $f_1$  (Rosenbrock) 和  $f_2$  (Rastrigin), 两个函数的最优值均为 0; 后两个测试函数为 CEC2014<sup>[26]</sup> 中两个测试函数  $F_3$  (Rotated Discus Function) 以及  $F_4$  (Shifted and Rotated Rosenbrock's Function), 两个函数的

最优值分别为 300、400. 表 1 给出了它们的表达式和自变量取值范围.

表 1 数值优化测试问题

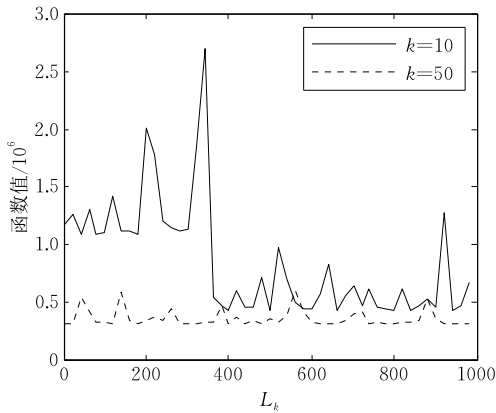
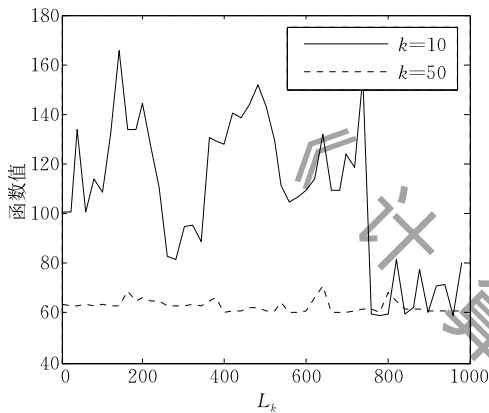
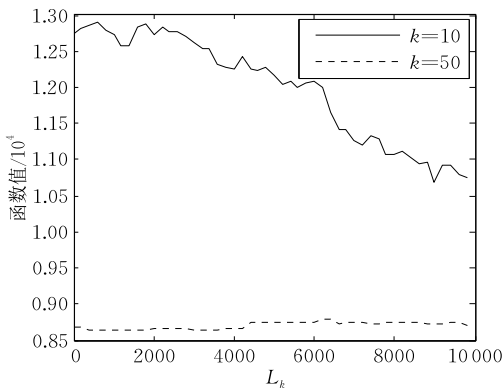
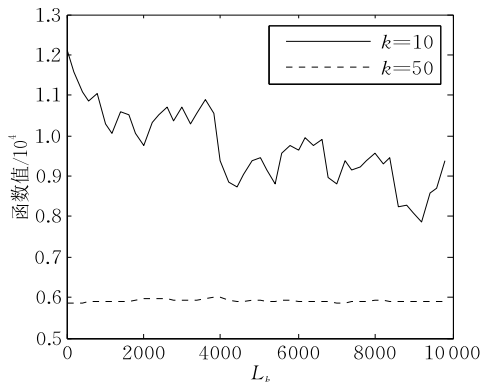
函数	表达式	变量取值范围
$f_1$	$\sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2$	$[-10, 10]$
$f_2$	$\sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	$[-5.12, 5.12]$
$F_3$	$f_3(M(x - o_3)) + F_3^*$	$[-100, 100]$
$F_4$	$f_4\left(M\left(\frac{2.048(x - o_4)}{100}\right) + 1\right) + F_4^*$	$[-100, 100]$

另一个是旅行商问题(TSP). 作为非数值优化测试问题, 它是最典型的组合优化问题, 也是著名的 NP 难题. 简单描述如下: 设有  $N$  个城市和距离矩阵  $\mathbf{D} = [d_{ij}]$ , 其中  $d_{ij}$  表示城市  $i$  到城市  $j$  的距离,  $i, j = 1, \dots, N$ , 问题是要找到遍访每个城市恰好一次的一条回路, 且其路径长度为最短, 此即目标函数, 记为  $f_5$  (TSP).

实验中运用 2.1 节描述的基本模拟退火算法, 其参数根据具体测试函数设置如下:  $f_1$  的初始控制参数设为  $T = 1.00\text{E} + 06$ ,  $f_2$  为  $T = 200$ ,  $F_3$  的参数设置为  $T = 1.00\text{E} + 09$ ,  $F_4$  为  $T = 3000$ , 四个数值函数的马尔科夫链长度均设置为  $L_k = 1000$ ,  $f_5$  设置初始控制参数为  $T = 1000$ , 马尔科夫链长度为  $L_k = 10000$ , 算法在测试函数上的运行终止条件根据实际问题选取与问题最小值相同的数量级均可, 本实验均取为  $T_{\text{min}} = 1\text{e} - 3$ . 所有测试问题的退火速率均为  $\beta = 0.95$ , 即  $T_{k+1} = \beta T_k$ . 取函数  $f_1, f_2$  的维数均为  $n = 20$ ,  $F_3, F_4$  的维数均为  $n = 10$ . 对四个数值优化问题, 其新解随机扰动产生方式采用文献[27]中每次随机对一个分量产生扰动的策略; 对 TSP 问题  $f_5$ , 分别取 50 和 100 个城市进行实验验证, 城市的坐标在  $[0, 1000] \times [0, 1000]$  的正方形内随机产生. 对每个测试问题均在上述参数设置下算法独立运行 30 次, 取 30 次运行结果的平均值作为求解结果.

### 4.2.2 实验结果与分析

图 1 和图 2 显示的是模拟退火算法求解测试问题  $f_1$  (Rosenbrock) 和  $f_2$  (Rastrigin) 的前、中期搜索性态, 图中  $k$  表示外层循环次数, 纵轴表示函数值, 横轴  $L_k$  是马尔科夫链的长度, 即内层循环次数. 以下图 3、图 4 中的图例说明同上, 图 3 和图 4 显示的是求解测试问题  $f_5$  (TSP) 的前、中期搜索性态, 图 3 和图 4 分别显示对城市数为  $N = 50$  和  $N = 100$  两个实例的求解.

图 1 数值优化问题  $f_1$  的前期搜索轨迹图 2 数值优化问题  $f_2$  的前期搜索轨迹图 3 TSP 问题  $f_5$  在  $N=50$  的前期搜索轨迹图 4 TSP 问题  $f_5$  在  $N=100$  的前期搜索轨迹

从图 1、图 2 和图 3、图 4 可以看出,虽然求解的问题不同,但是图中反映的搜索性态基本一致,图中实曲线和虚曲线分别描绘的是外层循环次数为  $k=10$  和  $k=50$  时内层循环  $L_k$  次的搜索轨迹。 $k=10$  时是算法运行的前期,图中看到内层循环的前阶段曲线波动较大,而后渐渐趋于稳定,到了循环的后阶段曲线下降,而且波动变小,只是在某个平衡值附近振荡。 $k=50$  时,算法由前期向中后期过渡,整个内层循环的搜索曲线波动不大,搜索过程只在某个平衡值附近振荡。这种运行性态正是 3.2 节依据方程(3)所论证的局部收敛性,因此实验结果验证了理论分析的结论。

图 5~图 10 分别描绘了模拟退火算法求解五个测试问题的函数值收敛曲线、弹性系数曲线和弹性势能曲线,图中的横轴是外层循环的次数  $k$ ,即整个算法的迭代步数。同样,虽然求解的问题不同,但是图中反映的算法的收敛行为基本一致。

先看弹性势能曲线,虽说有些波动但整体趋势是下降的,这就印证了在第 3 节中基于弹性力学中的极小位(势)能原理而得到势能下降的前提条件,从而推得函数值序列单调下降的结论,然后证明算法的收敛性。再看函数值收敛曲线,事实上,6 幅图中均显示函数值收敛曲线与弹性势能曲线的趋势是一致的,这是因为理论上弹性势能的单调下降才能保证函数值序列的单调下降。虽然具体的数值实验受到算法中随机性的影响,得不到如理论推导的单调下降序列,但从微积分中的极限理论知道,只要有单调下降的子序列就能保证算法的收敛性,图中可以看出这样的子序列肯定存在。因此,根据实验结果的分析,理论推导中的条件能够得到满足。最后看弹性系数曲线,虽然有些小的波动,但整体上呈下降趋势。弹性系数下降从理论上讲,体现的是动力系统模型方程(3)向模型方程(9)的转变,即从简谐振动逐步转变为阻尼振动;从算法运行行为上看,体现的是其搜索行为从前期的“探测”(搜寻有发展潜力的可行解)向后期的“开采”(搜寻最优解)转变。

从实验结果图可以看出,每个问题的三条曲线的走势基本一致,实际上反映了对同一问题从不同角度去观察所得到的结果,它们之间有着必然的内在联系。函数值收敛曲线呈收敛态势,但可能是收敛到某个局部最优值,因为这里使用的是没有任何改进的基本模拟退火算法,而且,理论上的收敛性定理也不能等同于算法实际运行时的收敛状态。



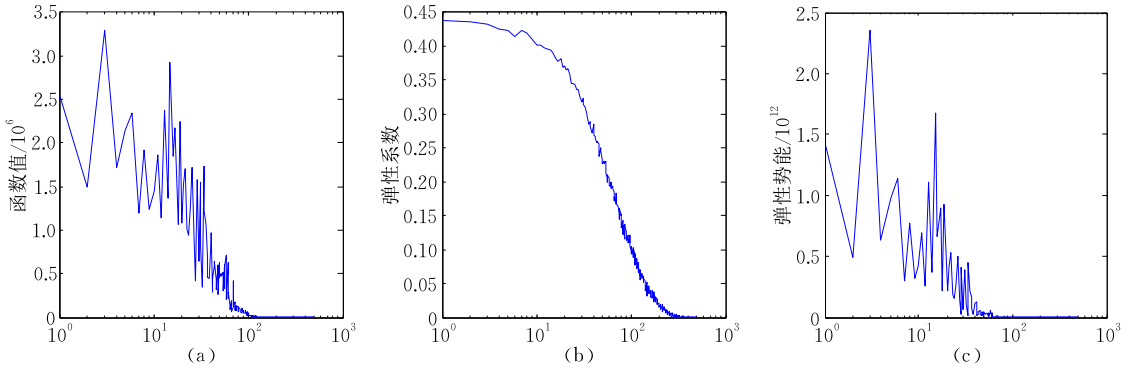


图 5  $f_1 (n=20)$  的收敛曲线、弹性系数曲线和弹性势能曲线

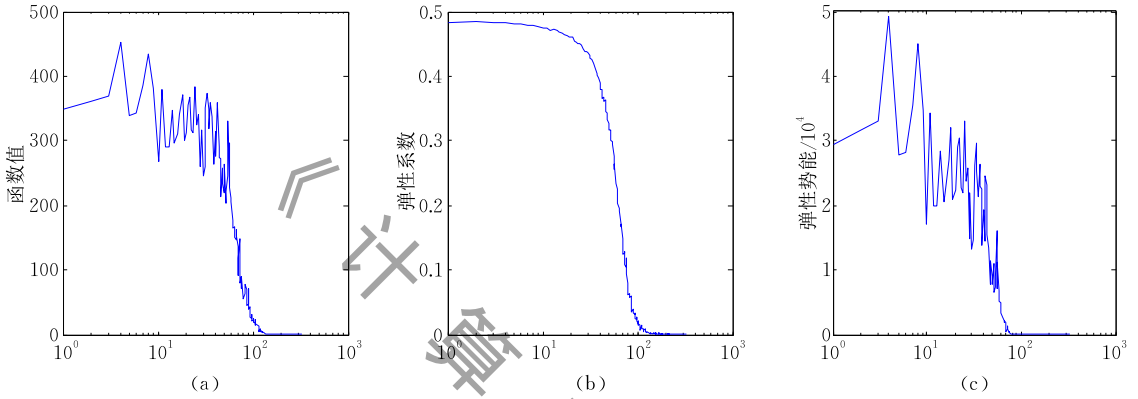


图 6  $f_2 (n=20)$  的收敛曲线、弹性系数曲线和弹性势能曲线

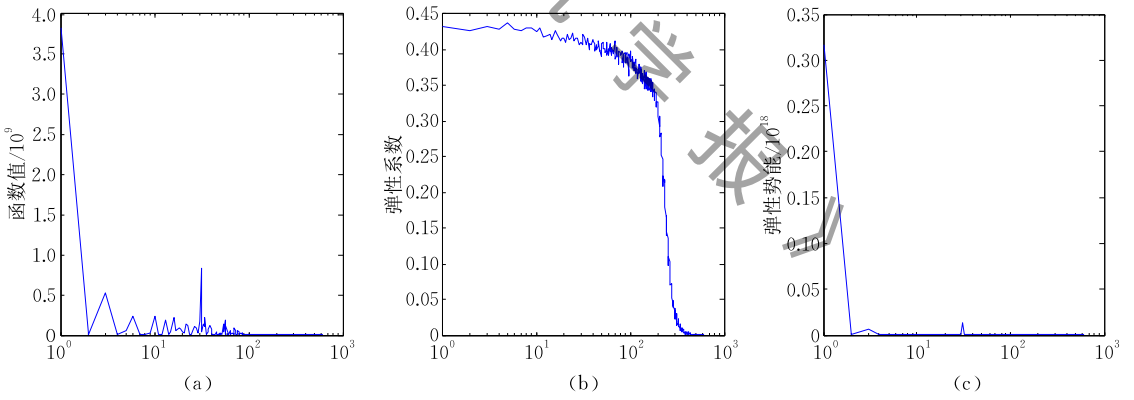


图 7  $F_3 (n=10)$  的收敛曲线、弹性系数曲线和弹性势能曲线

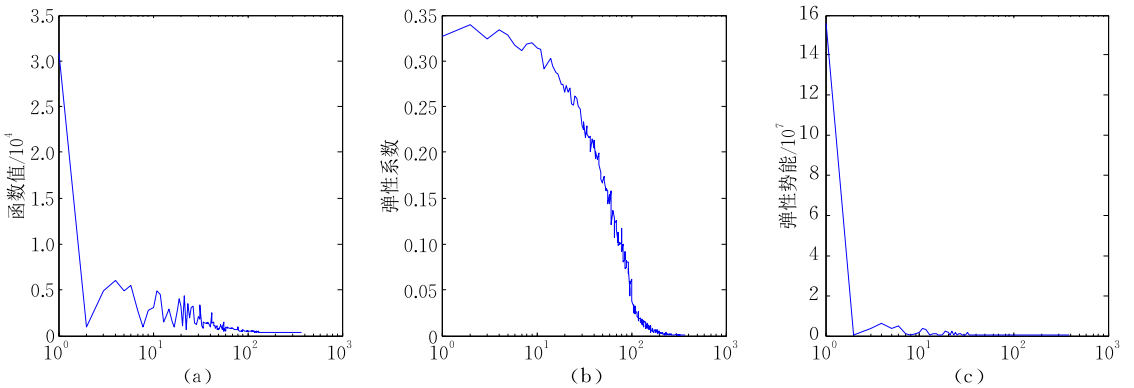
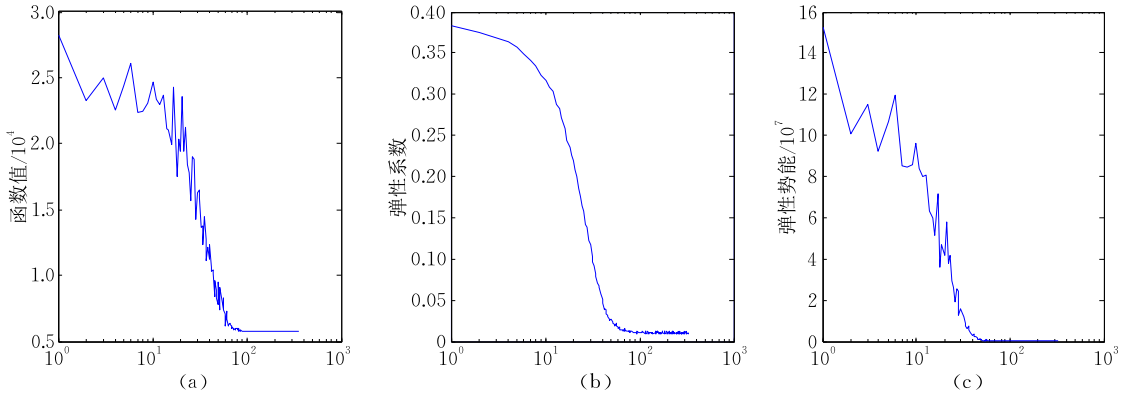
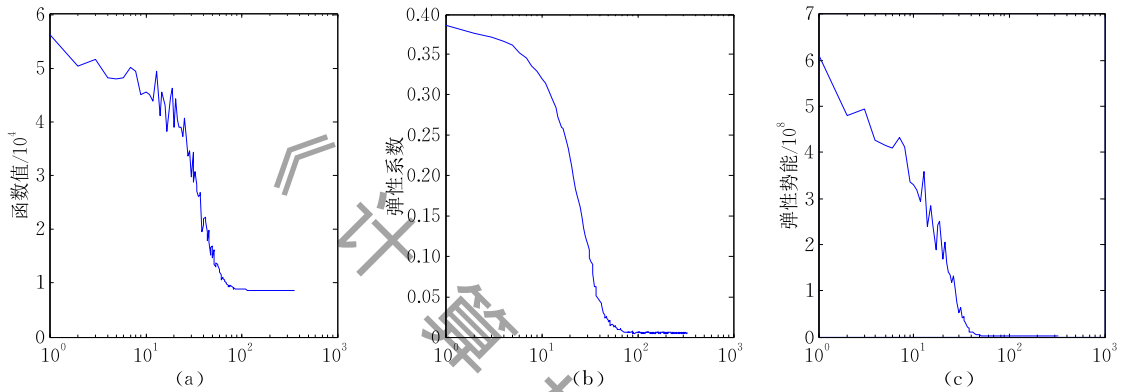


图 8  $F_4 (n=10)$  的收敛曲线、弹性系数曲线和弹性势能曲线

图 9  $f_5 (N=50)$  的收敛曲线、弹性系数曲线和弹性势能曲线图 10  $f_5 (N=100)$  的收敛曲线、弹性系数曲线和弹性势能曲线

在模拟退火算法的研究和发展之中有一个非常重要的改进,就是退火回火算法,改进策略是在算法运行陷入局部收敛时再升高“温度” $T$ ,即称“回火”,以跳出局部收敛,重新取势再次退火搜索,如此反复多次.然而,怎样判断算法的运行陷入了局部收敛呢?对比实验结果中的三条曲线,可以得到一个简单易行的判据,那就是弹性系数的估算值,在弹性系数趋于 0 的某个时刻即可采用回火策略.

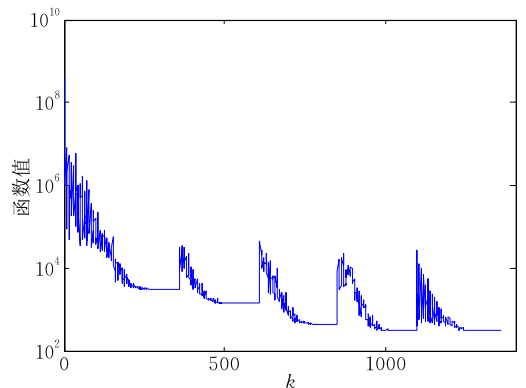
#### 4.3 回火退火与收敛速度

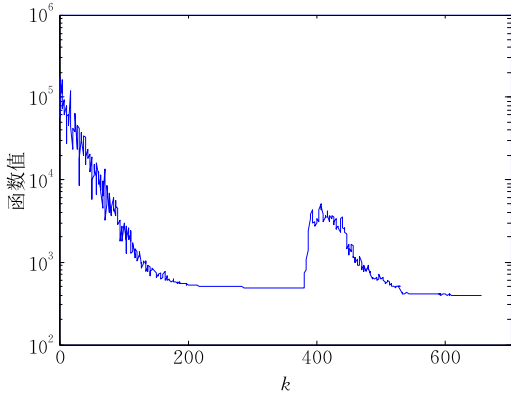
前已说明函数值曲线呈收敛性态不一定意味着全局收敛,往往可能是局部收敛,因此在实际问题的求解中经常运用回火策略来增强全局收敛性.但是,如何确定回火时刻通常是基于数值实验来获得一些经验依据,如比较相继两次外层循环后函数值的变化情况,这就需要回火之前多进行一次或若干次外层循环,增加了较多的额外计算开销.前已述及本文提出的弹性系数可作为退火时刻判据,且无需过多的计算开销,只要对第  $k$  次循环的  $L_k$  个转移概率做累加平均即可.

对比分析图 5~图 10 中每组图的三条曲线,可以看出当弹性系数小于 0.02 时,函数值曲线均显示算法接近局部收敛状态,因此可以选取弹性系数

$\bar{\mu}_k < 0.02$  的某个时刻采取回火策略.本小节的实验选取当弹性系数  $\bar{\mu}_k < 0.01$  时的外层循环  $k$  作为回火时刻点,选取 CEC2014 中的  $F_3 (n=10)$ 、 $F_4 (n=30)$  作为回火测试函数.根据测试函数不同以及实际经验,此处选取的“回火温度”设置为原始“温度” $T_0$  的  $1/10$ .为了充分利用当前解的良好信息,回火退火过程中适当缩小邻域搜索步长,本节的实验中,回火退火过程取邻域搜索步长为原始步长的  $1/10$  进行新解的生成.

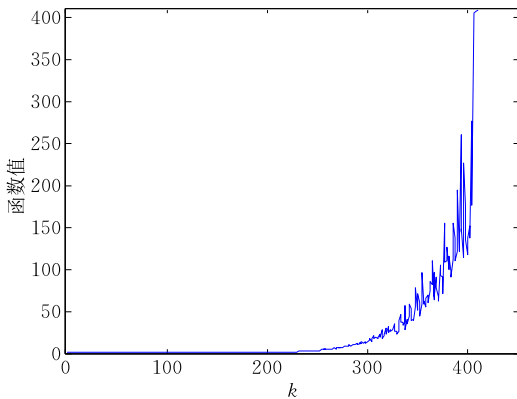
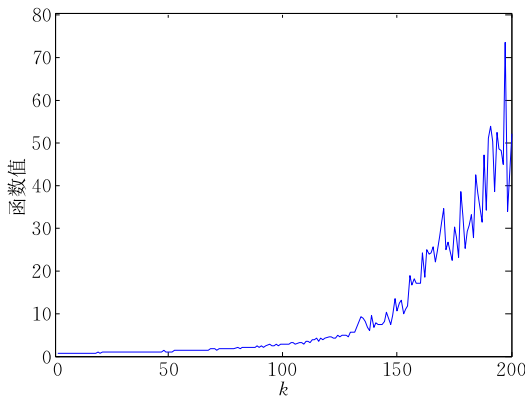
图 11 和图 12 分别为函数  $F_3 (n=10)$ 、 $F_4 (n=30)$  的回火退火收敛曲线,其回火时刻的判据均是

图 11  $F_3 (n=10)$  4 次回火退火曲线

图 12  $F_4(n=30)$  1 次回火退火曲线

$\bar{\mu}_k < 0.01$ . 从图上可以看到  $F_3$  的函数值曲线经过 4 次回火退火到达全局最优值 300 附近 (误差  $2.4732E-3$ ),  $F_4$  经过 1 次回火退火即可收敛到全局最优值 400 附近 (误差  $3.051E-3$ ). 实验结果表明本文提出根据弹性系数的估计值确定回火时刻简单易行, 能给出恰当的回火的时刻, 而且获得了较好的回火退火效果, 算法能较快地收敛到全局最优解附近.

此外, 本小节还运用函数  $F_3(n=10)$ 、 $F_4(n=30)$  进行了收敛速度的统计 (未运用回火策略), 图 13 和图 14 分别为函数  $F_3(n=10)$ 、 $F_4(n=30)$  收

图 13  $F_3(n=10)$  收敛速度曲线图 14  $F_4(n=30)$  收敛速度曲线

敛速度曲线图, 纵坐标为收敛速度估计值, 横坐标为外层循环次数  $k$ . 图中虽然有些波动, 但趋势表明, 随着算法外层循环次数  $k$  的增加, “温度”  $T_k$  下降, 算法的收敛速度在加大, 直至算法收敛, 当然这里也可能是局部收敛. 于是, 这里的实验结果验证了 3.4 节中理论分析的结论.

根据式 (15) 以及 3.4 节的分析, 收敛速度取决于阻尼系数  $\eta_k$ , 它随着“温度”的下降而增加, 收敛速度  $\alpha$  也在增加, 而在算法的收敛过程中  $x(t)$  的衰减因子为  $\gamma_k = e^{-\alpha t}$  在减小, 因此, 理论分析及其实验验证表明, 本文提出的收敛速度的估计具有合理性, 可以度量模拟退火算法的收敛速度.

本节实验分析的目的就是对理论分析工作的验证, 包括建立的模型、理论上提出的前提条件、分析结论以及伴随模型而导出的参数的作用和意义. 具体而言, 进行了两方面的验证工作: (1) 收敛性验证, 这方面的工作表明了动力系统模型的适用性, 理论分析方法的可行性, 收敛行为分析的合理性与结论可靠性; (2) 导出参数的验证, 这方面的工作表明了与模型方程 (3) 相伴的弹性系数能作为模拟退火回火算法中回火时刻的判据, 且简单易行; 同时也表明了模型方程 (9) 中的阻尼系数与收敛速度的关系.

其实, 这里的实验分析本质上就是对动力系统方程 (3) 和动力系统方程 (9) 的验证, 它们分别刻画了算法前后阶段的收敛性态, 而各自参数的数值变化又具体描绘了收敛性态的转变过程. 注意到弹性系数  $\mu_k$  和阻尼系数  $\eta_k$  的表达式, 它们直接与转移概率  $r_k$  相关, 因此, 这两个方程体现了模拟退火算法的基本特征.

## 5 结论与展望

模拟退火算法的核心思想源于统计热力学, 基于 Boltzmann 分布的 Metropolis 准则是其灵魂, 它驱动和控制算法的运行. 而本文基于弹性力学原理建立其动力系统模型, 对一个“老”算法, 从一个“新”的角度诠释其内涵. 从前面的理论分析和实验分析结果可知, 这类模型适合于描述模拟退火算法, 能够刻画它的运行性态. 并且, 由理论与实验分析相结合, 得到了一个简单易行的回火时刻判据, 可用于回火退火算法之中, 关于此点本文中只是做了初步的实验验证与分析. 前已述及, 回火是增强算法全局收敛性的重要策略, 此后, 还将结合收敛速度与时间复杂性分析进行多方面的理论与实验分析, 是本文后

续工作的一部分. 这里有必要澄清一个疑问, 就是 2.2 小节从方程(3)推得方程(9)的过程. 原则上, 建模是假设与演绎推导的结合, 数学上的严密性不是这里必需的, 重要的是看建立的模型是否适用. 那里作形式推导的目的是为了沟通方程(3)与方程(9)之间的联系, 同时也便于进行循序渐进的理论分析.

本文以模拟退火算法为典型例子, 进行动力系统建模与分析, 获得了有价值的理论和实验分析结果, 本文的后续工作还将对收敛速度和时间复杂性进行更细致的理论与实验分析. 而后, 希望在此基础上逐步推广, 对以群体进化为特征的群智能演化算法进行动力系统建模与分析, 此其一; 其二是在算法动力学行为分析的基础上, 从理论与实验两方面获得对算法进行改进的思路与方法, 或者以此为契机发展出新的算法; 由此, 探讨自然计算的数学物理基础, 发展自然计算算法设计与分析的动力系统理论与方法.

### 参 考 文 献

- [1] Dowsland K A, Thompson J M. Simulated annealing. *Handbook of Natural Computing*. Berlin, Germany: Springer, 2012, 1623-1655
- [2] Xie Yun. A summary on the simulated annealing algorithm. *Application Research of Computers*, 1998, 15(5): 7-9 (in Chinese)  
(谢云. 模拟退火算法综述. *计算机应用研究*, 1998, 15(5): 7-9)
- [3] Fu Wen-Yuan, Ling Chao-Dong. Brownian motion based simulated annealing algorithm. *Chinese Journal of Computers*, 2014, 37(6): 1301-1308 (in Chinese)  
(傅文渊, 凌朝东. 布朗运动模拟退火算法. *计算机学报*, 2014, 37(6): 1301-1308)
- [4] Sun S, Zhuge F, Rosenberg J, et al. Learning-enhanced simulated annealing: Method, evaluation, and application to lung nodule registration. *Applied Intelligence*, 2008, 28(1): 83-99
- [5] Xavier-de-Souza S, Suykens J A K, Vandewalle J, et al. Coupled simulated annealing. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2010, 40(2): 320-335
- [6] Geng X, Chen Z, Yang W, et al. Solving the traveling salesman problem based on an adaptive simulated annealing algorithm with greedy search. *Applied Soft Computing*, 2011, 11(4): 3680-3689
- [7] Suman B, Kumar P. A survey of simulated annealing as a tool for single and multiobjective optimization. *Journal of the Operational Research Society*, 2006, 57(10): 1143-1160
- [8] García-Martínez C, Lozano M, Rodríguez-Díaz F J. A simulated annealing method based on a specialized evolutionary algorithm. *Applied Soft Computing*, 2012, 12(2): 573-588
- [9] Bandyopadhyay S, Saha S, Maulik U, et al. A simulated annealing-based multiobjective optimization algorithm: AMOSA. *IEEE Transactions on Evolutionary Computation*, 2008, 12(3): 269-283
- [10] Smith K I, Everson R M, Fieldsend J E, et al. Dominance-based multiobjective simulated annealing. *IEEE Transactions on Evolutionary Computation*, 2008, 12(3): 323-342
- [11] Mafarja M M, Mirjalili S. Hybrid whale optimization algorithm with simulated annealing for feature selection. *Neurocomputing*, 2017, 260: 302-312
- [12] Torkaman S, Ghomi S M T F, Karimi B. Hybrid simulated annealing and genetic approach for solving a multi-stage production planning with sequence-dependent setups in a closed-loop supply chain. *Applied Soft Computing*, 2018, 71: 1085-1104
- [13] Lundy M, Mees A. Convergence of an annealing algorithm. *Mathematical Programming*, 1986, 34(1): 111-124
- [14] Romeo F, Sangiovanni-Vincentelli A. A theoretical framework for simulated annealing. *Algorithmica*, 1991, 6(1-6): 302
- [15] Aarts E, Korst J. *Simulated Annealing and Boltzmann Machines*. New York, USA: John Wiley and Sons Inc., 1988
- [16] Granville V, Krivánek M, Rasson J P. Simulated annealing: A proof of convergence. *IEEE transactions on pattern analysis and machine intelligence*, 1994, 16(6): 652-656
- [17] Huang Han, Lin Zhi-Yong, Hao Zhi-Feng, et al. Convergence analysis and comparison of evolutionary algorithms based on relation model. *Chinese Journal of Computers*, 2011, 34(5): 801-811 (in Chinese)  
(黄翰, 林智勇, 郝志峰等. 基于关系模型的进化算法收敛性分析与对比. *计算机学报*, 2011, 34(5): 801-811)
- [18] He J, Yao X. Drift analysis and average time complexity of evolutionary algorithms. *Artificial Intelligence*, 2001, 127(1): 57-85
- [19] He J, Yao X. Towards an analytic framework for analysing the computation time of evolutionary algorithms. *Artificial Intelligence*, 2003, 145(1-2): 59-97
- [20] Yu Y, Qian C, Zhou Z H. Switch analysis for running time analysis of evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 2015, 19(6): 777-792
- [21] Mori N, Yoshida J, Tamaki H, et al. A thermodynamical selection rule for the genetic algorithm//*Proceedings of the IEEE International Conference on Evolutionary Computation*. Perth, Australia. 1995, 1: 188
- [22] Li Y X, Zou X F, Kang L S, et al. A new dynamical evolutionary algorithm based on statistical mechanics. *Journal of Computer Science and Technology*, 2003, 18(3): 361-368
- [23] Cornforth T W, Lipson H. A hybrid evolutionary algorithm for the symbolic modeling of multiple-time-scale dynamical systems. *Evolutionary Intelligence*, 2015, 8(4): 149-164

- [24] Teschl G. Ordinary Differential Equations and Dynamical Systems. Providence, USA: American Mathematical Society, 2012
- [25] Ma Zhi-En, Zhou Yi-Cang, Li Cheng-Zhi. Qualitative Theory and Stability Theory of Ordinary Differential Equations. Beijing: Science Press, 2001 (in Chinese)  
(马知恩, 周义仓, 李承治. 常微分方程稳定性与稳定性方法. 北京: 科学出版社, 2001)
- [26] Liang J J, Qu B-Y, Suganthan P N. Problem definitions and evaluation criteria for the CEC 2014 special session and competition on single objective real-parameter numerical optimization. Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou: Technical Report 201311, 2013
- [27] Jin Li-Xia, Tang Huan-Wen, Li Bin, et al. A simulated annealing algorithm for continuous functions and its convergence properties. Mathematica Numerica Sinica, 2005, 27(1): 19-30 (in Chinese)  
(靳利霞, 唐焕文, 李斌等. 一类连续函数模拟退火算法及其收敛性分析. 计算数学, 2005, 27(1): 19-30)



**LI Yuan-Xiang**, Ph. D., professor. His main research interests include evolutionary computation and parallels computation.

**XIANG Zheng-Long**, Ph. D. candidate. His main research interests include evolutionary computation and search-based software engineering.

**XIA Jie-Ning**, Ph. D. candidate. His main research interests include evolutionary computation and parallels computation.

## Background

Simulated annealing algorithm, proposed by Metropolis in 1953, has been achieved many results in algorithm design and application researches. Currently, with the decades of research and development, the theoretical researches for the algorithm on performance analysis and convergence analysis are based primarily on the Markov chain theory in the random process, which only obtains the convergence theorem in probability. It is hardly to get the interpretation of the operating mechanism and deduce instructional improved strategy. Due to mathematics and physics which have robust theories and rich methods, modern complex dynamical systems such as dynamical systems and statistical mechanics are compatible with the evolutionary dynamics of evolutionary algorithms. These methods can be implemented to the theoretical analysis of nature inspired algorithms.

In this paper, the dynamic system theory is utilized to model and analyze the simulated annealing algorithm. Thus, by comparing the searching process of the algorithm to the procedure which a particle doing the elastic motion, the

dynamical system models of the ordinary differential equation are established for the algorithm. Solving and analyzing the model by using the properties of ordinary differential equations, the convergence verification of the algorithm and the operating mechanism are obtained. Additionally, a simple and easy “tempering” strategy is given. The numerical experiments show that the dynamical system models established in this paper are appropriate for describing the simulated annealing algorithm.

The work is supported by the National Nature Science Foundation of China (No. 61672391). The project is intended to study the mathematical-physical models for nature inspired computation and to design algorithms of nature inspired computation by using theory and methods of mathematics and physics. Up to now, the project team has proposed some useful models which derived meaningful results for analyzing evolutionary algorithms. And progress is partly made for this project.