

相关性粒子群优化模型*

申元霞^{1,2,3}, 王国胤^{1,2+}, 曾传华³

¹(西南交通大学 信息科学与技术学院, 四川 成都 610031)

²(重庆邮电大学 计算机科学与技术研究所, 重庆 400065)

³(重庆文理学院 计算机学院, 重庆 402160)

Correlative Particle Swarm Optimization Model

SHEN Yuan-Xia^{1,2,3}, WANG Guo-Yin^{1,2+}, ZENG Chuan-Hua³

¹(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China)

²(Institute of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

³(School of Computer Science, Chongqing University of Arts and Sciences, Chongqing 402160, China)

+ Corresponding author: E-mail: wanggy@ieee.org, http://www.cqupt.edu.cn

Shen YX, Wang GY, Zeng CH. Correlative particle swarm optimization model. *Journal of Software*, 2011, 22(4): 695-708. <http://www.jos.org.cn/1000-9825/3728.htm>

Abstract: In the study of particle swarm optimization, properly using the individual experience and social sharing information of particles has always been a problem. To solve this problem, this paper analyzes the random factors in updating the velocity equation in the view of cognition and creates the intrinsic cognitive relation between individual experience and social sharing information. First, a correlative particle swarm optimization model is developed, which uses the Copula function to measure the dependence among random factors. In the new model, the different correlation structures and degrees of correlation between random factors can denote different strategies, which are used to process individual experience and social sharing experience. Meanwhile, this paper provides a flowchart of the correlative particle swarm optimization model, based on Gaussian Copula. Second, the relationship between the degrees of correlation and population diversity is presented, which shows that the random factors with positive linear correlation avail to maintain population diversity. Finally, the relationship between the degrees of correlation and convergence is analyzed and the convergence conditions of the correlative particle swarm optimization model are provided. Experimental simulations show that the correlation of random factors have a much greater influence on the performance of the new model, which can greatly improve convergence velocity and precision when the random factors are a completely positive linear correlation.

Key words: PSO (particle swarm optimizer); correlation; Copula; convergence

摘要: 在粒子群优化算法中,粒子如何合理地利用自身经验信息和群体共享信息的问题一直未能有效解决.针对这一问题,基于认知论的观点,对速度更新公式中的随机因子进行了分析,建立了粒子对自身经验信息和群体共享

* 基金项目: 国家自然科学基金(60773113); 重庆市杰出青年科学基金(2008BA2041); 重庆市自然科学基金重点项目(2008BA2017)

收稿时间: 2009-03-26; 修改时间: 2009-07-06; 定稿时间: 2009-08-26

信息认知的内在联系,提出了相关性粒子群优化模型.该模型采用 Copula 函数去刻画随机因子间的相关结构,而不同的相关结构和相关性程度反映了粒子对自身经验信息和群体共享信息的利用策略的差异,同时给出了基于 Gaussian Copula 的相关性粒子群优化模型的实现方法.理论上给出了随机因子间相关程度与群体多样性的关系式,表明了当随机因子间正线性相关时有利于维持群体的多样性.证明了随机因子间相关程度与算法收敛性的关系,同时给出了相关性粒子群优化模型的收敛条件.仿真实验结果表明,随机因子间相关程度的水平设置对模型的优化性能有非常显著的影响,当粒子的自身经验信息和群体共享信息被同等利用时,模型表现出优良的整体性能.

关键词: 粒子群优化;相关性;Copula;收敛性

中图法分类号: TP183 文献标识码: A

粒子群优化(particle swarm optimizer,简称 PSO)^[1]是一种基于群体智能的优化算法,受到人工生命和社会心理学的启发,由美国社会心理学家 Kennedy 和电气工程师 Eberhart 于 1995 年提出,其基本概念源于对鸟群和鱼群捕食行为的研究.由于 PSO 算法具有概念简单、实现容易、不受问题解空间限制性假设的约束等优点,使其在电力系统、机器学习、模式识别、图像处理等领域得到了广泛的应用^[2].PSO 算法的第 i 个粒子在第 d 维空间的运动方程,可以由一组差分方程来描述,如公式(1)、公式(2)所示:

$$V_{id}(t+1)=\omega V_{id}(t)+c_1 r_1(t)(P_{id}(t)-X_{id}(t))+c_2 r_2(t)(P_{gd}(t)-X_{id}(t)) \quad (1)$$

$$X_{id}(t+1)=X_{id}(t)+V_{id}(t+1) \quad (2)$$

其中: X_{id} 和 V_{id} 分别代表第 i 个粒子第 d 维的位置与速度分量, $|V_{id}| \leq V_{\max}$; P_{id} 为粒子个体所经历的历史最佳位置(p_{best})分量; P_{gd} 为群体所经历的历史最佳位置(g_{best})分量; ω 为惯性权重; c_1 和 c_2 为加速系数; r_1 和 r_2 为 $[0,1]$ 之间均匀分布的随机数,称为随机因子.算法的速度更新方程可分成 3 项:第 1 项为动量部分,表示粒子以先前速度所进行的惯性运动;第 2 项为“认知”部分,表示粒子本身的思考;第 3 项为“社会”部分,表示粒子间的信息共享与相互合作.

PSO 算法根据个体粒子和群体粒子的搜索经验向着最优解的方向运动,合理地利用粒子的个体经验信息 p_{best} 和群体共享信息 g_{best} 为精确有效地寻找最优解提供一条重要途径^[3].粒子对这两个信息的利用依赖于下面两个因素:加速系数和随机因子.加速系数 c_1 和 c_2 代表将每个粒子推向其 p_{best} 和 g_{best} 的统计加速项的权重,加速系数 c_1 和 c_2 通常取值为 2^[1],Clerc 建议, $(c_1+c_2)/2$ 的取值为 1.494^[4].文献[5]表明,认知参数 c_1 选择大些,社会参数 c_2 选择小些,但当 $c_1+c_2 \leq 4$ 时能够得到更好的结果.Ratnaweera^[3]提出加速系数随迭代次数线性变化的策略,文献[6]提出加速系数为 g_{best} 与 p_{best} 之比的非线性函数.文献[7]通过度量多样性来调整加速系数.上述策略从一定程度上改善了 PSO 算法的全局优化能力,但是只针对于加速系数的调整,而忽略了随机因子对算法性能的影响.目前对随机因子的研究甚少,随机因子 r_1, r_2 可以保证微粒群体的多样性和搜索的随机性^[1].但在经典 PSO 算法中, r_1, r_2 的独立性假设使得群体在各自认知及信息共享的部分完全随机地去适应历史,没有加以区别地利用个体经验和群体共享的信息.为了进一步研究粒子对 p_{best} 和 g_{best} 的利用,对随机因子进行深入的探讨是非常必要的.

本文从认知论角度对随机因子进行分析,给出随机因子 r_1, r_2 的相关性假设,提出基于 Copula 相关性 PSO 模型及具体的模型实现方法.之后,从理论上分析随机因子间相关程度与群体多样性之间的关系,剖析相关性 PSO 模型的内在机理.同时,对随机因子间相关程度与算法收敛性之间的关系进行理论分析,并给出其收敛条件.最后通过仿真实验表明,粒子对 p_{best} 和 g_{best} 不同的认知策略的相关性 PSO 模型的优化性能.

1 相关性 PSO 模型

1.1 随机因子的认知分析

PSO 算法的心理学假设是个体在寻求一致的认知过程中记住它自身信念的同时考虑同事们的信念^[2]. p_{best}, g_{best} 分别是粒子根据自身的经验持有的信念和粒子对同伴共享信息持有的信念.随机因子 r_1 和 r_2 刻画了粒

子对持有信念的倾向强度. r_1 的随机性反映了粒子对持有的自身信念 p_{best} 是否会影响算法收敛速度的不确定性; r_2 的随机性则反映了粒子对共享信念 g_{best} 是否为全局最优的不确定性. r_1, r_2 的随机性综合体现了粒子在认知过程中的非理性行为,反映出粒子作为认知主体的情绪^[8].加速系数视为粒子对这种情绪的心理效应^[8],加速系数的增大或减小是对该情绪的放大或减弱.

粒子对 p_{best} 和 g_{best} 信息的利用实质上就是对这两个信息进行认知和加工.由上面的分析可知:随机因子 r_1, r_2 的独立性假设忽视了粒子对 p_{best} 和 g_{best} 认知的联系;同时,单纯的加速系数调整也不能体现 p_{best} 和 g_{best} 在信息加工系统中的内在联系.如何权衡信息加工系统中对 p_{best} 和 g_{best} 的利用,需要建立粒子对这两个信息持有态度的关联性,即随机因子 r_1, r_2 的相关性描述和分析.

1.2 基于Copula相关性的PSO模型

相关性分析是研究事物的相互关系,测定它们联系的紧密程度,揭示其变化的具体形式和规律性的统计方法.在相关性理论中,线性相关系数是分析随机变量间相关关系的一个重要工具.但是近年来却发现,线性相关系数在描述相关性时存在一定的缺陷^[9],而 Copula 函数^[10]为相关性的全面分析提供了一条崭新的途径.因此,本文采用 Copula 来刻画随机因子的相关性.Copula 由 Sklar 于 1959 年首度引入,在 20 世纪 90 年代后期得到了极大的重视并得以迅速发展,在金融风险分析的相关性建模中得到成功的应用^[11].

1.2.1 Copula 的相关知识

Copula 函数实际上是一种将联合分布与其各自的边缘分布连接在一起的函数,称为连接函数.鉴于本文仅讨论两个随机因子 r_1, r_2 的相关性,下面给出二元 Copula 函数的有关定义及定理.

定义 1^[10]. 一个二元函数 $C: [0,1]^2 \rightarrow [0,1]$, 如果满足以下两个条件:

(1) 对于 $\forall t \in [0,1], C(t,0)=C(0,t)=0$ 且 $C(t,1)=C(1,t)=t$;

(2) 对于 $u_1, u_2, v_1, v_2 \in [0,1]$ 且 $u_1 \leq u_2, v_1 \leq v_2$, 有 $C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$,

则称函数 C 为一个 Copula.

Copula 理论中的奠基性工作就是著名的 Sklar 定理,它是一切基于 Copula 的相关性分析的理论支撑.

定理 1^[10]. 设随机变量 X, Y 的联合分布函数为 $H: R^2 \rightarrow [0,1]$, 其边缘分布函数分别为 F_X, F_Y , 则存在一个 Copula $C: [0,1]^2 \rightarrow [0,1]$, 对任意 $x=(x_1, x_2) \in R^2$, 有

$$H(x_1, x_2) = C(F_X(x_1), F_Y(x_2)) \quad (3)$$

相反地, 如果 $C: [0,1]^2 \rightarrow [0,1]$ 是一个 Copula 函数, F_X, F_Y 分别是随机变量 X, Y 的分布函数, 那么存在一个以 F_X, F_Y 为边缘分布的联合分布函数 H , 满足对任意的 $(u_1, u_2) \in [0,1]^2$, 有

$$C(u_1, u_2) = H(F_X^{-1}(u_1), F_Y^{-1}(u_2)) \quad (4)$$

而且, 如果 F_X, F_Y 是连续的, 则 Copula 函数 C 是唯一的; 否则, C 在 $Ran(F_X) \times Ran(F_Y)$ 上唯一确定. $Ran(F)$ 表示函数 F 的值域.

Sklar 定理指出 Copula 函数的存在性, 同时给出如何利用联合分布函数求解 Copula 函数的方法. Copula 理论中有 3 种基本的 Copula: Fréchet-Hoeffding 下界 $W(u, v) = \max(u+v-1, 0)$, Fréchet-Hoeffding 上界 $M(u, v) = \min(u, v)$ 和乘积 Copula: $\Pi(u, v) = uv$. 如果将这 3 种 Copula 函数均视为两个 $[0,1]$ 均匀分布随机变量 U, V 的联合分布函数, 则这 3 种特殊的 Copula 函数依次对应着 3 种特殊的相关关系: 完全负线性相关 ($U=1-V$), 完全正线性相关 ($U=V$) 和独立. Copula 函数与其 Fréchet-Hoeffding 上、下界之间有如下 Fréchet-Hoeffding 界值定理:

定理 2^[10]. 设 C 为任意给定的一个 Copula 函数, 则对任意 $u, v \in [0,1]^2$, 有

$$W(u, v) \leq C(u, v) \leq M(u, v) \quad (5)$$

Copula 函数有很多类型, 不同类型的 Copula 有不同的特征, 可以描述不同的相关结构. 在常见的 Copula 中, Gaussian Copula 用于线性相关的结构, Gumbel Copula 用于极值分布, 而 Archimedean Copula 和 t -Copula 用于尾部相关性研究^[10].

Gaussian Copula 是椭圆 Copula 簇中的一种, 也是一种最常见且应用最为广泛的 Copula. 利用 Sklar 定理, 可以得出二元 Gaussian Copula 的定义.

定义 2^[10]. 对于任意 $(u,v) \in [0,1]^2$, 二元 Gaussian Copula 可定义为

$$C_\rho(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)) \quad (6)$$

式中的 Φ 是标准正态变量 X 和 Y 的分布函数; Φ_ρ 是二维正态变量 X, Y 的联合分布函数; Φ^{-1} 是 Φ 的逆函数; ρ 是随机变量 X 和 Y 的线性相关系数, 且 $-1 < \rho < 1$. ρ 的不同取值反映了随机变量 X 和 Y 之间不同的相关特性. 当 $\rho = -1$ 时, 令 $C_\rho(u,v) = W(u,v)$; 当 $\rho = 0$ 时, 令 $C_\rho(u,v) = \Pi(u,v)$; 当 $\rho = 1$ 时, 令 $C_\rho(u,v) = M(u,v)$.

1.2.2 模型描述

在经典的 PSO 算法中, 随机因子 r_1, r_2 的概率假设为两个独立的 $[0,1]$ 均匀分布随机变量, 模型描述简单. 刻画随机因子 r_1, r_2 的相关性则是建立相关性 PSO 模型的关键. 由 Copula 函数的定义可知, 二元 Copula 函数的实质是两个 $[0,1]$ 均匀分布随机变量的联合分布函数. 因此, 本文将采用 Copula 函数描述随机因子 r_1, r_2 的相关结构, 具体如下:

$$H(r_1, r_2) = C(r_1, r_2) \quad (7)$$

式中的 H 为随机因子 r_1, r_2 的联合分布函数, C 为相应的 Copula 函数. 在给出随机因子 r_1, r_2 的相关性假设公式(7)后, 就可以得出相关性 PSO 模型的定义.

定义 3. 由公式(1)、公式(2)和公式(7)共同描述粒子运动规律的算法模型称为相关性 PSO 模型.

在相关性 PSO 模型中, 随机因子 r_1, r_2 的不同相关性结构(Copula 函数)和同一相关性结构下不同的相关性程度都体现了粒子对 p_{best} 和 g_{best} 不同的认知策略. 随机因子的相关性假设为探求粒子对 p_{best} 和 g_{best} 的利用平衡问题提供了一条新的途径.

Gaussian Copula 是关于线性相关系数 ρ 的分布函数簇, 可以全面反映随机因子 r_1, r_2 的线性相关特性. 同时, 具有 Gaussian Copula 的相关随机数的生成算法较为简单. 本文将采用 Gaussian Copula 描述随机因子 r_1, r_2 的相关结构, 对应的线性相关系数 ρ 刻画了随机因子间的相关程度以及不同的相关系数对应粒子对 p_{best} 和 g_{best} 的不同的认知策略.

1.2.3 模型实现

在经典 PSO 算法中, 随机因子 r_1, r_2 为独立的 $[0,1]$ 均匀分布随机变量, 算法实现方法简单. 在相关性 PSO 模型中, 随机因子 r_1, r_2 为相关随机变量, 模型实现的难点是如何生成两个相关的具有给定 Copula 函数的随机因子. 相关随机变量的生成方法有很多, 其中最常用的是直接利用 Copula 函数 $C(u,v)$ 的方法. 直接利用 Copula 函数的生成方法具有很大的局限性, 它需要计算 C 的偏导 C_u (或 C_v) 及偏导的拟逆 $C_u^{(-1)}$ (或 $C_v^{(-1)}$)^[10]. 由于 Copula 函数表达的复杂性, 这些计算往往仅能通过数值方法来实现. 采用 Gaussian Copula 描述的相关性 PSO 模型无需直接借助 Gaussian Copula 的函数表达, 通过线性变换就可以产生具有 Gaussian Copula 的相关的 $[0,1]$ 均匀分布随机数.

下面给出基于 Gaussian Copula 相关性 PSO 模型实现的具体流程.

第 1 步. 初始化算法参数;

第 2 步. 产生初始群体, 其中所有子粒的位置和速度都是随机的, 并给 g_{best} 粒子赋予初值;

第 3 步. 群体的进化操作:

- 1) 给定相关系数 ρ , 生成相关的 $[0,1]$ 均匀分布随机变量 r_1, r_2 ;
 - (a) 产生两个独立的 $[0,1]$ 均匀分布的随机变量 u_1, u_2 ;
 - (b) 计算 $x = \Phi^{-1}(u_1), y = \Phi^{-1}(u_2)$. 其中, $\Phi(\cdot)$ 是标准正态分布的分布函数;
 - (c) 由 Cholesky 分解获得两个相关的服从正态分布的随机数 $x, y: x = \rho x + (1 - \rho^2)^{0.5} y$;
 - (d) 计算 $r_1 = \Phi(x), r_2 = \Phi(y)$. r_1, r_2 就是具有 Gaussian Copula $C_\rho(u,v)$ 的两个随机数;
- 2) 用公式(1)、公式(2)分别更新粒子的速度和位置;
- 3) 计算粒子的目标函数;
- 4) 更新粒子的 g_{best} 和 p_{best} ;

第 4 步. 算法收敛性判断;

第 5 步. 若未达到结束条件(通常为预设的运算精度或迭代次数), 则返回步骤 3, 开始新一轮迭代计算; 否则,

取当前 g_{best} 作为最优解.

2 相关性 PSO 模型的粒子行为分析

2.1 粒子认知策略分析

Gaussian Copula 是一类特殊的椭圆 Copula.由 Gaussian Copula 的定义 2 可知, $C_\rho(r_1,r_2)=C_\rho(r_2,r_1)$,即对于不同的相关系数 ρ 而言,随机因子 r_1,r_2 在矩形区域 $[0,1] \times [0,1]$ 中的取值是关于对角线 $r_1=r_2$ 对称的,且随着相关系数 $\rho \in [-1,1]$ 的增大,随机因子 r_1,r_2 的取值在对角线 $r_1=r_2$ 附近的聚集程度有所增加.特别地,当 $\rho=1$ 时,随机因子 r_1,r_2 的取值完全落在对角线 $r_1=r_2$ 上.图 1 给出了具有 Gaussian Copula 且相关系数分别为 $-1,-0.4,-0.8,0.4,0.8$ 和 1 的随机因子 r_1,r_2 的取值分布散点图.该图反映了随机因子 r_1,r_2 的分布关于对角线 $r_1=r_2$ 的对称性以及不同相关系数所对应的 r_1,r_2 的取值在对角线 $r_1=r_2$ 附近的聚集程度.

对于基于 Gaussian Copula 的相关性 PSO 模型,随机因子 r_1,r_2 的取值关于对角线 $r_1=r_2$ 的对称性反映了粒子在整个寻优过程中对极值信息 p_{best} 和 g_{best} 利用的均衡性.随着相关系数 ρ 的增加,随机因子 r_1,r_2 的取值在对角线 $r_1=r_2$ 附近的聚集程度有所提高,这进一步反映了粒子在局部寻优过程中对极值信息 p_{best} 和 g_{best} 的均衡利用程度随相关系数 ρ 的增加而增大.

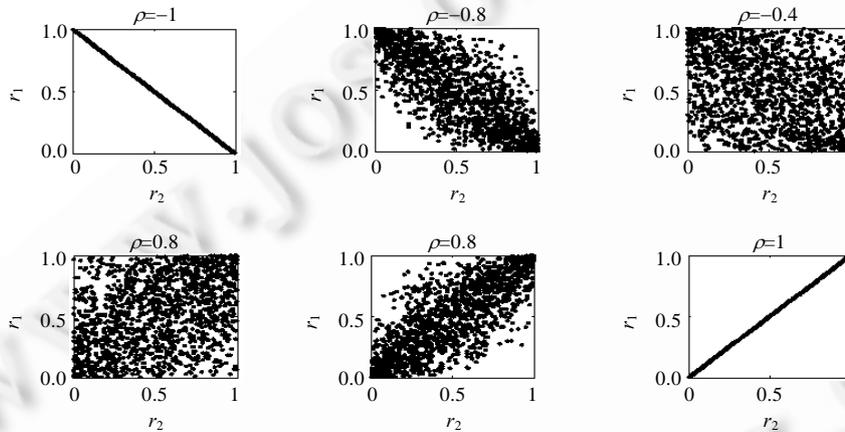


Fig.1 Bivariate scatter diagram of Gaussian Copula with different correlation coefficients

图 1 不同相关系数的 Gaussian Copula 二维随机变量的散点图

2.2 群体多样性分析

群体多样性是用来表征群体内个体特征的差异性,是刻画群体进化过程中的一个重要特征.其本质是粒子的位置相对于某个几何中心的离差程度,几何中心的选取通常是重心. Shi 给出了一种广义的群体多样性计算公式^[12]:

$$div(X(t)) = \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D [X_{id}(t) - \overline{X_d(t)}]^2 \tag{8}$$

其中, $\overline{X_d(t)} = \sum_{i=1}^N X_{id}(t) / N$ 是所有粒子位置 $X(t)$ 重心的 d 维分量, N 是种群规模, D 是待求解问题的维数.

为了研究相关系数 ρ 对粒子群体多样性的影响,本文在当前状态下考察粒子群体下一时刻的群体多样性的变化. $t+1$ 时刻,群体多样性的表达式为

$$div(X(t+1)) = \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D [X_{id}(t+1) - \overline{X_d(t+1)}]^2 \tag{9}$$

其中, $\overline{X_d(t+1)} = \sum_{i=1}^N X_{id}(t+1) / N$. 此时,各个粒子在 t 时刻以及 t 时刻以前的位置和速度等信息都是已知的,粒

子在 $t+1$ 时刻的位置由公式(1)、公式(2)、公式(7)共同确定.显然,由于随机因子 r_1, r_2 的随机性,公式(9)中 $X_{id}(t+1), \overline{X_d}(t+1)$ 均为 r_1, r_2 的随机函数.因而,群体在 $t+1$ 时刻的种群多样性 $div(X(t+1))$ 也是随机的.

鉴于 $div(X(t+1))$ 的随机性,本文采用期望效用最大的策略来确定相关系数 ρ 的选取.对公式(9)两端求数学期望可得:

$$\begin{aligned}
 E[div(X(t+1))] &= \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D E[X_{id}(t+1) - \overline{X_d}(t+1)]^2 \\
 &= \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D E[X_{id}(t+1) - E(X_{id}(t+1)) + E(X_{id}(t+1)) - \overline{X_d}(t+1)]^2 \\
 &= \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D \{E[X_{id}(t+1) - EX_{id}(t+1)]^2 - 2E[(X_{id}(t+1) - EX_{id}(t+1))[\overline{X_d}(t+1) - EX_{id}(t+1)]] + \\
 &\quad E[\overline{X_d}(t+1) - EX_{id}(t+1)]^2\} \\
 &= \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D \left\{ Var[X_{id}(t+1)] + \frac{1}{N} Var[X_{id}(t+1)] + \frac{1}{N^2} \sum_{j=1}^N Var[X_{jd}(t+1)] + \right. \\
 &\quad \left. \left(\frac{1}{N} \sum_{j=1}^N E[X_{jd}(t+1)] - E[X_{id}(t+1)] \right)^2 \right\} \\
 &= \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D \left\{ \left(1 - \frac{1}{N} \right) Var[X_{id}(t+1)] \right\} + \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D (E[X_{id}(t+1)] - \frac{1}{N} \sum_{j=1}^N E[X_{jd}(t+1)])^2 \\
 &= \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D \left\{ \left(1 - \frac{1}{N} \right) Var[X_{id}(t+1)] \right\} + div(E[X(t+1)])
 \end{aligned} \tag{10}$$

又 $E(X_{id}(t+1)) = X_{id}(t) + V_{id}(t) + 0.5c_1(P_{id}(t) - X_{id}(t)) + 0.5c_2(P_{gd}(t) - X_{id}(t))$, 则有

$$Var[X_{id}(t+1)] = \frac{1}{12} \{c_1^2 [P_{id}(t) - X_{id}(t)]^2 + 2\rho c_1 c_2 [P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)] + c_2^2 [P_{gd}(t) - X_{id}(t)]^2\} \tag{11}$$

综合公式(10)、公式(11)可得:

$$E[div(X(t+1))] = E_0[div(X(t+1))] + \frac{\rho c_1 c_2}{6ND} \sum_{i=1}^N \sum_{d=1}^D \{[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)]\} \tag{12}$$

其中, $E_0[div(X(t+1))] = \frac{N-1}{12N^2D} \sum_{i=1}^N \sum_{d=1}^D \{c_1^2 [P_{id}(t) - X_{id}(t)]^2 + c_2^2 [P_{gd}(t) - X_{id}(t)]^2\} + div(E[X(t+1)])$ 是相关系数 $\rho=0$ 时群体在下一时刻 $t+1$ 的多样性的期望,即经典 PSO 的预期的群体多样性.

由公式(12)可知,为了分析相关系数 ρ 对群体多样性的影响,仅需考虑公式(12)第 2 项取值的符号问题.第 2 项的和式 $\sum_{i=1}^N \sum_{d=1}^D [P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)]$ 是 ND 个乘积项 $[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)]$ 的和.根据 $X_{id}(t)$ 与 $P_{id}(t), P_{gd}(t)$ 的位置关系,当 $X_{id}(t)$ 位于 $P_{id}(t)$ 与 $P_{gd}(t)$ 的同侧时,乘积项 $[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)] > 0$; 当 $X_{id}(t)$ 位于 $P_{id}(t)$ 与 $P_{gd}(t)$ 之间时,乘积项 $[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)] < 0$.又对于所有的粒子而言,至少有一个粒子(不妨假定为第 i 个粒子的 P_i 与 P_g 是相等的,此时,对应的所有乘积项 $[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)] > 0 (d=1, \dots, N)$.因此,

$$P([P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)] > 0) > 2/3.$$

不失一般性,假定 $[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)] = \theta + \eta$, 其中, θ 为正常数, $\eta \sim N(0, \delta)$ 为白噪声,则有

$$\mu = E[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)] > 0.$$

依据概率论中的大数定理,对任意 $\varepsilon > 0$ 有

$$\lim_{ND \rightarrow \infty} P \left(\left| \frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D \{[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)]\} - \mu \right| < \varepsilon \right) = 1,$$

即

$$\lim_{ND \rightarrow \infty} P \left(\frac{1}{ND} \sum_{i=1}^N \sum_{d=1}^D \{[P_{id}(t) - X_{id}(t)][P_{gd}(t) - X_{id}(t)]\} > 0 \right) = 1 \tag{13}$$

综合公式(12)、公式(13)可得如下定理:

定理 3. 对于相关性 PSO 模型,群体预期多样性的期望 $E[\text{div}(X(t+1))]$ 随着相关系数 $\rho(0 \leq \rho \leq 1)$ 的增加而增加.

由定理 3 可知,为使群体在下一时刻 $t+1$ 时保持较高的多样性,相关性 PSO 模型应选择尽可能大的相关系数,极端情形时, $\rho=1$.结合第 2.1 节有关基于 Gaussian Copula 的相关性 PSO 模型的认知策略分析,可以得出如下推论:

推论. 对于基于 Gaussian Copula 的相关性 PSO 模型,群体预期的多样性的期望 $E[\text{div}(X(t+1))]$ 随着粒子对极值信息 p_{best} 和 g_{best} 的均衡利用程度的提高而有所增加.

3 相关性 PSO 模型的收敛性分析

PSO 的收敛性分析方法已有了较为深入的研究,文献[13,14]基于线性离散系统理论进行 PSO 收敛性分析,并给出了收敛条件.但在稳定性分析过程中,由于上述方法回避了 PSO 系统是一个带有随机因素的线性时变系统这一问题,文献[15,16]利用随机分析的方法讨论了 PSO 的收敛性,进而给出了更为合理的在考虑随机因素情况下的收敛条件.本文将在此基础上,利用随机过程的分析手段对相关性 PSO 模型进行收敛性分析.

3.1 粒子位置的期望收敛分析

相关性 PSO 模型是一个带有“认知”部分和“社会”部分两个随机项且考虑了两个随机因子 r_1, r_2 之间相关性的线性时变系统.为了便于分析和表达,将问题空间简化为一维,且仅研究 PSO 中的某一个粒子 i 的运动过程.在粒子寻优后期(收敛期),假设 P_i 和 P_g 在粒子 i 的运动过程中不发生变化,由公式(1)、公式(2)和公式(7)可得粒子 i 运动的状态方程为

$$V(t+1) = \omega V(t) + c_1 r_1(t)(P_i - X(t)) + c_2 r_2(t)(P_g - X(t)) \quad (14)$$

$$X(t+1) = X(t) + V(t+1) \quad (15)$$

式中的随机因子 r_1 和 r_2 是 $[0,1]$ 均匀分布且相关系数为 ρ 的相关随机变量(在收敛性分析过程中,由于不涉及随机因子间的相关结构,因此,以下有关的收敛性分析结果对随机因子不同的相关结构均是适合的).

将公式(14)的时间项向前推一步,并将公式(15)代入公式(14)可得:

$$X(t+1) = (1 + \omega - (c_1 r_1(t) + c_2 r_2(t)))X(t) - \omega X(t-1) + c_1 r_1(t)P_i + c_2 r_2(t)P_g \quad (16)$$

在公式(16)中, $X(t)$ 是随机变量,那么迭代过程 $\{X(t)\}$ 可看成是一个随机过程.对公式(16)两端求数学期望,可得关于 $X(t)$ 的一个二阶差分方程如下:

$$E(X(t+1)) = (1 + \omega - (c_1 + c_2)/2)E(X(t)) - \omega E(X(t-1)) + (c_1 P_i + c_2 P_g)/2 \quad (17)$$

特征方程为

$$\lambda^2 - (1 + \omega - (c_1 + c_2)/2)\lambda + \omega = 0 \quad (18)$$

$\{E(X(t))\}$ 的收敛性分析将以二阶差分方程(18)为讨论基础,相关结论如下:

定理 4. 如果 $\omega, c_1, c_2 \geq 0, 0 \leq \omega < 1$ 且 $0 < c_1 + c_2 < 4(1 + \omega)$, 序列 $\{E(X(t))\}$ 收敛于 $(c_1 P_i + c_2 P_g)/(c_1 + c_2)$.

证明:根据公式(18)可以得到特征方程的特征根: $\lambda = 0.5[(1 + \omega - (c_1 + c_2)/2) \pm ((1 + \omega - (c_1 + c_2)/2)^2 - 4\omega)^{0.5}]$, 如果需要序列 $\{E(X(t))\}$ 收敛,那么特征根的最大值将小于 1,从下面两个方面来考虑:

1. 当 $(1 + \omega - (c_1 + c_2)/2)^2 < 4\omega$ 时,有两个复根,令其为 λ_1, λ_2 , 则有 $|\lambda_1|^2 = |\lambda_2|^2 = \omega$. 由于 $0 \leq \omega < 1$, 所以 $\max\{|\lambda_1|, |\lambda_2|\} < 1$, 因此可得 $2(1 + \omega - 2\omega^{0.5}) < c_1 + c_2 < 2(1 + \omega + 2\omega^{0.5})$.

2. 当 $(1 + \omega - (c_1 + c_2)/2)^2 \geq 4\omega$, 即 $c_1 + c_2 \geq 2(1 + \omega + 2\omega^{0.5})$ 或 $c_1 + c_2 \leq 2(1 + \omega - 2\omega^{0.5})$ 时,有两个实根 λ_1, λ_2 , 当 $c_1 + c_2 \leq 2(1 + \omega - 2\omega^{0.5})$ 时,需要 $\max\{|\lambda_1|, |\lambda_2|\} < 1$, 则有 $0.5|(1 + \omega - (c_1 + c_2)/2) \pm ((1 + \omega - (c_1 + c_2)/2)^2 - 4\omega)^{0.5}| < 1$, 得到 $0 < c_1 + c_2 \leq 2(1 + \omega - 2\omega^{0.5})$. 当 $c_1 + c_2 \geq 2(1 + \omega + 2\omega^{0.5})$ 时,需要 $\max\{|\lambda_1|, |\lambda_2|\} < 1$, 则有 $0.5|(1 + \omega - (c_1 + c_2)/2) \pm ((1 + \omega - (c_1 + c_2)/2)^2 - 4\omega)^{0.5}| > -1$, 得到 $2(1 + \omega + 2\omega^{0.5}) < c_1 + c_2 < 4(1 + \omega)$.

综合上述两个方面,序列 $\{E(X(t))\}$ 收敛的条件是 $0 \leq \omega < 1$ 和 $0 < c_1 + c_2 < 4(1 + \omega)$. 当序列 $\{E(X(t))\}$ 收敛时,可以得到 $EX = (1 + \omega - (c_1 + c_2)/2)EX - \omega EX + (c_1 P_i + c_2 P_g)/2$, 整理可得 $EX = (c_1 P_i + c_2 P_g)/(c_1 + c_2)$. \square

3.2 粒子位置方差的收敛分析

为了进一步讨论粒子位置期望收敛的稳定性,下面需要考察粒子位置变量方差的收敛性.

为了简化表达,令

- $v=(c_1+c_2)/2$;
- $\mu=(c_1P_i+c_2P_g)/(c_1+c_2)$;
- $\psi=1+\omega-v$;
- $Z(t)=c_1r_1(t)+c_2r_2(t)-v$;
- $U(t)=c_1c_2(r_2(t)-r_1(t))(P_g-P_i)/(c_1+c_2)$;
- $Y(t)=X(t)-\mu$.

由公式(16)可以得到:

$$Y(t+1)=(\psi-Z(t))Y(t)-\omega Y(t-1)+U(t) \quad (19)$$

显然, $Y(t)$ 也是随机变量,且 $E(Y(t))=E(X(t))-\mu, D(Y(t))=D(X(t))$.

由于随机因子 $r_1(t)$ 和 $r_2(t)$ 是相关的 $[0,1]$ 均匀分布随机变量,且相关系数为 ρ ,于是可得:

- $E(Z(t))=E(U(t))=0$;
- $D(Z(t))=D(c_1r_1(t)+c_2r_2(t))=c_1^2D(r_1(t))+c_2^2D(r_2(t))+2c_1c_2cov(r_1(t),r_2(t))=(c_1^2+c_2^2+2c_1c_2\rho_{r_1,r_2})/12$;
- $D(U(t))=(c_1c_2(P_g-P_i)/(c_1+c_2))^2(1-\rho)/6$;
- $E(Z(t)U(t))=c_1c_2(P_g-P_i)(c_2-c_1)(1-\rho)/(12(c_1+c_2))$.

注意,这里的 ρ 对于序列 $\{X(t)\}$ 是固定的.记 $Z=D(Z(t)), U=D(U(t))$ 和 $T=E(Z(t)U(t))$.由于 $Y(t)$ 和 $Y(t-1)$ 相关,而 $Y(t), Y(t-1)$ 和 $U(t), Z(t)$ 之间是相互独立的,所以, $E(Y(t+1)^2), E(Y(t+2)^2)$ 和 $E(Y(t+1)Y(t))$ 有如下计算结果:

$$E(Y(t+1)^2)=(\psi^2+Z)E(Y(t)^2)+\omega^2E(Y(t-1)^2)+U-2\omega\psi E(Y(t)Y(t-1))-2TE(Y(t)) \quad (20)$$

$$E(Y(t+2)^2)=(\psi^2+Z)E(Y(t+1)^2)+\omega^2E(Y(t)^2)+U-2\omega\psi E(Y(t+1)Y(t))-2TEY(t+1) \quad (21)$$

$$E(Y(t+1)Y(t))=\psi E(Y(t)^2)-\omega E(Y(t)Y(t-1)) \quad (22)$$

由公式(22)消去公式(20)和公式(21)的混合项 $E(Y(t)Y(t-1))$ 和 $E(Y(t+1)Y(t))$ 可得:

$$E(Y(t+2)^2)=(\psi^2+Z-\omega)E(Y(t+1)^2)-\omega(\psi^2-Z-\omega)E(Y(t)^2)+\omega^3E(Y(t-1)^2)+U(1+\omega)-2T(E(Y(t+1))+\omega E(Y(t))) \quad (23)$$

将 $D(Y(t))=E(Y(t)^2)-(E(Y(t)))^2, E(Y(t+2))=\psi E(Y(t+1))-\omega E(Y(t))$ 和 $\omega E(Y(t-1))=\psi E(Y(t))-E(Y(t+1))$ 代入公式(23)可得:

$$D(Y(t+2))=(\psi^2+Z-\omega)D(Y(t+1))-\omega(\psi^2-Z-\omega)D(Y(t))+\omega^3D(Y(t-1))+Z[(E(Y(t+1)))^2+\omega(E(Y(t)))^2]-2T(E(Y(t+1))+\omega E(Y(t)))+U(1+\omega) \quad (24)$$

再由公式(24), $D(Y(t))=D(X(t))$ 和 $E(Y(t))=E(X(t))-\mu$ 可得 $D(X(t))$ 的迭代方程:

$$D(X(t+2))=(\psi^2+Z-\omega)D(X(t+1))-\omega(\psi^2-Z-\omega)D(X(t))+\omega^3D(X(t-1))+Z[(E(X(t+1))-\mu)^2+\omega(E(X(t))-\mu)^2]-2T(E(X(t+1))-\mu+\omega(E(X(t))-\mu))+U(1+\omega) \quad (25)$$

由公式(25)可知, $\{D(X(t))\}$ 是一个三阶差分定常系统,其特征方程为

$$\lambda^3-(\psi^2+Z-\omega)\lambda^2+\omega(\psi^2-Z-\omega)\lambda-\omega^3=0 \quad (26)$$

为了分析上述系统的稳定性条件,借助文献[16]辅助方程的手段.令 $\lambda_1, \lambda_2, \lambda_3$ 是特征方程的3个根:

$$h(\lambda)=\lambda^3-(\psi^2+Z-\omega)\lambda^2+\omega(\psi^2-Z-\omega)\lambda-\omega^3,$$

$$h(1)=1-(\psi^2+Z-\omega)+\omega(\psi^2-Z-\omega)-\omega^3.$$

下面给出辅助定理.

引理. 当 $0 \leq \omega < 1$ 和 $c_1+c_2 > 0$ 时, $\max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\} < 1$ 的充分必要条件是 $h(1) > 0$.

证明:首先考虑两种特殊的情况:

1. 当 $\omega=0$ 时,特征方程式(26)可化为 $\lambda^2(\lambda-\psi^2-Z)=0$ 的形式,其特征根分别为 $\lambda_1=\psi^2+Z \geq 0, \lambda_2=\lambda_3=0$.

此时, $\max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\} < 1$ 的充要条件是 $|\lambda_1| < 1$ 或 $h(1)=1-|\lambda_1| > 0$.

2. 当 $\psi=0$, 即 $c_1+c_2=(1+\omega)/2$ 时, 特征方程(26)可化为 $(\lambda+\omega)(\lambda^2-Z\lambda-\omega^2)=0$ 的形式, 其特征根分别为 $\lambda_1=-\omega, \lambda_2=0.5(Z-(Z^2+4\omega^2))^{0.5}, \lambda_3=0.5(Z+(Z^2+4\omega^2))^{0.5}$. 显然, $|\lambda_1|=\omega<1$. 由于 $Z>0, \omega^2\geq 0, \max\{|\lambda_2|, |\lambda_3|\}=0.5(Z+(Z^2+4\omega^2))^{0.5}$. 因此, $\max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}<1$ 的充要条件是 $0.5(Z+(Z^2+4\omega^2))^{0.5}<1$. 进一步简化 $0.5(Z+(Z^2+4\omega^2))^{0.5}<1$, 得到 $1-Z-\omega^2>0$. 此时, $h(1)=(1-Z-\omega^2)(1+\omega)>0$. 因此, 在假设 $0\leq\omega<1$ 和 $c_1+c_2>0$ 的前提下, $\max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}<1$ 的充要条件是 $h(1)>0$.

考虑更为一般的情况, 当 $0<\omega<1, c_1+c_2>0$ 且 $c_1+c_2\neq 2(1+\omega)$ 时, 由方程式(26)可知, $\lambda_1\lambda_2\lambda_3=\omega^3>0$, 这表明, 该方程有 3 个正根或者两个负根和一个正根. 下面从必要性和充分性两个方面予以证明:

必要性的证明. 采用反证法. 假设 $h(1)=0$, 那么 1 是方程的根, $\max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}<1$ 不可能成立. 假设 $h(1)<0$, 注意到 $h(0)=-\omega^3<0, h(\omega)=-2\omega^2Z<0, h(-\omega)=-2\psi^2\omega^2<0$, 那么, $h(1)$ 与 $h(-\omega), h(\omega), h(0)$ 同号, 方程 $h(\lambda)=0$ 的根的数目在区间 $(-\omega, 0), (0, \omega), (\omega, 1)$ 是偶数, 那么另一个根将落在区间 $(1, \infty)$ 内, 使 $\max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}<1$ 不可能成立. 因此, $h(1)>0$.

充分性的证明. 由于 $h(1)>0, h(\omega)<0$, 因此方程的根的数目在区间 $(\omega, 1)$ 是奇数, 但不会是 3 个, 因为如果这 3 个根都在 $(\omega, 1)$, 那么使 $\lambda_1\lambda_2\lambda_3>\omega^3$, 所以在区间 $(\omega, 1)$ 只有一个根. 又因为 $h(-\omega), h(\omega), h(0)$ 同号, 方程的根的数目在区间 $(-\omega, 0), (0, \omega)$ 是偶数. 那么另两个根要么落在区间 $(-\omega, 0)$ 或 $(0, \omega)$ 内, 要么落在区间 $(-\infty, -\omega)$ 或 $(1, \infty)$ 内. 由于 $\lambda_1\lambda_2\lambda_3=\omega^3$, 方程的根不会落在 $(-\infty, -\omega)$ 和 $(1, \infty)$. 综合前面对方程的根的位置的讨论, 必有 $\max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}<1$. \square

定理 5. 如果 $\omega, c_1, c_2\geq 0$, 当 $0\leq\omega<1, c_1+c_2>0$ 和 $h(1)>0$ 被满足时, 序列 $\{Var(X(t))\}$ 收敛于

$$Var(X)=((c_1c_2(P_g-P_i)/(6(c_1+c_2)))^2(1+\omega)(1-\rho))/h(1),$$

其中, $h(1)=-(c_1+c_2)\omega^2+\omega[(c_1+c_2)^2+c_1c_2(1-\rho)]/6+c_1+c_2-(c_1+c_2)^2/3+c_1c_2(1-\rho)/6$.

证明: 由引理可知, 当 $\omega, c_1, c_2\geq 0, 0\leq\omega<1, c_1+c_2>0$ 和 $h(1)>0$ 时, 方程(26)的特征根满足不等式 $\max\{|\lambda_1|, |\lambda_2|, |\lambda_3|\}<1$. 此时, 方差序列 $\{Var(X(t))\}$ 收敛. 对公式(25)两端, 当 $t\rightarrow\infty$ 取极限时, 则有

$$\begin{aligned} \lim_{t\rightarrow\infty} Var(X(t+2)) &= (\psi^2 + Z - \omega) \lim_{t\rightarrow\infty} Var(X(t+1)) - \omega(\psi^2 + Z - \omega) \lim_{t\rightarrow\infty} Var(X(t)) + \\ &\quad \omega^3 \lim_{t\rightarrow\infty} Var(X(t-1)) + Z[(\lim_{t\rightarrow\infty} E(X(t+1)) - \mu)^2 + \omega(\lim_{t\rightarrow\infty} E(X(t)) - \mu)^2] - \\ &\quad 2T(\lim_{t\rightarrow\infty} E(X(t+1)) - \mu) + \omega(\lim_{t\rightarrow\infty} E(X(t)) - \mu) + U(1 + \omega) \end{aligned} \quad (27)$$

将 $\lim_{t\rightarrow\infty} EX_t = \mu$ 带入公式(27)可得:

$$Var(X) = \lim_{t\rightarrow\infty} Var(X(t)) = (c_1c_2(P_g - P_i)/6(c_1+c_2))^2(1+\omega)(1-\rho)/h(1) \quad (28)$$

其中, $h(1)=-(c_1+c_2)\omega^2+\omega[(c_1+c_2)^2+c_1c_2(1-\rho)]/6+c_1+c_2-(c_1+c_2)^2/3+c_1c_2(1-\rho)/6$. \square

当 $\rho=0$ 时, PSO 相关性模型退化为经典 PSO 模型. 此时, $\lim_{t\rightarrow\infty} Var(X) = (c_1c_2(P_g - P_i)/6(c_1+c_2))^2(1+\omega)/h(1)$. 当 $\rho=1$ 时, $\lim_{t\rightarrow\infty} Var(X) = 0$, 粒子位置变量 X_t 均方收敛于 $\mu=(c_1P_i+c_2P_g)/(c_1+c_2)$.

由相关性 PSO 模型的收敛性分析, 可得如下结论: (1) 综合定理 4 和定理 5, 对于不同的相关系数, 当 $0\leq\omega<1, 0<c_1+c_2<4(1+\omega)$ 且 $h(1)>0$ 被满足时, 粒子位置期望均收敛于 $(c_1P_i+c_2P_g)/(c_1+c_2)$; (2) 从定理 4 给出的粒子位置方差极限与随机因子相关系数 $\rho_{\eta, \eta}$ 的关系表达式可知, 在粒子收敛的过程中, 粒子位置的波动幅度与随机因子的相关程度有关. 当随机因子间完全正线性相关时, 粒子位置变量均方收敛于其位置变量的均值.

4 仿真实验

4.1 实验设计

为了测试粒子在对 p_{best} 和 g_{best} 不同的认知策略(即不同相关系数水平)下, 基于 Gaussian Copula 相关性 PSO 模型的性能, 本文设计了两个实验, 分别为相关系数的方差分析实验和相关性 PSO 模型优化性能实验. 第 1 个实验用于检验在相同的实验条件下, 不同的相关系数水平对算法性能是否具有显著性的影响; 第 2 个实验用于检验模型的综合优化(全局、局部搜索能力)性能. 实验选用 6 个常用于优化算法比较的单峰和多峰基准函数, 其函数形式、维数、搜索范围和理论极值见表 1. 第 1 个实验选用了两个测试函数, 分别为单峰函数 Sphere 和多峰函数 Schaffer; 第 2 个实验选用了所有的测试函数.

Table 1 Benchmark function used to test the algorithms**表 1** 用于测试算法的基准函数

Name	Fomula	Dim	Range (x_{\min}, x_{\max})	Optimal f
Sphere f_1	$f_1(x) = \sum_{i=1}^D x_i^2$	30	[-100,100]	0
Quadric's f_2	$f_2(x) = \sum_{j=1}^D \left(\sum_{k=1}^j x_k \right)^2$	30	[-30,30]	0
Griewank's f_3	$f_3(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[-600,600]	0
Rastrigin's f_4	$f_4(x) = \sum_{i=1}^D (x_i^2 - 10 \cos(2\pi x_i) + 10)$	30	[-5.12,5.12]	0
Ackley's f_5	$f_5(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2}\right) - \exp\left(\frac{1}{D} \sum_{i=1}^D \cos(2x_i)\right) + 20 + e$	30	[-32.768,-32.768]	0
Shaffer f_6	$f_6(x) = 0.5 - \frac{(\sin \sqrt{x_1^2 + x_2^2})^2 - 0.5}{(1.0 + 0.001(x_1^2 + x_2^2))^2}$	2	[-100,100]	0

具体参数设置为:种群规模为 40,进化代数数为 2 000,速度的最大值 $V_{\max}=X_{\max}$,粒子初始化范围均为基准函数的自变量界限.为了减小加速系数和惯性权重的不同取值策略对算法性能的影响,本文对上述参数均设常值: $c_1=c_2=2$; $\omega=0.7$ (在常用的 ω 取值区间[0.4,1]中取中间值).

相关系数的设置:由于相关系数的取值范围为 $-1 \leq \rho \leq 1$,实验采用了均匀抽样法.即在区间[-1,1]以 0.2 为步长取值,即因素水平数 $l=11$.本文采用同一个因素水平下进行相同的实验次数,因此,对于同一个因素水平重复进行 10 次实验.

4.2 实验结果分析

4.2.1 相关系数方差分析结果

基于 Gaussian Copula 相关性 PSO 模型的相关系数方差分析的结果见表 2 和表 3.表中计算量 $SS(Columns)$ 和 $SS(Error)$ 分别表示单因素条件相关系数下组间平方和和误差平方和; $df(Columns)$ 和 $df(Error)$ 分别表示组间平方和的自由度和误差平方和的自由度; $MS(Columns)$ 和 $MS(Error)$ 分别表示组间平均平方和和误差平均平方和; F 为组间平均平方和与误差平均平方和之比; p 为检验值,当 $0.01 < p \leq 0.05$ 时,认为影响显著;当 $p < 0.01$ 时,认为影响特别显著.由表 2 和表 3 可知,对于 Sphere 函数和 Schaffer 函数方差分析的检验值 p 远远小于 0.01,即随机因子间相关系数水平设置对模型的收敛精度有着非常显著的影响.这揭示了粒子认知策略中考虑随机因子相关性的意义和引入相关性 PSO 模型的必要性.

Table 2 Variance analysis of optimization performance for Sphere function**表 2** Sphere 函数优化效果的单因子方差分析

	SS	df	MS	F	P
Columns	5.062e+8	10	5.062e+7	9.47	7.073e-11
Error	5.293e+8	99	5.347e+6	—	—
Total	1.035e+9	109	—	—	—

Table 3 Variance analysis of optimization performance for Schaffer function**表 3** Schaffer 函数优化效果的单因子方差分析

	SS	df	MS	F	P
Columns	8.392e+1	10	8.392e+0	10.62	5.098e-12
Error	7.827e+1	99	7.906e-1	—	—
Total	1.622e+2	109	—	—	—

4.2.2 基于 Gaussian Copula 相关性 PSO 模型的优化性能分析

基于不同相关系数的 Gaussian Copula 相关性 PSO 模型的优化结果见表 4.表中给出了 6 个基准函数的最小值、均值和标准方差.由表 4 可知,随着相关系数的增加,6 个基准函数的最小值、均值和标准方差有明显降低的趋势.从整体上看,当随机因子间呈正线性相关时,模型的优化性能要优于随机因子间呈负线性相关的模型的优化性能.特别地,当随机因子间呈完全负线性相关(即相关系数为-1)时,基于 Gaussian Copula 相关性 PSO 模型的优化性能最差;而当随机因子间呈完全正线性相关(即相关系数为 1)时,基于 Gaussian Copula 相关性 PSO 模型不仅获得了最佳的优化性能,而且对于函数 f_1, f_2, f_3, f_4 和 f_5 达到了全局最优值.

图 2 给出基于不同相关系数的 Gaussian Copula 相关性 PSO 模型的收敛曲线(注:为了方便进化曲线的显示和观察,本文只显示了相关系数为-1,-0.6,-0.2,0,0.2,0.6 和 1 时函数的收敛曲线;同时,对函数的适应度取以 10 为底的对数).从图 2 可知,当相关系数为 1 时,基于 Gaussian Copula 相关性 PSO 模型不仅具有较高的收敛精度,同时具有较快的收敛速度,与相关系数为 0(即相关系数独立情况)时相比有非常显著的提高.由于随机因子间呈正线性相关时有利于维持群体多样性,使粒子避免陷入局部最优,因此明显提高了模型的综合性能.

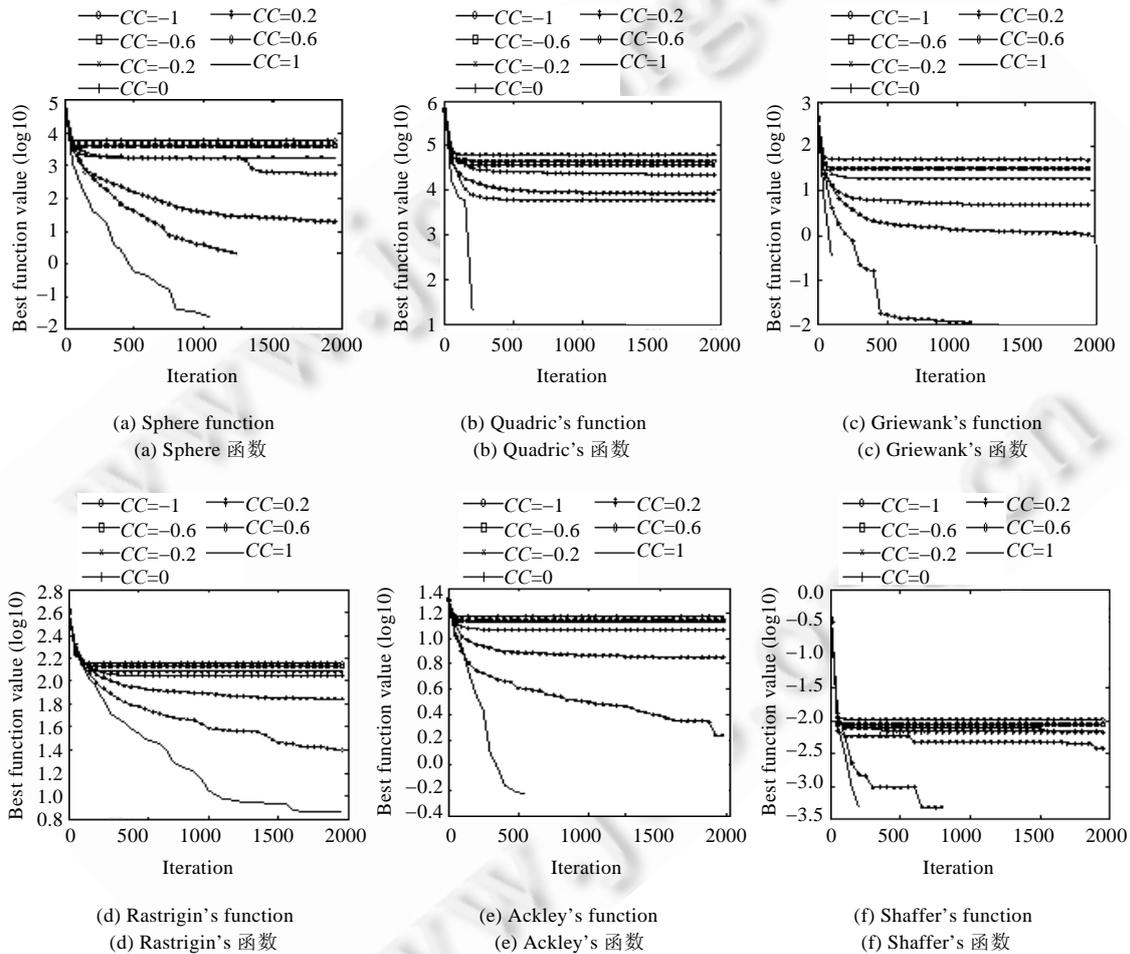


Fig.2 Convergence curve of functions for different correlation coefficients

图 2 不同相关系数下函数的收敛曲线

Table 4 Results of optimization performance for the different correlation coefficient**表 4** 不同相关系数下模型的优化性能

F \ CC	f_1			f_2			f_3		
	Min	Mean	Std	Min	Mean	Std	Min	Mean	Std
-1	1.336e+2	1.684e+3	2.078e+3	3.652e+4	6.188e+4	3.360e+4	1.879e+1	3.555e+1	1.291e+1
-0.8	2.831e+2	1.339e+3	1.172e+3	3.606e+4	7.026e+4	2.708e+4	1.325e+1	3.976e+1	2.109e+1
-0.6	4.006e+1	7.573e+2	5.954e+2	2.498e+4	4.605e+4	1.975e+4	1.976e+1	4.159e+1	2.872e+1
-0.4	1.183e+1	8.335e+2	7.326e+2	2.860e+4	6.919e+4	5.628e+4	1.600e+1	2.617e+1	5.878e+0
-0.2	7.970e+1	7.631e+2	8.014e+2	1.458e+4	3.471e+4	1.658e+4	1.498e+1	2.095e+1	4.267e+0
0	2.618e+1	2.954e+2	3.619e+2	4.525e+3	2.228e+4	3.086e+4	2.209e+0	4.874e+0	2.603e+0
0.2	0	3.527e+1	3.665e+1	5.816e+1	8.503e+3	2.191e+4	4.342e-1	1.100e+0	2.510e-1
0.4	0	1.290e+1	1.423e+0	0	4.005e+3	9.662e+3	0	1.077e-1	1.992e-1
0.6	0	1.131e-1	3.636e-3	0	3.000e+3	1.578e+4	0	0	0
0.8	0	0	0	0	8.342e+2	4.641e+3	0	0	0
1	0	0	0	0	0	0	0	0	0

F \ CC	f_4			f_5			f_6		
	Min	Mean	Std	Min	Mean	Std	Min	Mean	Std
-1	1.107e+2	1.443e+2	2.379e+1	1.315e+1	1.509e+1	1.046e+0	0	1.150e-2	9.287e-3
-0.8	1.150e+2	1.422e+2	1.911e+1	9.985e+0	1.347e+1	1.802e+0	0	8.662e-3	7.969e-3
-0.6	1.234e+2	1.348e+2	1.474e+1	1.096e+1	1.429e+1	1.891e+0	0	8.259e-3	3.559e-3
-0.4	8.497e+1	1.244e+2	2.249e+1	1.094e+1	1.351e+1	1.630e+0	0	9.634e-3	7.398e-3
-0.2	8.751e+1	1.220e+2	1.673e+1	5.110e+0	1.198e+1	2.890e+0	0	9.634e-3	7.398e-3
0	8.463e+1	1.113e+2	1.819e+1	9.836e+0	1.159e+1	1.126e+0	0	8.177e-3	8.194e-3
0.2	3.466e+1	7.025e+1	2.671e+1	0	7.416e+0	3.282e+0	0	3.401e-3	4.755e-3
0.4	7.419e+0	4.539e+1	2.758e+1	0	2.211e+0	1.962e+0	0	2.357e-3	1.145e-3
0.6	7.232e+0	2.526e+1	1.690e+1	0	1.403e+0	4.438e+0	0	0	0
0.8	7.214e+0	1.277e+1	1.432e+1	0	2.131e-1	1.324e-1	0	0	0
1	7.214e+0	7.228e+0	2.336e-2	0	0	0	0	0	0

4.2.3 与参考文献中的优化性能比较

将完全线性正相关的 PSO 模型(CPCPSO)与目前具有代表性的 7 种算法进行性能比较.这 7 种算法分别是时变加速系数 PSO 算法(PSO-TVAC)^[3]、压缩因子 PSO 算法^[4](PSO-CF)、具有开采能力的 PSO 算法^[6](ePSO)、多样性反馈的自组织 PSO 算法^[7](SOPSO)、适应度距离比 PSO 算法^[17](FDR)、全信息 PSO 算法^[18](FIPS)和全面学习 PSO 算法^[19](CLPSO).本文的实验条件(如粒子种群数、进化迭代次数等)比参考文献中要苛刻更多.

对比结果见表 5,每个函数最佳的实验结果用加粗字体表示.对于函数 f_1, f_2, f_3 和 f_6 ,CPCPSO 比其他几种算法获得更为优良的性能,均达到全局最优值.对于函数 f_4 ,CPCPSO 和 CLPSO 都获取全局最优值,但 CLPSO 花费大且优化结果依赖于引入的参数.对于函数 f_5 ,CPCPSO 的优化结果比 CLPSO 和 ePSO 优化结果要差,但优于其他 5 种算法.综上所述,对比结果说明,CPCPSO 在没有引入任何其他操作和参数的情况下明显提高了算法的收敛精度.

Table 5 Comparison of some modified PSOs**表 5** 与改进的 PSO 算法的比较

Results	f_1	f_2	f_3	f_4	f_5	f_6
PSO-CF	3.79e-11±2.67	4.49e-11±2.98e-11	5.20e-2±8.70e-2	5.62e+1±9.76	3.65±1.54	1.19e+1±0
PSO-TVAC	1.00e-2±0	9.40e-5±6.70e-6	1.42e-2±1.85e-2	3.62e+1±8.13e+1	5.22e-1±7.34e-1	2.20e-3±3.83e-3
SOPSO	—	—	2.14e-2±5.30e-3	1.03e+1±2.36	2.93e-2±1.75e-2	—
FDR	4.88e-10±1.53e-10	1.87e-15±0	1.01e-2±1.23e-2	2.84e+1±8.71	2.84e-14±4.10e-15	—
FIPS	2.69e-12±6.84e-13	—	1.16e-6±1.87e-6	7.30e+1±1.24e+1	4.81e-7±9.17e-8	1.00e-5±0
CLPSO	4.46e-14±1.73e-14	—	3.14e-10±4.64e-10	4.85e-10±3.63e-10	0±0	—
ePSO	2.95e-5±1.12e-5	6.40e-5±1.42e-5	8.31e-5±1.43e-5	8.24e-5±1.33e-5	1.38e-4±9.09e-5	4.46e-4±1.93e-3
CPCPSO	0±0	0±0	0±0	7.22±2.12e-2	0±0	0±0

5 结束语

在随机因子认知分析的基础上,本文给出了随机因子的相关性假设,提出了相关性 PSO 模型.该模型采用 Copula 函数描述了粒子对自身经验信息和群体共享信息持有态度的关联性,即随机因子 r_1, r_2 的相关性.给出了

基于 Gauss Copula 的相关性 PSO 模型的实现方法,通过分析 Gauss Copula 的概率特性解释了基于 Gauss Copula 的相关性 PSO 模型的认知行为.进一步的理论结果包括两个方面:在惯性权重 ω 、加速系数 c_1 和 c_2 都保持固定策略的前提下,群体预期的多样性随着相关系数 $\rho(0 \leq \rho \leq 1)$ 的增加而增大;粒子位置的波动幅度随着随机因子相关程度的变化而变化.特别地,当随机因子间呈完全正线性相关时,粒子位置变量均方收敛于其位置变量的均值.仿真实验中相关系数的方差分析表明,随机因子间相关系数的水平设置对模型的性能有着显著性影响.优化实验(不同相关系数的 Gaussian Copula 相关性 PSO 模型)结果显示:随着相关系数 ρ 的增加,模型的优化性能有明显提高的趋势;当随机因子呈完全正线性相关时,模型的综合优化性能达到最高.

相关性 PSO 模型的建立,为研究粒子对自身经验信息和群体共享信息的合理利用问题提供了一条新的途径.后续工作将围绕相关性 PSO 模型给出进一步的理论分析,并探讨相关结构(Copula)的选取对优化性能的影响.

References:

- [1] Kennedy J, Eberhart RC. Particle swarm optimization. In: Proc. of the Int'l Conf. on Neural Networks. Perth: IEEE Press, 1995. 1942–1948. <http://www.engr.iupui.edu/~shi/Coference/psopap4.html>
- [2] Kennedy J, Eberhart RC, Shi Y. Swarm Intelligence. Singapore: Elsevier Science Press, 2001. 202–210.
- [3] Ratnaweera A, Halgamuge SK, Watson HC. Self-Organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. IEEE Trans. on Evolutionary Computation, 2004,8(3):240–255. [doi: 10.1109/TEVC.2004.826071]
- [4] Clerc M, Kennedy J. The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. on Evolutionary Computation, 2002,6(1):58–73. [doi: 10.1109/4235.985692]
- [5] Parsopoulos KE, Vrahatis MN. Recent approaches to global optimization problems through particle swarm optimization. Natural Computing, 2002,1(2-3):235–306. [doi: 10.1023/A:1016568309421]
- [6] Arumugam MS, Rao MVC, Tan AWC. A novel and effective particle swarm optimization like algorithm with extrapolation technique. Applied Soft Computing, 2009,9(1):308–320. [doi: 10.1016/j.asoc.2008.04.016]
- [7] Jie J, Zeng JC, Han CZ. Self-Organized particle swarm optimization based on feedback control of diversity. Journal of Computer Research and Development, 2008,45(3):464–471 (in Chinese with English abstract).
- [8] Wang P. Social Cognitive Psychology. Beijing: China Social Sciences Press, 2006. 72–86 (in Chinese).
- [9] Embrechts P, Mcneil AJ, Straumann D. Correlation and dependency in risk management: Properties and pitfalls. In: Proc. of the Risk Management: Value at Risk and Beyond. Cambridge University Press, 2001. 176–223. <http://www.math.ethz.ch/%7Ebaltes/ftp/papers.html>
- [10] Nelsen RB. An Introduction to Copulas. 2nd ed., New York: Springer-Verlag, 2006. 10–28.
- [11] Mendes BVDM, Souza RMD. Measuring financial risks with copulas. Int'l Review of Financial Analysis, 2004,13(1):27–45. [doi: 10.1016/j.irfa.2004.01.007]
- [12] Shi YH, Eberhart RC. Population diversity of particle swarms. In: Wang J, *et al.*, eds. Proc. of the IEEE World Congress on Computational Intelligence. Hong Kong: IEEE Press, 2008. 1063–1067.
- [13] Trelea IC. The particle swarm optimization algorithm: Convergence analysis and parameter selection. Information Processing Letters, 2003,85(6):317–325. [doi: 10.1016/S0020-0190(02)00447-7]
- [14] Bergh FVD, Engelbrecht AP. A study of particle swarm optimization particle trajectories. Information Sciences, 2006,176(8): 937–971. [doi: 10.1016/j.ins.2005.02.003]
- [15] Jin XL, Ma LH, Wu TJ. Convergence analysis of the particle swarm optimization based on stochastic processes. Acta Automatica Sinica, 2007,33(12):1263–1268 (in Chinese with English abstract).
- [16] Jiang M, Luo YP, Yang SY. Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. Information Processing Letters, 2007,102(1):8–16. [doi: 10.1016/j.ipl.2006.10.005]
- [17] Peram T, Veeramachaneni K, Mohan CK. Fitness-Distance-Ratio based particle swarm optimization. In: Proc. of the Swarm Intelligence Symp. 2003. 174–181. http://www.ecs.syr.edu/research/dreamsnet/publications/Kalyan_SIS%20_2003.pdf [doi: 10.1109/SIS.2003.1202264]

- [18] Mendes R, Kennedy J, Neves J. The fully informed particle swarm: Simpler, maybe better. IEEE Trans. on Evolutionary Computation, 2004,7(8):204-210. [doi: 10.1109/TEVC.2004.826074]
- [19] Liang JJ, Qin AK, Suganthan PN. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. IEEE Trans. on Evolutionary Computation, 2006,10(3):281-294. [doi: 10.1109/TEVC.2005.857610]

附中文参考文献:

- [7] 介婧,曾建潮,韩崇昭.基于群体多样性反馈控制的自组织微粒群算法.计算机研究与发展,2008,45(3):464-471.
- [8] 王沛.社会认知心理学.北京:中国社会科学出版社,2006.72-86.
- [15] 金欣磊,马龙华,吴铁军.基于随机过程的 PSO 收敛性分析.自动化学报,2007,33(12):1263-1268.



申元霞(1979-),女,安徽六安人,博士生,主要研究领域为计算智能,智能信息处理.



曾传华(1975-),男,讲师,主要研究领域为概率与统计理论,偏微分方程数值求解.



王国胤(1970-),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为粗糙集,粒计算,认知计算,智能信息处理,数据挖掘,智能信息安全.