

第3章

量子行为粒子群优化算法

在 PSO 算法中,粒子的运动状态由位置和速度描述,随着时间的演化,粒子的运动轨迹是既定的;同时粒子的速度受到一定限制,使得粒子的搜索空间是一个有限的并逐渐减小的区域,不能覆盖整个可行解空间,从而 PSO 算法不能保证全局收敛,这个结论已被 Frans Van de Bergh 所证明^[1,2]。针对 PSO 算法的这个缺点,根据粒子群的基本收敛性质^[3],利用量子力学中的相关理论作为背景知识,我们首先提出了基于 δ 势阱模型的量子行为粒子群优化(QDPSO)算法^[4];在此基础上,为了进一步完善 QDPSO 算法,针对算法的重要参数即波函数的特征长度的特性,设计了一种新的基于全局水平的参数控制方法,从而提出了量子行为粒子群优化(QPSO)算法^[5]。本章将首先给出量子力学的背景知识,然后分析 QPSO 算法的几种模型并进行比较,而后提出 QPSO 算法的进化方程,最后对 QPSO 算法的学习模式作了讨论。

3.1 量子力学背景

20 世纪的科学革命广泛地发生在宇观、宏观、微观三大层次上,使整个自然科学形成一个前沿不断扩大的多层次的综合的整体。概要说来,由相对论表征的科学革命是关于高速及宇观领域的;由量子力学和分子生物学表征的科学革命是关于微观领域的;由分形理论、混沌理论等一系列学科表征的科学革命,则是介于两者之间的宏观的领域。它们分别从宇观、微观、宏观三大层次上揭示了自然界的本质和规律^[6]。

1900 年普朗克提出的量子假说,1913 年玻尔建立的量子化的原子结构模型,1923 年德布罗意提出德物质波概念,1925 年海森堡建立的矩阵力学,1926 年薛定谔(Schrodinger, 1887—1961)建立的波动力学,以及之后玻恩对量子力学和波函数的统计诠释,揭示了崭新的、不同于宏观客体规律的微观客体规律,阐明了连续性与间断性、波动性与粒子性的对立统一,实现了量子(微观)世界的概率随机性,从而根本改变了精确确定的连续轨迹的经典概念,经典理论中的严格决定论,被因果律仅作为一种近似的和统计趋势的概率所替代^[7]。

经典力学中,一个粒子的运动状态,可用它在每一时刻 t 的坐标和动量(即相空间中一个点)给出确切的描述,而运动状态随着时间的演化,遵守 Newton 方程(或与之等价的正则方程等)。因此,如粒子在初始($t=0$)时刻的坐标和动量一经给定,则以后任何 $t>0$ 时刻粒子的运动状态就随之而定,这是一个确定性的(deterministic)描述。

不同于经典力学中的粒子,微观粒子的运动状态的描述方式及其随时间演化的规律,具有波动-粒子二象性(wave-particle duality)。粒子性,是微观粒子的“原子性”或“颗粒性”,并不意味着粒子在空间中的运动具有确切的轨道;波动性,是波动现象最本质的要素,即波的“相干叠加性”,并不意味着这种波动一定是某种实在的物理量的波动。能把实物粒子的“原子性”和波动的“相干叠加性”统一起来的,唯一自洽的方案是 M. Born 提出的“概率波”(probability wave)概念,即波函数的统计诠释。波函数所预言的概率分布,只是对粒子测量结果的一种预期(expectation),并非粒子已经具有那样的分布等待人们去观测它。

3.2 量子行为粒子群优化算法的基本模型

3.2.1 思想来源

J. Kennedy 和 R. Eberhart 在他们的著作 *Swarm Intelligence* 中论述道:“随机性的程度决定了智能的高低。”因此在他们提出的 PSO 算法中,随机因子被引入进化方程来体现鸟类等动物群体的智能性。这种随机的轨道搜索模型只能用来模拟低智能的动物群体,不能描述人类的群体智能行为。我们发现,人类的智能行为与量子空间中粒子的行为很相似。量子系统由于态叠加性而具有很强的不确定性,而人类思维也是具有不确定性,因而用量子模型描述人类思维和智能是合乎逻辑的,关键问题是如何建立一个有效量子模型。

研究表明,聚集性是群体智能最基本的特点。所谓聚集性就是群体中个体的差异是有限的,不可能趋向无穷大,这是群体智能最基本的特点。聚集性是由群体中的个体具有相互学习的特点决定的,个体的学习有以下特点:(1)追随性,即个体总是倾向于学习群体中最优的知识,这种性质使个体的差异减小;(2)记忆性,亦即个体在学习过程中,受到自身经验知识的约束,而这种特性使个体差异增加,由于这两个特性,个体在学习过程中同时受到群体最好知识和本身经验知识的影响,通过学习获得一种介于群体最好和个体经验之间的知识,但总体而言,具有这两种性质的学习可使个体间差异减少,群体多样性降低;(3)创造性,创造性使个体远离现有知识,使个体的差异扩大,群体多样性增加。聚集性是趋同和趋异两种趋势共同作用的结果,但趋同的趋势更大,否则就没有聚集性了。从算法的角度分析,追随性和记忆性的共同作用代表局部搜索能力,创造性代表全局搜索能力。

在考虑建立量子行为粒子群优化算法的模型时,决策变量同样用粒子的当前位置表示(用向量 X_i 表示),代表个体的当前思维状态;粒子经验中搜索到的具有最好适应值(目标函数值)的位置代表个体经验知识(即 $pbest$);当前群体中的具有最好适应值的粒子位置代表群体最好知识(即 $gbest$)。聚集性在力学中,用粒子的束缚态来描述。产生束缚态的原因是在粒子运动的中心存在某种吸引势场。为此可以建立一个量子化的吸引势场来束缚粒子(个体)以使群体具有聚集态。处于量子束缚态的粒子可以以一定的概率密度出现在空间任何点,它只要求当粒子与中心的距离趋向无穷时,概率密度趋近 0。因此量子模型的随

机性更大,关键问题就是如何建立以及采用何种形式的势能场。

M. Clerc 通过代数和数学分析方法,对 PSO 算法中粒子收敛行为进行了分析。研究表明,粒子 i 的收敛过程以点 $p_i = (p_{i1}, p_{i2}, \dots, p_{iN})$ 为吸引子,其坐标为:

$$p_{i,j}(t) = \frac{c_1 r_{1,i,j}(t) P_{i,j}(t) + c_2 r_{2,i,j}(t) G_j(t)}{c_1 r_{1,i,j}(t) + c_2 r_{2,i,j}(t)}, \quad 1 \leq j \leq N \quad (3.1)$$

或者

$$p_{i,j}(t) = \varphi_{i,j}(t) \cdot P_{i,j}(t) + [1 - \varphi_{i,j}(t)] \cdot G_j(t) \quad (3.2)$$

其中

$$\varphi_{i,j}(t) = c_1 r_{1,i,j}(t) / [c_1 r_{1,i,j}(t) + c_2 r_{2,i,j}(t)] \quad (3.3)$$

实际上,当 $c_1 = c_2$ 时, $\varphi_{i,j}(t)$ 本身就是一个区间 $(0, 1)$ 上均匀分布的随机数,即 $\varphi_{i,j}(t) \sim U(0, 1)$ 。因此在实际计算过程中可以直接由随机数发生器产生,这时式(3.2)和式(3.3)写为

$$p_{i,j}(t) = \varphi_{i,j}(t) \cdot P_{i,j}(t) + [1 - \varphi_{i,j}(t)] \cdot G_j(t), \quad \varphi_{i,j}(t) \sim U(0, 1) \quad (3.4)$$

在收敛过程中,随着速度的减小,粒子 i 不断地接近 p_i 点,最后跌落到 p_i 点。因此在整个过程中,在 p_i 点处实际上存在某种形式的吸引势吸引该粒子,这正是整个群体保持聚集性的原因。但由于在经典的 PSO 系统中,粒子的收敛是以轨道的形式实现的,并且粒子的速度总是有限的,因此在搜索过程中粒子每个迭代步的搜索空间是一个有限的区域,不能覆盖整个可行空间。因此一般的 PSO 算法不能保证以概率 1 收敛到全局最优解,这正是一般 PSO 算法的最大缺陷。而在量子空间中,粒子的聚集性通过在粒子运动中心存在的某种吸引势产生的束缚态来描述,而处于量子束缚态的粒子可以以一定的概率密度出现在空间任何点,满足聚集态的性质的粒子可以在整个可行解空间中进行搜索,但不会发散到无穷远处。

根据以上基本思想,考虑在 p_i 点建立一个吸引势来对群体中的粒子产生影响。通过比较发现, δ 势阱场可以产生比较好的效果,于是我们提出了基于 δ 势阱的量子行为的粒子群算法(QDPSO)。下面将详细描述如何通过建立 δ 势阱来得到粒子在量子空间中的运动方程,此方程也是 QDPSO 算法的粒子基本进化方程。

3.2.2 粒子群势阱模型的建立

在量子空间中,粒子的速度和位置是不能同时确定的,因此粒子的状态必须用波函数 $\Psi(X, t)$ 来描述,其中 $X = (x, y, z)$ 是粒子在三维空间中的位置向量。波函数的物理意义是:波函数模的平方是粒子在空间某一点出现的概率密度,即

$$|\Psi|^2 dx dy dz = Q dx dy dz \quad (3.5)$$

其中 Q 为概率密度函数。当然这个概率密度函数满足归一化条件,即

$$\int_{-\infty}^{+\infty} |\Psi|^2 dx dy dz = \int_{-\infty}^{+\infty} Q dx dy dz = 1 \quad (3.6)$$

在量子力学中粒子运动的动力学方程是 Schrödinger 方程,形式如下:

$$i\hbar \frac{\partial}{\partial t} \Psi(X, t) = \hat{H} \Psi(X, t) \quad (3.7)$$

其中, $\hat{H}$ 是哈密顿算子, $\hbar$ 称为普朗克常数。而哈密顿算子 $\hat{H}$ 具有以下形式:

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V(X) \quad (3.8)$$

其中, m 是粒子的质量, $V(X)$ 是粒子所在的势场。

现在假定粒子群系统是一个量子粒子系统, 每一个粒子具有量子行为, 由波函数来描述其状态。根据 PSO 算法中粒子收敛行为的分析, 必然存在以点 p_i 为中心某种形式的吸引势。为简单起见, 先考虑单个粒子在一维空间中运动的情形。设粒子质量为 m , 沿 X 轴运动, 势能为 $V(X)$ 。Schrödinger 方程为

$$i\hbar \frac{\partial}{\partial t} \Psi(X, t) = \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial X^2} + V(X) \right] \Psi(X, t) \quad (3.9)$$

假定粒子处于定态, 即具有一定能量的状态。波函数形式为

$$\Psi(X, t) = \phi(X) e^{-iEt/\hbar} \quad (3.10)$$

E 为相应的能量。式(3.10)代入式(3.9), $\phi(X)$ 满足

$$\frac{d^2 \phi(X)}{dX^2} + \frac{2m}{\hbar^2} [E - V(X)] \phi(X) = 0 \quad (3.11)$$

上述方程称为定态 Schrödinger 方程, $\phi(X)$ 称为定态波函数, 与 $\Psi(X, t)$ 只相差一个相位因子 $e^{-iEt/\hbar}$, 它们的模是相等的。因此对于定态问题, 只需研究 $\phi(X)$ 即可。

现将粒子的吸引子 p_i 记为 p , 粒子的位置为 X 。在 p 点建立一维 δ 势阱, 其势能函数表示为:

$$V(X) = -\gamma \delta(X - p)$$

做坐标变换 $Y = X - p$ 。则定态波函数表为 $\phi(Y)$ 势能函数变为

$$V(Y) = -\gamma \delta(Y)$$

因此该问题的哈密顿算子为:

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dY^2} - \gamma \delta(Y)$$

粒子在 δ 势阱中的定态 Schrödinger 方程为:

$$\frac{d^2 \phi}{dY^2} + \frac{2m}{\hbar^2} [E + \gamma \delta(Y)] \phi = 0 \quad (3.12)$$

其中 E 是粒子的能量。于是我们得到以下定理。

定理 3.1 粒子在以 p 点为中心的一维 δ 势阱中运动, 对应的定态 Schrödinger 方程的解为

$$\phi(Y) = \frac{1}{\sqrt{L}} e^{-|Y|/L} \quad (3.13)$$

其中, $L = 1/\beta = \hbar^2/m\gamma$ 。

证明：对式(3.12)两边求积分 $\int_{-\epsilon}^{+\epsilon} dY$, 当 $\epsilon \rightarrow 0^+$ 时可得：

$$\psi'(0^+) - \psi'(0^-) = -\frac{2m\gamma}{\hbar^2} \psi(0) \quad (3.14)$$

当 $Y \neq 0$, 式(3.12)可写为：

$$\frac{d^2 \psi}{dY^2} - \beta^2 \psi = 0 \quad (3.15)$$

其中 $\beta = \sqrt{-2mE/\hbar} (E < 0)$, 为了满足束缚态条件：

$$|Y| \rightarrow \infty, \quad \psi \rightarrow 0 \quad (3.16)$$

式(3.15)的解必须具有以下形式：

$$\psi(Y) \approx e^{-\beta|Y|} (Y \neq 0) \quad (3.17)$$

由于波函数必须满足束缚态条件(3.16), 因此式(3.15)的解为：

$$\psi(Y) = \begin{cases} Ae^{-\beta Y} & Y > 0 \\ Ae^{\beta Y} & Y < 0 \end{cases} \quad (3.18)$$

其中 A 是归一化常数。根据条件式(3.14)可得：

$$-2A\beta = -\frac{2m\gamma}{\hbar^2} A \quad (3.19)$$

这样得到

$$\beta = m\gamma / \hbar^2 \quad (3.20)$$

和

$$E = E_0 = -\frac{\hbar^2 \beta^2}{2m} = -\frac{m\gamma^2}{2\hbar^2} \quad (3.21)$$

函数 $\psi(Y)$ 需满足归一化条件以决定常数 A , 即：

$$\int_{-\infty}^{+\infty} |\psi(Y)|^2 dY = |A|^2 / \beta = 1 \quad (3.22)$$

于是得到 $|A| = \sqrt{\beta}$, $L = 1/\beta = \frac{\hbar^2}{m\gamma}$ 是 δ 势阱的特征长度。代入式(3.18), 则归一化波函数表示为：

$$\psi(Y) = \frac{1}{\sqrt{L}} e^{-|Y|/L}$$

证毕。 ■

由此我们可以得到相应的概率密度函数 Q 为：

$$Q(Y) = |\psi(Y)|^2 = \frac{1}{L} e^{-2|Y|/L} \quad (3.23)$$

概率分布函数 F 为：

$$F(Y) = 1 - e^{-2|Y|/L} \quad (3.24)$$

3.2.3 粒子的基本进化方程

从定理 3.1 的证明过程,我们得到了粒子在一维 δ 势阱的量子束缚态波函数和相应的位置概率分布函数。在算法的实际设计中,为评价适应值(目标函数值),我们需要了解粒子精确的位置。而量子状态函数 $\psi(Y)$ 仅仅给出粒子出现在相对于 p 点位置 Y 的概率密度函数 $|\psi(Y)|^2$ 或 $Q(Y)$ 。因此为了给出粒子的精确位置,必须将量子状态塌缩到经典状态,这可以通过蒙特卡罗随机模拟的方式来测量粒子的位置,这种方法好比对粒子进行拍照,在按快门的一瞬间粒子的位置就被定格在照片上了。这里我们采用的蒙特卡罗方法又称为逆变换法,这是针对概率分布函数形式比较简单的情况下最常用的随机模拟方法。推导过程如定理 3.2。

定理 3.2 粒子在以 p 点为中心的一维 δ 势阱中运动,其位置由以下随机方程确定,即:

$$X = p \pm \frac{L}{2} \ln(1/u) \quad (3.25)$$

其中, $L=1/\beta=\hbar^2/m\gamma$, u 为区间 $(0,1)$ 上的均匀分布随机数,即 $u \sim U(0,1)$ 。

证明: 令 v 是在区间 $(0,1)$ 上均匀分布的随机数,即

$$v \sim U(0,1) \quad (3.26)$$

用 v 代替式(3.24)中的左边,即令:

$$1-v = 1-F(Y) \quad (3.27)$$

由于 $1-v \sim U(0,1)$, 令 $u=1-v$, 则有 $u \sim U(0,1)$ 。于是可以得到:

$$u = e^{-2|Y|/L} \quad (3.28)$$

用逆变换求出 Y

$$Y = \pm \frac{L}{2} \ln(1/u) \quad (3.29)$$

由于 $Y=X-p$, 因此我们可以测量粒子的位置的随机方程:

$$X = p \pm \frac{L}{2} \ln(1/u) \quad (3.30)$$

其中 u 是 $(0,1)$ 区间的均匀分布的随机数。

证毕。 ■

式(3.25)是量子行为粒子群算法的基本进化方程,其中 L 是 δ 势阱的特征长度,是进化方程中最重要的参量,选择 L 的控制方式是设计算法很重要的一步,具体的方法我们将在后面叙述,接下来先探讨为什么选择 δ 势阱模型。

3.2.4 δ 势阱模型与其他模型的比较

除了 δ 量子势阱模型,还有许多其他量子势模型可用于 PSO 算法。以量子谐振子 (Quantum harmonic Oscillator, QO) 模型为例。QO 模型的波函数为:

$$\psi_0(Y) = \frac{\sqrt{\sigma}}{\pi^{\frac{1}{4}}} e^{-\frac{1}{2}\sigma^2 Y^2} \quad (3.31)$$

概率密度函数 Q 为:

$$Q(Y) = |\psi_0(Y)| = \frac{\sigma}{\sqrt{\pi}} e^{-\sigma^2 Y^2} \quad (3.32)$$

概率分布函数为

$$F(Y) = \int_{-\infty}^Y \frac{\sigma}{\sqrt{\pi}} e^{-\sigma^2 Y^2} dY \quad (3.33)$$

通过蒙特卡罗随机模拟,粒子的位置为:

$$Y = \pm \frac{1}{\sigma} \lambda \quad (3.34)$$

令 $Y = X - p$ 得:

$$X = p \pm \frac{1}{\sigma} \lambda \quad (3.35)$$

其中 σ 是特征长度, λ 是服从标准正态分布的随机数, 即 $\lambda \sim N(0, 1)$ 。

图 3.1 给出了量子谐振子模型和 δ 势阱模型的概率密度函数的波形图。

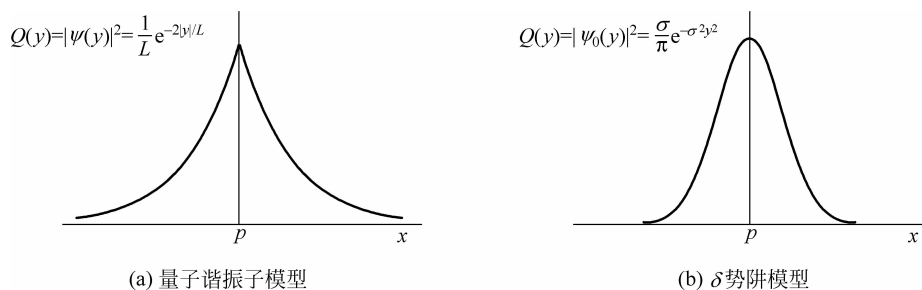


图 3.1 概率密度函数波形图

由图 3.1 可见,量子谐振子模型的概率密度函数相对于 δ 势阱模型的概率密度函数较窄,粒子以较高的概率出现在 p 点附近,这意味着在参数控制相同的情况下,粒子以谐振子模型的进化方程(式 3.35)运行将比以 δ 势阱模型的进化方程(式 3.25)运行的收敛速度快。如果运用到算法中,这样的收敛速度是以牺牲算法的全局搜索能力而获得的。我们对以式(3.35)为进化方程设计的量子行为粒子群优化算法进行测试函数求解,结果显示这种算法容易早熟。而其他的势能模型有的波函数太复杂难以模拟,有的易早熟,因此,我们选择基于 δ 势阱模型提出了量子行为粒子群优化算法。

3.3 量子行为粒子群优化算法

3.3.1 粒子的进化方程

3.2 节阐述了量子行为粒子群优化算法的基本思想、 δ 势阱模型的建立、粒子基本进化

方程以及收敛性条件,但这些都是在一维情形下考虑的,现在我们将以上结果扩展到 N 维搜索空间的情况。

对于粒子 i ,将上面的 p 点,即吸引子写成 $p_i = (p_{i,1}, p_{i,2}, \dots, p_{i,N})$ 。在每一维上以 p_i 坐标 $p_{i,j}$ 为中心建立一个一维 δ 势阱,应用定理 3.1,考虑到粒子位置坐标 $X_{i,j}$ 随时间变化,对于给定的 $p_{i,j}$,可以得到粒子 i 每一维的波函数为

$$\psi[X_{i,j}(t+1)] = \frac{1}{\sqrt{L_{i,j}(t)}} \exp[-|X_{i,j}(t+1) - p_{i,j}(t)| / L_{i,j}(t)] \quad (3.36)$$

相应的概率密度函数和概率分布函数分别为

$$\begin{aligned} Q[X_{i,j}(t+1)] &= |\psi[X_{i,j}(t+1)]|^2 \\ &= \frac{1}{L_{i,j}(t)} \exp[-2|X_{i,j}(t+1) - p_{i,j}(t)| / L_{i,j}(t)] \end{aligned} \quad (3.37)$$

$$F[X_{i,j}(t+1)] = 1 - \exp[-2|X_{i,j}(t+1) - p_{i,j}(t)| / L_{i,j}(t)] \quad (3.38)$$

确切地讲,以上波函数、概率密度函数和概率分布函数都应该写成条件概率的形式,即

$$\psi[X_{i,j}(t+1) | p_{i,j}(t)] = \frac{1}{\sqrt{L_{i,j}(t)}} \exp[-|X_{i,j}(t+1) - p_{i,j}(t)| / L_{i,j}(t)] \quad (3.39)$$

$$\begin{aligned} Q[X_{i,j}(t+1) | p_{i,j}(t)] &= |\psi[X_{i,j}(t+1) | p_{i,j}(t)]|^2 \\ &= \frac{1}{L_{i,j}(t)} \exp[-2|X_{i,j}(t+1) - p_{i,j}(t)| / L_{i,j}(t)] \end{aligned} \quad (3.40)$$

$$F[X_{i,j}(t+1) | p_{i,j}(t)] = 1 - \exp[-2|X_{i,j}(t+1) - p_{i,j}(t)| / L_{i,j}(t)] \quad (3.41)$$

但对于给定了的 $p_{i,j}$ (由式(3.4)确定),两种写法对于用 Monte Carlo 法求粒子位置的坐标是没有影响的。这样对于粒子 i 的每一维应用定理 3.2,可得第 j 维的坐标基本进化方程:

$$X_{i,j}(t+1) = p_{i,j}(t) \pm \frac{L_{i,j}(t)}{2} \ln[1/u_{i,j}(t)] \quad u_{i,j}(t) \sim U(0,1) \quad (3.42)$$

同样对于每一维应用定理 3.3,很容易得到以下单个粒子收敛条件。

定理 3.3 N 维空间中,以式(3.42)进化的粒子 i 的位置依概率收敛到其吸引子 $p_i = (p_{i,1}, p_{i,2}, \dots, p_{i,N})$ 的充分必要条件是,每一维坐标 $X_{i,j}(t)$ 都依概率收敛于 $p_{i,j}$ 。

证明: (1) 必要性。设 $X_i(t) \xrightarrow{P} p_i(p_{i,1}, p_{i,2}, \dots, p_{i,N})$, 即对于任意 $\epsilon > 0$, 都有

$$\lim_{t \rightarrow \infty} P(|X_i(t) - p_i| < \epsilon) = 1 \quad (3.43)$$

对于每一个 $j \in \{1, 2, \dots, N\}$, 由于

$$|X_{i,j}(t) - p_{i,j}| \leq \left(\sum_{j=1}^N |X_{i,j}(t) - p_{i,j}|^2 \right)^{\frac{1}{2}} = |X_i(t) - p_i|$$

因此 $|X_i(t) - p_i| < \epsilon$ 必有 $|X_{i,j}(t) - p_{i,j}| < \epsilon$, 即有以下事件的包含关系

$$\{|X_i(t) - p_i| < \epsilon\} \subset \{|X_{i,j}(t) - p_{i,j}| < \epsilon\}$$

从而

$$P\{|X_{i,j}(t) - p_{i,j}| < \epsilon\} \geq P\{|X_i(t) - p_i| < \epsilon\}$$

上式两边求极限得

$$\lim_{t \rightarrow \infty} P\{|X_{i,j}(t) - p_{i,j}| < \epsilon\} \geq \lim_{t \rightarrow \infty} P\{|X_i(t) - p_i| < \epsilon\} = 1$$

则必有

$$\lim_{t \rightarrow \infty} P\{|X_{i,j}(t) - p_{i,j}| < \epsilon\} = 1$$

即 $X_{i,j}(t) \xrightarrow{P} p_{i,j}$ 。

(2) 充分性。设 $X_{i,j}(t) \xrightarrow{P} p_{i,j}$, 即对于 $j \in \{1, 2, \dots, N\}$, 以及任意 $\epsilon > 0$, 都有

$$\lim_{t \rightarrow \infty} P\left\{|X_{i,j}(t) - p_{i,j}| < \frac{\epsilon}{\sqrt{N}}\right\} = 1$$

如果 $|X_{i,j}(t) - p_{i,j}| < \frac{\epsilon}{\sqrt{N}}$, 必有

$$|X_i(t) - p_i| = \left(\sum_{j=1}^N |X_{i,j}(t) - p_{i,j}|^2\right)^{\frac{1}{2}} < \left[N \cdot \left(\frac{\epsilon}{\sqrt{N}}\right)^2\right]^{\frac{1}{2}} = \epsilon$$

从而

$$\left\{|X_{i,j}(t) - p_{i,j}| < \frac{\epsilon}{\sqrt{N}}\right\} \subset \{|X_i(t) - p_i| < \epsilon\}$$

因此

$$P\{|X_i(t) - p_i| < \epsilon\} \geq P\left\{|X_{i,j}(t) - p_{i,j}| < \frac{\epsilon}{\sqrt{N}}\right\}$$

两边求极限可得

$$\lim_{t \rightarrow \infty} P\{|X_i(t) - p_i| < \epsilon\} \geq \lim_{t \rightarrow \infty} P\left\{|X_{i,j}(t) - p_{i,j}| < \frac{\epsilon}{\sqrt{N}}\right\} = 1$$

故 $\lim_{t \rightarrow \infty} P\{|X_i(t) - p_i| < \epsilon\} = 1$, 即 $X_i(t) \xrightarrow{P} p_i$ 。

证毕。 ■

定理 3.3 表明 N 维空间中, 满足式(3.42)的单个粒子位置的收敛性可以归结为一维空间的粒子收敛性, 这是一个很重要的结论。后面将会证明粒子的位置依概率收敛到吸引子的充要条件为 $\lim_{t \rightarrow \infty} L_{i,j}(t) = 0$ 。

3.3.2 两种搜索迭代策略

有了粒子的基本进化方程(3.42), 接下来的关键问题是如何控制 $L_{i,j}(t)$, 使得它能够收敛到 0, 从而使粒子收敛到局部吸引子 p_i 。但最终所有粒子的吸引子 $p_i (i=1, 2, \dots, m)$, 个体最好位置 P_i , 以及全局最好位置 G 都将收敛到一点, 这与基本 PSO 算法是一样的。与遗传算法有所不同, 如果变异概率不为零的话, 最后所有的个体是不会收敛到一点的。但

粒子收敛于局部吸引子,或者说粒子群收敛于一点不是算法全局收敛性的充要条件,但它能使算法有更好的收敛速度和精度,其原因不仅牵涉算法收敛性分析,更涉及算法的搜索机制和复杂度的研究。目前我们对 QPSO 算法和 PSO 算法的搜索机制的研究,两者与十进制遗传算法搜索机制的比较,以及 QPSO 算法的复杂度研究有了一定的成果,但本书限于篇幅不讨论这两方面的问题。

关于 $L_{i,j}(t)$ 的控制,一种容易想到的方法是像控制模拟退火算法中的温度一样,让 $L_{i,j}(t)$ 随迭代次数即时间不断减小,比如线性减小的方式。但实验结果表明,这种方法效果并不好,主要是对不同的问题,用相同的控制方法优化结果大相径庭。原因是这种方法可能缺乏自适应性,到搜索的后期,粒子跳出局部最优解的概率下降得太快。这里我们提出了两种方法,下面将进行详细阐述。

第一种方法是用以下公式来评价 $L_{i,j}(t)$,即

$$L_{i,j}(t) = 2\alpha \cdot |p_{i,j}(t) - X_{i,j}(t)| \quad (3.44)$$

则粒子的进化方程变为

$$X_{i,j}(t+1) = p_{i,j}(t) \pm \alpha \cdot |p_{i,j}(t) - X_{i,j}(t)| \cdot \ln[1/u_{i,j}(t)] \quad u_{i,j}(t) \sim U(0,1) \quad (3.45)$$

第二种方法是在算法中引入了平均最好位置(mean best position, mbest),记为 $C(t)$ 。它定义为所有粒子个体最好位置的平均,即

$$\begin{aligned} C(t) = (C_1(t), C_2(t), \dots, C_n(t)) &= \frac{1}{M} \sum_{i=1}^M P_i(t) \\ &= \left(\frac{1}{M} \sum_{i=1}^M P_{i,1}(t), \frac{1}{M} \sum_{i=1}^M P_{i,2}(t), \dots, \frac{1}{M} \sum_{i=1}^M P_{i,N}(t) \right) \end{aligned} \quad (3.46)$$

$L_{i,j}(t)$ 通过以下公式来评价

$$L_{i,j}(t) = 2\alpha \cdot |C_j(t) - X_{i,j}(t)| \quad (3.47)$$

于是粒子的进化公式变为

$$X_{i,j}(t+1) = p_{i,j}(t) \pm \alpha \cdot |C_j(t) - X_{i,j}(t)| \cdot \ln[1/u_{i,j}(t)] \quad u_{i,j}(t) \sim U(0,1) \quad (3.48)$$

式(3.44)、式(3.45)、式(3.47)和式(3.48)中的 α 称为收缩-扩张系数(Contraction-Expansion coefficient, CE coefficient),它是算法除群体规模和迭代次数以外的唯一控制参数。对参数 α 控制可以采用固定取值和线性减小的方式,详细情况可参考文献[8]。

我们把应用公式(3.44)的粒子群优化算法称为基于 δ 势阱模型的量子行为粒子群优化(QDPSO)算法,而把应用式(3.48)的粒子群优化算法称为量子行为粒子群优化(Quantum-behaved Particle Swarm Optimization, QPSO)算法。通过对测试函数和其他应用问题的结果比较,后者的性能远优于前者,所以现在使用的 QPSO 算法都指的是后者。进行单个粒子收敛行为分析时,前者较容易分析并且是后者的基础。在第4章,我们还将对这两类 QPSO 算法的收敛率进行比较。