

通俗理解粒子群优化算法

原创 2017-12-01 火星十一郎 机器学习算法与自然语言处理

推荐阅读时间8min~12min
主要内容：粒子群优化算法简介

1 背景介绍

人工生命

人工生命：研究具有某些生命基本特征的人工系统。包括两方面的内容：

- 1、研究如何利用计算技术研究生物现象；
- 2、研究如何利用生物技术研究计算问题。

我们关注的是第二点。已有很多源于生物现象的计算技巧，例如神经网络和遗传算法。现在讨论另一种生物系统---社会系统：由简单个体组成的群落和环境及个体之间的相互行为。

群智能

模拟系统利用局部信息从而可以产生不可预测的群行为。我们经常能够看到成群的鸟、鱼或者浮游生物。这些生物的聚集行为有利于它们觅食和逃避捕食者。它们的群落动辄以十、百、千甚至万计，并且经常不存在一个统一的指挥者。它们是如何完成聚集、移动这些功能呢？

Millonas在开发人工生命算法时(1994年)，提出群体智能概念并提出五点原则：

- 1、接近性原则：群体应能够实现简单的时空计算；
- 2、优质性原则：群体能够响应环境要素；
- 3、变化相应原则：群体不应把自己的活动限制在一狭小范围；
- 4、稳定性原则：群体不应每次随环境改变自己的模式；
- 5、适应性原则：群体的模式应在计算代价值得的时候改变。

模拟群

对鸟群行为的模拟：Reynolds、Heppner和Grenader提出鸟群行为的模拟。他们发现，鸟群在行进中会突然同步的改变方向，散开或者聚集等。那么一定有某种潜在的能力或规则保证了这些同步的行为。这些科学家都认为上述行为是基于不可预知的鸟类社会行为中的群体动态学。在这些早期的模型中仅仅依赖个体间距的操作，也就是说，这中同步是鸟群中个体之间努力保持最优的距离的结果。

对鱼群行为的研究：生物社会学家E.O.Wilson对鱼群进行了研究。提出：“至少在理论上，鱼群的个体成员能够受益于群体中其他个体在寻找食物的过程中的发现和以前的经验，这种受益超过了个体之间的竞争所带来

的利益消耗，不管任何时候食物资源不可预知的分散。”这说明，同种生物之间信息的社会共享能够带来好处。这是PSO的基础。

2 算法介绍

粒子群优化算法的基本思想是通过群体中个体之间的协作和信息共享来寻找最优解。

PSO的优势在于简单容易实现并且没有许多参数的调节。目前已被广泛应用于函数优化、神经网络训练、模糊系统控制以及其他遗传算法的应用领域。

问题提出

设想这样一个场景：一群鸟在随机的搜索食物。在这个区域里只有一块食物，所有的鸟都不知道食物在那。但是它们知道自己当前的位置距离食物还有多远。那么找到食物的最优策略是什么？最简单有效的就是搜寻目前离食物最近的鸟的周围区域。

问题抽象

鸟被抽象为没有质量和体积的微粒(点)，并延伸到N维空间，粒子 i 在N维空间的位置表示为矢量 $X_i = (x_1, x_2, \dots, x_N)$ ，飞行速度表示为矢量 $V_i = (v_1, v_2, \dots, v_N)$ 。每个粒子都有一个由目标函数决定的适应值(fitness value)，并且知道自己到目前为止发现的最好位置(pbest)和现在的位置 X_i 。这个可以看作是粒子自己的飞行经验。除此之外，每个粒子还知道到目前为止整个群体中所有粒子发现的最好位置(gbest)(gbest是pbest中的最好值)。这个可以看作是粒子同伴的经验。粒子就是通过自己的经验和同伴中最好的经验来决定下一步的运动。

算法描述

PSO初始化为一群随机粒子(随机解)。然后通过迭代找到最优解。在每一次的迭代中，粒子通过跟踪两个“极值”(pbest,gbest)来更新自己。

在找到这两个最优值后，粒子通过下面的公式来更新自己的速度和位置。

方程中pbest和gbest分别表示微粒群的局部和 全局最优位置，当C1=0时，则粒子没有了认知能力， 变为只有社会的模型(social-only):

$$V_i = \omega \times V_i + c_2 \times rand() \times (gbest_i - x_i)$$

被称为全局PSO算法.粒子有扩展搜索空间的能力，具有 较快的收敛速度，但由于缺少局部搜索，对于复杂问题 比标准PSO 更易陷入局部最优。

当C2=0时，则粒子之间没有社会信息，模型变为 只有认知 (cognition-only)模型：

$$V_i = \omega \times V_i + c_1 \times rand() \times (pbest_i - x_i)$$

算法优化

1998年shi等人在进化计算的国际会议上 发表了一篇论文《A modified particle swarm optimizer》对前面的公式进行了修正。引入 惯性权重因子。值较大，全局寻优能力强，局部寻优能力弱； 值较小反之。

初始时，shi将 ω 取为常数，后来实验发现，动 态 能够获得比固定值更好的寻优结果。动态 可以在PSO搜索过程中线性变化，也可根据PSO 性能的某个测度函数动态改变。目前，采用较多的是shi建议的线性递减权值 (linearly decreasing weight, LDW)策略。

标准PSO算法的流程：

Step1:初始化一群微粒(群体规模为m)，包括随机位置和 速度；

Step2:评价每个微粒的适应度；

Step3:对每个微粒，将其适应值与其经过的最好位置 pbest作比较，如果较好，则将其作为当前的 最好位置pbest;

Step4:对每个微粒，将其适应值与其经过的最好位置 gbest作比较，如果较好，则将其作为当前的 最好位置gbest;

Step5:根据(2)、(3)式调整微粒速度和位置；

Step6：未达到结束条件则转Step2。

迭代终止条件根据具体问题一般选为最大迭代次数Gk或(和)微粒群迄今为止搜索到的最优位置满足预定最小适应阈值。

参数分析

方程中pbest和gbest分别表示微粒群的局部和 全局最优位置，当 $C1=0$ 时，则粒子没有了认知能力， 变为只有社会的模型(social-only):

$$V_i = \omega \times V_i + c_2 \times rand() \times (gbest_i - x_i)$$

被称为全局PSO算法.粒子有扩展搜索空间的能力，具有 较快的收敛速度，但由于缺少局部搜索，对于复杂问题 比标准PSO 更易陷入局部最优。

当 $C2=0$ 时，则粒子之间没有社会信息，模型变为 只有认知(cognition-only)模型：

$$V_i = \omega \times V_i + c_1 \times rand() \times (pbest_i - x_i)$$

被称为局部PSO算法。由于个体之间没有信息的 交流，整个群体相当于多个粒子进行盲目的随机 搜索，收敛速度慢，因而得到最优解的可能性小。

群体规模m 一般取20 ~ 40，对较难或特定类别的问题 可以取到100 ~ 200。

最大速度Vmax决定当前位置与最好位置之间的区域的 分辨率(或精度)。如果太快，则粒子有可能越过极小点;如果太慢，则粒子不能在局部极小点之外进行足 够的探索，会陷入到局部极值区域内。这种限制可以 达到防止计算溢出、决定问题空间搜索的粒度的目的。

权重因子 包括惯性因子 和学习因子 $c1$ 和 $c2$ 。使粒子 保持着运动惯性，使其具有扩展搜索空间的趋势，有能力探索新的区域。 $C1$ 和 $c2$ 代表将每个粒子推向Pbest 和gbest位置的统计加速项的权值。较低的值允许粒子在被拉回之前可以在目标区域外徘徊，较高的值导致粒 子突然地冲向或越过目标区域。

3 优化PSO

引入收敛因子，不要惯性权重

通常设 $c1 = c2 = 2$ 。Suganthan的实验表明： $c1$ 和 $c2$ 为常数时可以得到较好的解，但不一定必须等于2。Clerc引入收敛因子(constriction factor) K来保证 收敛性。

$$V_i = K [V_i + \varphi_1 \text{rand}() (pbest_i - x_i) + \varphi_2 \text{rand}() (gbest_i - x_i)]$$

其中

$$K = \frac{2}{2 - \varphi - \sqrt{\varphi^2 - 4\varphi}}, \varphi = \varphi_1 + \varphi_2, \varphi > 4.$$

通常取为4.1,则 $K = 0.729$.实验表明,与使用惯性权重的PSO算法相比,使用收敛因子的PSO有更快的收敛速度。其实只要恰当的选取 φ_1 、 φ_2 ,两种算法是一样的。因此使用收敛因子的PSO可以看作使用惯性权重PSO的特例。恰当的选取算法的参数值可以改善算法的性能。

离散二进制粒子群

基本PSO是用于实值连续空间,然而许多实际问题是组合优化问题,因而提出离散形式的PSO。速度和位置更新式为:



PSO与GA比较

共性: (1) 都属于仿生算法。(2) 都属于全局优化方法。(3) 都属于随机搜索算法。(4) 都隐含并行性。(5) 根据个体的适配信息进行搜索,因此不受函数约束条件的限制,如连续性、可导性等。(6) 对高维复杂问题,往往会遇到早熟收敛和收敛性能差的缺点,都无法保证收敛到最优点。

差异: (1) PSO有记忆,好的解的知识所有粒子都保存,而GA,以前的知识随着种群的改变被改变。(2) PSO中的粒子仅仅通过当前搜索到最优点进行共享信息,所以很大程度上这是一种单共享项信息机制。而GA中,染色体之间相互共享信息,使得整个种群都向最优区域移动。(3) GA的编码技术和遗传操作比较简单,而PSO相对于GA,没有交叉和变异操作,粒子只是通过内部速度进行更新,因此原理更简单、参数更少、实现更容易。

GA可以用来研究NN的三个方面:网络连接权重、网络结构、学习算法。优势在于可处理传统方法不能处理的问题,例如不可导的节点传递函数或没有梯度信息。缺点:在某些问题上性能不是特别好;网络权重的编码和遗传算子的选择有时较麻烦。已有利用PSO来进行神经网络训练。研究表明PSO是一种很有潜力的神经网络算法。速度较快且有较好的结果。且没有遗传算法碰到的问题。

4 PSO实现



各算法对应的问题如下：

- PSO 用基本粒子群算法求解无约束优化问题
- YSPSO 用带压缩因子的粒子群算法求解无约束优化问题
- LinWPSO 用线性递减权重粒子群优化算法求解无约束优化问题
- SAPSO 用自适应权重粒子群优化算法求解无约束优化问题
- RandWPSO 用随机权重粒子群优化算法求解无约束优化问题
- LnCPSO 用学习因子同步变化的粒子群优化算法求解无约束优化问题
- AsyLnCPSO 用学习因子异步变化的粒子群优化算法求解无约束优化问题
- SecPSO 用二阶粒子群优化算法求解无约束优化问题
- SecVibratPSO 用二阶振荡粒子群优化算法求解无约束优化问题
- CLSPSO 用混沌粒子群优化算法求解无约束优化问题
- SelfPSO 用基于选择的粒子群优化算法求解无约束优化问题
- BreedPSO 用基于交叉遗传的粒子群优化算法求解无约束优化问题
- SimuAPSO 用基于模拟退火的粒子群优化算法求解无约束优化问题

csdn链接（包含了基本PSO和12种优化PSO算法，绝对能用）。

参考文献：西电姚新正老师课件

推荐阅读：

[精选干货|近半年干货目录汇总](#)

[干货|掌握机器学习数学基础之优化\[1\]（重点知识）](#)

[【直观详解】什么是PCA、SVD](#)

欢迎关注公众号学习交流~



长按二维码扫描关注

机器学习算法与自然语言处理

ID: yizhenotes

通俗笔记， 分享交流

欢迎加入交流群交流学习~



机器学习&NLP3群

扫一扫二维码，加入该群。