

随机游走

李昱辰 (21721221)

2018 年 5 月 14 日

摘要

本文将从随机游走算法的定义、思想、算法流程等方面全面分析这种优化算法，并阐述我对这种算法的理解以及这种算法在实际工作中的部分应用场景。

1 定义

随机游走 (random walk) 也称随机漫步，随机行走等是指基于过去的表现，无法预测将来的发展步骤和方向。核心概念是指任何无规则行走者所带的守恒量都各自对应着一个扩散运输定律，接近于布朗运动，是布朗运动理想的数学状态，现阶段主要应用于互联网链接分析及金融股票市场中。

2 思想

随机游走算法的设计思想主要有三点，分别是全局最优化，逻辑简单，不易陷入局部最优值。

首先，全局最优化。在优化问题中，需要求解的最终结果是特定条件下的最优值。在优化算法的发展历程中，出现了种类繁多的各类优化方式，但至今并没有出现一种可以对任意复杂函数都能保证求解最优值的通用算法，随机游走算法的设计初衷是求解全局最优值，但与其他优化算法相同，在一些情况下它会面临着失效的可能性，所以在实际使用中需要通过测试数据来验证随机游走算法在特定条件下是否能有效发挥作用。

第二，逻辑简单。随机游走算法中不存在非常复杂的逻辑思想，核心思想可以概括为随机和迭代。所以，这种算法在各种应用场景下都有着易于快速实现、易于定位问题等优势。

第三，不易陷入局部最优值。各类优化算法普遍面临着一个不可忽视的问题，就是陷入局部最优值。如果一类算法在某类应用场景下频繁陷入局部最优值，那么这种优化算法就是没有应用价值的。随机游走算法依托其不断迭代验证的特征，将陷入局部最优值的可能性降到一个较低的水平，这本质上得益于随机游走算法运行过程中频繁的验证机制。

3 算法基本原理

随机游走是一种“尝试型”的优化算法，也就是说，它与梯度下降等优化算法不同的是，它对于下一个取值的选取是没有具体规则的，下一个取值是更好还是更坏，只有在取完之后才知道，然后根据实际情况去进行实时判断，从而决定下一步的运行方式。

在一些时候，可能会出现一种情况，就是以指定步长进行随机探索的过程中，足够长的一段时间里都无法找到更优质的点，一般来说，可能有两种原因导致了这种现象的发生：

一、步长过大，超过精度要求过多；

二、步长在精度要求范围内，当前点与实际上的全局最优点非常近。

针对第一种情况，随机游走算法采取的策略是将步长减为原有的一半，然后继续迭代寻找最优点；针对第二种情况，随机游走算法会将当前所在点近似认为是全局最优点。由此也可以看出，随机游走算法的最终结果是一个误差在精度要求范围内的近似值，而非精确值。

4 算法流程

- 1、设 $f(x)$ 是一个含有 n 个变量的多元函数, $x=(x_1, x_2, \dots, x_n)$ 为 n 维向量。
- 2、给定初始迭代点 x ，初次行走步长 λ ，控制精度 ϵ 。
- 3、给定迭代控制次数 N ， k 为当前迭代次数，置 $k=1$ 。
- 4、当 $k < N$ 时，随机生成一个 $(-1, 1)$ 之间的 n 维向量 $u=(u_1, u_2, \dots, u_n)$, $(-1 < u_i < 1, i=1, 2, \dots, n)$ 。
- 5、将 n 维向量标准化得到

$$u' = \frac{u}{\sqrt{\sum_{i=1}^n u_i^2}} \quad (1)$$

- 6、令

$$x_1 = x + \lambda u' \quad (2)$$

从而完成第一步游走。

7、计算函数值，如果 $f(x_1) < f(x)$ ，即找到了一个比初始值好的点，那么 k 重新置为 1，将 x_1 变为 x ，回到第 2 步；否则 $k=k+1$ ，回到第 3 步。

8、如果连续 N 次都找不到更优的值，则认为，最优解就在以当前最优解为中心，当前步长为半径的 N 维球内。

9、此时，如果 $\lambda < \epsilon$ ，则结束算法；否则，令 λ 减半，回到第 1 步，开始新一轮游走。

5 算法应用经验

在优化算法的应用过程中，往往需要一些针对性的经验来使得算法的使用效果更好。在随机游走算法领域，最显著的一种经验丰富的体现就是对于初始步长的选取。这其中的原因是，如果初始步长的选取过大，则初始搜索范围将会很大，从而导致算法的收敛非常慢；如果初始步长的选取过小，则容易导致随机游走算法难以定位到全局最优点。一般来说，在实际应用过程中，在硬件条件允许的情况下，会选取一个较大的初始步长，力求最终的结果足够贴近全局最优值。

6 算法应用场景综述

实际上，随机游走算法的深刻思想是可以在物理学中找到依据的。随机游走算法的核心概念是指任何无规则行走者所带的守恒量都各自对应着一个扩散运输定理，这是接近于布朗运动的，实际上就是布朗运动理想的数学状态。现阶段，随机游走算法主要应用于各类互联网技术与金融股票市场中。

7 互联网技术应用实例

在互联网技术中，体现随机游走算法能效的一个典型例子就是对等计算搜索。在对等计算搜索过程中，请求者发出若干个查询请求给随机挑选的若干个相邻节点，然后每个查询请求在之后的随机游走过程中保持与请求者的信息传输，这种信息传输主要是用来决定相应的查询请求是否还需要继续进行随机游走。如此迭代，直到完成搜索过程。

除了上述的对等计算搜索之外，许多链接分析算法中也充分体现了随机游走算法的思想，例如应用广泛的 PageRank 算法等。

8 金融股票市场应用实例

随机游走算法的本质思想是金融数学领域的一项重要假设，它把市场中的各类随机变化与物理学中的布朗运动建立起了联系，在这个基础之上，出现了许多以随机过程为核心的数学算法，这类数学算法一同组成了一类重要的金融领域的数学分析策略。例如，英国统计学家肯德尔曾经提出过一个有关股票价格波动模式的重要结论：股票价格的变动实际上没有什么规律可循，就像一个随意走动的醉汉，每周会出现一个随机数字，加在股价之上，从而形成下周的股价。这个结论实际上就是指股票价格波动模式遵从的就是一种随机游走模式。

9 总结

随机游走算法作为一种算法逻辑比较简洁的全局最优化算法，虽然在实现上没有什么特别的难点，但各类基于基础随机游走算法进行改进的数学模型却在很多领域中得到了成功的应用。这种算法本质上将许多领域中的数学规律与物理学中的布朗运动联系在了一起，体现了数学和自然界之间精妙的联系，并以此为思想基础，解决了各个领域中的一些显著难题。