

基于旅行商问题的顺序添加研究

曾诗棋, 谢佳艳*

吉首大学数学与统计学院, 湖南 吉首

收稿日期: 2025年3月16日; 录用日期: 2025年4月9日; 发布日期: 2025年4月18日

摘 要

针对复杂多因子试验顺序优化中存在的路径成本高、实验迭代次数多等难题, 本文提出一种基于邻接关系模型的双循环排序算法。通过构建全循环模型矩阵与基线约束条件, 结合主因子分割策略, 实现了最优试验序列的快速确定与实验次数的最小化。理论分析表明: 该算法将传统旅行商问题的应用边界拓展至试验设计领域。这项研究为化学工程和生物医学研究等领域的实验序列优化提供了理论框架和跨学科解决方案。

关键词

试验设计, 顺序添加, 全循环设计, 旅行商问题

Research on Order-of-Addition in the Traveling Salesman Problem

Shiqi Zeng, Jiayan Xie*

College of Mathematics and Statistics, Jishou University, Jishou Hunan

Received: Mar. 16th, 2025; accepted: Apr. 9th, 2025; published: Apr. 18th, 2025

Abstract

To address the challenges of high path costs and excessive experimental iterations in complex multi-factor trial sequence optimization, this study proposes a dual-loop sorting algorithm based on an Adjacency Relationship Model. By constructing a full-cycle model matrix with baseline constraints and integrating a principal component partitioning strategy, the algorithm achieves rapid determination of the optimal trial sequence while minimizing experimental iterations. Theoretical

*通讯作者。

文章引用: 曾诗棋, 谢佳艳. 基于旅行商问题的顺序添加研究[J]. 应用数学进展, 2025, 14(4): 483-491.
DOI: 10.12677/aam.2025.144179

analysis demonstrates that the algorithm extends the application scope of the traditional Traveling Salesman Problem to experimental design domains. This research provides a robust theoretical framework and an interdisciplinary solution for experimental sequence optimization in fields such as chemical engineering and biomedical studies.

Keywords

Experimental Design, Order-of-Addition, Full-Cycle Design, Traveling Salesman Problem

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

旅行商问题在运筹学和计算数学领域占据着核心地位。随着计算技术的发展,特别是在顺序添加方面的研究引起了广泛关注这一研究领域不仅涉及经典的路径优化问题,还包括成本运算、基于特定标准(如距离)的顺序选择方法等多个方面。试验顺序添加的安排通常是随机进行的,这将导致不可控因素的产生。因为试验顺序添加的随机化可能会导致因子水平的大量变化或易受时间趋势的影响,从而使试验过程昂贵和耗时,因此如何合理安排试验顺序成为一个备受关注的问题。研究关于试验安排顺序的成果具有广泛的应用潜力,它可以在多个领域发挥重要作用,例如化学工程、航空航天等。这些成果不仅提高了这些领域的效率 and 创新能力,还促进了人类事业的整体繁荣与发展。

试验安排顺序的研究起源于 1951 年。文献[1]首次研究了试验安排顺序的改变对因子效应的估计产生显著的影响。文献[2]提出了试验实施中因子水平变化增加了试验成本的消耗。文献[3]研究了在一定时间顺序下进行测量时,为设计生成良好的试验安排顺序的方法。该方法给出的试验顺序比简单的随机顺序更有效,而且优于大多数特殊的经验规则。文献[4]构建了无时间趋势的两水平完全因子设计和部分因子设计,以便研究时间趋势对试验顺序的影响。通过折叠反转技术,文献[5]构造了无时间趋势且最小化成本的情况下,部分因子设计的试验安排顺序。文献[6]基于两水平饱和正规部分因子设计,构造了该类设计最优试验顺序的算法。

旅行商问题涉及寻找一系列位置的最短路径。这个问题的复杂性在于需要考虑所有可能的路径,并从中选择总长度最短的一条,所以关于成本运算的讨论显得尤为重要,通过采用适应性策略,如基于快速排序算法的方法,可以有效探索最优顺序添加。文献[7]提出的方法适用于处理具有大量因子的复杂旅行商问题,并能显著减少所需的运行次数和成本。此方法在不依赖于特定模型的情况下,通过比较因子间的相对顺序来确定最优解。文献[8]以及文献[9]的研究提供了进一步的见解,特别是在处理小规模问题时的效率和准确度。

在基于距离的顺序添加研究中,文献[10]研究强调了在确定最优路径时,考虑因子间的相对距离的重要性。这种方法在作业调度问题中尤为有效,它可以在减少计算时间的同时提高解决方案的准确性。文献[11]的研究也支持了这一观点,他们提出的方法在处理具有不同方差水平的问题时表现出色,进一步验证了基于距离的方法在解决复杂旅行商问题中的应用价值。除了传统的算法外,近年来在 TSP 领域中引入的新模型和算法,如高阶顺序添加模型、平衡不完全区块设计和阈值接受算法,为解决复杂的 TSP 问题提供了新的思路。文献[12]的研究对于理解更高阶的顺序添加模型起到了关键作用,而文献[9]通过平衡不完全区块设计提供了一种有效的方法来探索最优顺序。本文提出在已知因子的先验信息的情况下,

给出最优的试验顺序, 并且试验次数对比以往的数据都大幅度减少。

本文的余下结构安排如下。第 2 节将介绍相关的基础概念和使用的符号。第 3 节将展示主要的研究成果和结论。第 4 节将详细描述所采用算法的具体步骤。第 5 节通过一个实例来验证该算法的实际效果和应用价值。第 6 节是本文的总结。

2. 基本概念

假设有 $m (m \geq 3)$ 个因子, 可以以 $m!$ 种方式排序。任何这样的排序, 比如 $a = (a_1 a_2 \cdots a_m)$, 即 $123 \dots m$ 的一个排列, 都被视为一种处理。我们的模型包含了每对因子的顺序, 这与 Van Nostrand (1995) 和 Voelkel (2017) 的做法一致。

为了正式呈现这个模型, 我们用 Z 表示 $m!$ 种处理的集合。设 $a_{ij} = \phi_{(j)} - \phi_{(i)}$ 为 i 和 j 之间的有向距离。因此, 每个在 Z_m 上的排列都可以由给定的 i 确定的 $m-1$ 个有向距离 δ_{ij} 一一决定。例如, 给定顺序 $\{2, 1, 3, 4\}$ 。可以由 $\delta_{12} = -1$, $\delta_{13} = 1$, $\delta_{14} = 2$ 确定。考虑所有 $1 \leq i < j \leq m$ 的对 (i, j) , 我们得到一个新的成对排序距离(PWOD)数组。

依据我们提出的模型, 举例来说, 当设定参数 $m = 3$, 我们能够构建一个设计方案, 如表 1 所示, 该方案包含 $3!$ 次不同的试验安排。

Table 1. PWOD design

表 1. PWOD 设计

试验编号	试验顺序	δ_{12}	δ_{13}	δ_{23}
1	123	1	2	1
2	132	2	1	-1
3	213	-1	1	2
4	231	-2	-1	1
5	312	1	-1	-2
6	321	-1	-2	-1

现在提出一个邻接关系模型, 正如上文所讨论的, 每个城市在旅行商问题(TSP)中可以被看作是一个因子。注意, 这里的每个序列代表一个环路。因此, 无论起始城市是哪一个, 只有因子之间的相对顺序(AR)会影响响应。例如, 序列 13452 表示 1 与 2 和 3 相邻, 3 与 1 和 4 相邻, 4 与 3 和 5 相邻, 5 与 4 和 2 相邻, 以及 2 与 5 和 1 相邻。只要我们找到了一个最佳的环路, 不管它从哪里开始, 它仍然是最优的。假设总共有 m 个因子, $A = (a_{ij})$ 是因子矩阵, A 的每一行是 $\{1, 2, \dots, m\}$ 的一个排列, 代表一个旅行环路。我们为此类问题提出如下 AR 模型:

$$y = \beta_0 + \sum_{1 \leq i < j \leq m} \beta_{ij} I_{ij} + \varepsilon \quad (1)$$

如果 i 和 j 是相邻的, 则记为 $I_{ij} = |1|$, 当 i 在 j 在前面, 则记为 $I_{ij} = 1$, 反之 $I_{ij} = -1$; 如果 i 和 j 是不相邻的, 表示为 $I_{ij} = 0$, $\varepsilon \sim N(0, \sigma^2)$ 是随机误差。

定义 1 在一个旅行环路中, 每个因子只能与其他两个因子相邻, 所以我们有 m 个约束:

$$\sum_{j>i} I_{ij} + \sum_{j<i} I_{ji} = |2|, \text{ 对所有 } i = 1, 2, \dots, m. \quad (2)$$

我们可以添加基线约束:

$$\beta_{im} = 0, i = 1, 2, \dots, m-1, \beta_{(m-2)(m-1)} = 0. \quad (3)$$

在基线约束(2)下的 AR 模型可以表示为:

$$E_{2i-1}^*(D(X)) = \begin{cases} 0, & \text{for } i = 1, \dots, k \text{ and } i \neq \frac{k+1}{2}, \\ n, & \text{for } i = \frac{k+1}{2}. \end{cases} \quad (4)$$

对于一个 $n \times m$ 的因子矩 A , 由 $IA = (\mathbf{1}_n, \tilde{I}_{12}, \dots, \tilde{I}_{1(m-1)}, \dots, \tilde{I}_{(m-3)(m-1)})$ 表示相应的模型矩阵, 其中 $\mathbf{1}_n$ 是一个 $n \times 1$ 的全一向量, $\tilde{I}_{ij}$ 是对应于 I_{ij} 的 $n \times 1$ 向量。除非另有说明, 本文中提到的 AR 模型是指在基线约束(2)下的模型(3)。接下来, 将给出 OofA 全设计的模型矩阵, OofA 全设计, 记为 F_m , 包含了 m 个因子所有可能的 $m!$ 个排列。

引理 1 对于有 m 个组成部分的 AR 模型, 循环设计并不是唯一的, 但其对应的模型矩阵是同构的, 在本文中如果两个模型矩阵可以通过行的置换从一个得到另一个则称它们是同构的, 否则它们称为异构的。如果两个设计矩阵的模型矩阵是同构的, 那么这两个设计矩阵是同构的。

例如, 当 $m = 4$ 时, 两个循环设计是不同的, 如下:

$$\begin{pmatrix} 2 & 1 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 1 & 4 & 2 \end{pmatrix} \text{ 和 } \begin{pmatrix} 3 & 2 & 1 & 4 \\ 2 & 3 & 1 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$$

但它们的模型矩阵是同构的如下: 也就是说明其试验顺序是一致的。

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \text{ 和 } \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

3. 主要结论

本节将深入讨论针对旅行商问题的邻接关系模型, 特别是在面对任意数量因子的情况下。我们将重点分析该模型的性质和结构, 并探索确定最优试验安排顺序的方法和策略。

3.1. 基于 AR 模型下的全循环模型矩阵

定理 1 在 AR 模型下, 顺序添加完全设计 F_m 的模型矩阵 X_{OF} 具有以下性质:

- 1) 其大小为 $m! \times \frac{(m-1)(m-2)}{2}$ 。
- 2) 在 X_{OF} 的第 p 列 ($p = 2, 3, \dots, \frac{(m-1)(m-2)}{2}$), 包含 $m(m-2)!$ 个 1 和 -1, 以及 $m(m-3)(m-2)!$ 个 0。
- 3) 对于任意两列 $\tilde{I}_{ij}$ 和 $\tilde{I}_{ik}$ (或 $\tilde{I}_{jk}$), 其中 (i, j) 和 k 互不相同, (00) 出现 $m(m-4)(m-3)(m-3)!$ 次, (0,1)、(1,0)、(0,-1) 和 (-1,0) 分别出现 $m(m-3)(m-3)!$ 次, 而 (1,1) 和 (-1,-1) 出现 $m(m-3)!$ 次。
- 4) 对于任意两列 $\tilde{I}_{ij}$ 和 $\tilde{I}_{hk}$, 其中 i, j, h 和 k 互不相同, (00) 出现 $m(m^2 - 7m + 14)(m-3)!$ 次, (0,1)、(1,0)、(0,-1) 和 (-1,0) 分别出现 $m(m-4)(m-3)!$ 次, 而 (1,1) 和 (-1,-1) 出现 $2m(m-3)!$ 次。

证明: 1) 有 m 个因子, 则有 $m!$ 个排序, 故有 $m!$ 行, 根据 AR 模型及其基线约束, 其列的大小 n 为:

$$\begin{aligned} n &= \frac{(m-1)!}{2} - (m-2) - 1 + 1 \\ &= \frac{m^2 - m - 2m + 2}{2} \end{aligned}$$

$$= \frac{(m-1)(m-2)}{2}.$$

故其结构大小为: $m! \times \frac{(m-1)(m-2)}{2}$ 。

2) m 个因子在循环过程中, 有 m 个相邻的位置排列。当固定其相邻的两个因子的位置后, 剩余因子排列有 $(m-2)!$ 种情况, 故出现 $m(m-2)!$ 个 1 和 -1 列中 0 的次数为:

$$\begin{aligned} m! - 2m(m-2)! &= m(m-1)(m-2)! - 2m(m-2)! \\ &= (m^2 - m - 2m)(m-2)! \\ &= m(m-3)(m-2)! \end{aligned}$$

3) m 个因子在循环过程中, 取任意两列 $\tilde{I}_{ij}$ 和 $\tilde{I}_{ik}$, 其中下标 i 相同, 故假设固定下标 A , 其左右不是 j 和 k , 故因子有 m 个选择, j 和 k 的位置有 $(m-3)(m-4)$ 个情况。剩下的 $m-3$ 因子选择有 $(m-3)!$ 个情况。所以 $(0,0)$ 出现 $m(m-4)(m-3)(m-3)!$ 次。同理, 固定因子 A , 其左右是 j 和 k 中一个, 故因子有 $2m$ 个选择, 并且还要区分其在因子 A 的前后顺序, 故有 m 个选择, 其中 j 和 k 不在 A 左右位置的有 $m-3$ 个选择, 剩下的 $m-3$ 因子选择有 $(m-3)!$ 个情况。所以 $(0,1)$ $(1,0)$ $(0,-1)$ $(-1,0)$ 分别出现 $m(m-3)(m-3)!$ 次。同理, 固定因子 A , j 和 k 都在旁边且区分前后顺序, 故因子有 m 个选择, 剩下的 $m-3$ 因子选择有 $(m-3)!$ 个情况。所以 $(1,1)$ 和 $(-1,-1)$ 出现 $m(m-3)!$ 次。

4) m 个因子在循环过程中, 对于任意两列 $\tilde{I}_{ij}$ 和 $\tilde{I}_{hk}$, 其中 i, j, h 和 k 互不相同, (11) 、 $(-1,-1)$ 说明其都是相邻的位置关系, 故直接固定 4 个因子, 分别形成两两在一起的结构有 $2m$ 次, 且其他因子在循环结构中插空选择有 $(m-3)!$ 种情况。所以 $(1,1)$ 、 $(-1,-1)$ 各出现 $2m(m-3)!$ 次。

$(0,1)$ 、 $(1,0)$ 、 $(0,-1)$ 和 $(-1,0)$ 则表示只有两个因子相邻, 而另外两个因子不相邻。先考虑 2 个因子相邻的情况, 所以有 m 种选择, 然后是两个不相邻的因子, 先排除两个固定因子的情况, 再考虑不相邻的情况, 则有 $(m-3)(m-4)$ 种情况, 接下来就是剩余 $m-4$ 个因子的情况, 共有 $(m-4)!$ 种情况。所以 $(0,1)$ 、 $(1,0)$ 、 $(0,-1)$ 和 $(-1,0)$ 分别出现 $m(m-4)(m-3)!$ 次。

最后关于 $(0,0)$ 出现的次数等于:

$$\begin{aligned} m! - 4 \cdot m(m-4)(m-3)! - 2 \cdot 2m(m-3)! \\ &= [m(m-1)(m-2) - 4m(m-4) - 4m](m-3)! \\ &= [m(m^2 - 3m + 2 - 4m + 16 - 4)](m-3)! \\ &= m(m^2 - 7m + 14)(m-3)! \end{aligned}$$

证毕。

3.2. 最优试验顺序的试验次数

本节的核心在于利用地理位置信息的局部性, 递归地优化每个子集的顺序, 最终合并得到全局循环最优顺序, 考虑全循环顺序, 其试验次数为 $\frac{(m-1)!}{2}$ 。这种划分方式通过分治策略和空间连续性, 将复杂问题分解为较小的子问题, 从而减少试验次数并保证最优顺序。

定理 2 若因子数为偶数 $2m$, 根据经纬度地理位置信息, 确定一个主城市, 则左子集、右子集的因子数量按照其地理位置信息划分为 m 、 $m-1$, 则只需要进行 $\frac{(m-2)!m}{2}$ 次试验即可确定最优试验顺序。

证明: 定一个主城市因子后, 剩余因子数为 $2m-1$, 依据从小到大排序后, 左子集、右子集的因子数

量划分为 m 、 $m-1$ 或者 $m-1$ 、 m , 则试验次数 T 为:

$$\begin{aligned} T &= \frac{(m-1)!}{2} + \frac{(m-2)!}{2} \\ &= \frac{(m-2)!}{2} (m-1+1) = \frac{(m-2)!m}{2} \end{aligned}$$

证毕。

定理 3 若因子数为奇数 $2m+1$, 根据经纬度地理位置信息, 确定一个主城市, 则左子集、右子集的城市数量按照其地理位置信息划分为 m 、 m , 则只需要进行 $(m-1)!$ 次试验即可确定最优试验顺序。

证明: 定一个主城市因子后, 剩余因子数为 $2m$, 依据从小到大排序后, 左子集、右子集的因子数量划分为 m 、 m , 则试验次数 T 为:

$$T = \frac{(m-1)!}{2} + \frac{(m-1)!}{2} = (m-1)!$$

证毕。

这种划分方式能够保留空间分布的连续性, 避免将距离较远的城市划分到同一个子集中, 从而减少试验的复杂性, 且在根据地理位置信息时已经确保空间分布均匀选择中间因子作为主城市, 因此在此方式下可得到最优的试验顺序。

4. 最优试验顺序排序的算法

在经典的快速排序方法是通过从所有原始元素中选择一个枢轴元素, 然后将其他元素分为两个子集(一个包含比枢轴元素大的, 另一个包含比枢轴元素小的), 接着递归地对这两个子集进行排序。受到快速排序基本思想的启发, 提出的方法将所有因子分为两个子集: 左子集包含在枢轴因子之前的因子, 右子集包含在枢轴因子之后的因子。基于邻接关系对这两个子集按照循环的方法计算, 可以通过结合所有最优子集来获得最优顺序。该方法的过程在算法 1 中提出, 并在第 5 节中提供了一个示例。需要考虑的最优响应通常有三种类型: (a) “目标是最好的”; (b) “越大越好”; 和(c) “越小越好”。为了简单起见, 我们只关注情况(c); 其他两种情况可以以类似的方式进行。

在输入 m 个因子的经纬坐标后, 根据极差以及空间性, 所有因子可以被分为三个子集: (a) 主因子 a_1 , (b) 所有在主因子 a_1 之前的因子 $a_t (t=2, \dots, m)$ (将放在左子集中), (c) 所有被认为在主因子 a_1 之后的因子 a_t (将放在右子集中)。因此, 我们得到了一个原始的分区顺序, 命名为 O_1^* , 如下所示:

$$O_1^* = \{a_t : \text{因子 } a_t \text{ 在 } a_1 \text{ 之前}\}, a_1, \{a_t : \text{因子 } a_t \text{ 在 } a_1 \text{ 之后}\}. \quad (5)$$

y 是排列顺序 $(a_1 a_2 a_3 \dots a_m)$ 的函数, 实验者的目标是最小化响应。让 $a_j (j=1, \dots, m)$ 是第 j 个位置的因子, 记 a_1 是主因子。假设右子集初始运行以基线顺序 $O_1 = (a_1 a_2 a_3 \dots a_m)$ 开始记作较 y_1 。交换 a_m 和 a_{m-1} 来比较 y_1 和 y_2 。以这种方式, 进行相应次数的运行。通过全循环邻接关系对左右两个子集应用上述步骤, 可以获得最优顺序。所提出的程序从一个初始运行开始, 然后根据分区顺序进行一系列连续的运行, 直到最终得到最优顺序。详细信息在算法 1 中给出。

算法 1: 具有最优顺序的算法

输入: 城市的经纬坐标点

输出: 最优的试验安排顺序

第一步: 通过经纬坐标找出主城市

第二步: 通过定位主城市, 划分左子集、右子集

第三步: 将左子集、右子集的城市分别按照全循环邻接模型的结构找出最优的顺序

第四步: 将主城市分别放入左子集、右子集, 分别寻找最优的顺序

第五步: 固定主城市在左子集的末尾, 将主城市固定在右子集的开端啊, 将左子集、主城市、右子集的顺序连接确定, 获得最优顺序。

5. 数值例子

本节通过一个实际的数值示例展示了如何运用算法 1 来确定在 AR 模型中处理旅行商问题时的最优全循环试验序列, 个过程旨在展示该算法如何有效地找到问题的最佳解决方案。

例 1: 假设有 10 个城市 C1 (北京)、C2 (上海)、C3 (西安)、C4 (成都)、C5 (武汉)、C6 (香格里拉)、C7 (青岛)、C8 (南昌)、C9 (三亚)和 C10 (贵阳)需要访问, 它们的经纬坐标如表 2 所示。为了后续方便, 后续城市用字母表示。任意两个城市之间由经纬坐标可得出距离, 为了方便计算保留两位小数, 详情如表 3 所示。

Table 2. Longitude and latitude coordinates of China's top 10 tourist cities

表 2. 中国十大旅游城市的经纬坐标

城市	纬度	经度	城市	纬度	经度
三亚	18.2528	109.5119	武汉	30.5931	114.3054
贵阳	26.6477	106.6302	上海	31.2304	121.4737
香格里拉	27.8251	99.7080	西安	34.3416	108.9402
南昌	28.6829	115.8583	青岛	36.0671	120.3826
成都	30.5728	104.0668	北京	39.9042	116.4074

Table 3. Table of distances between two cities

表 3. 两两城市间距离表

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0.00	1067.31	905.39	1524.52	1052.48	2038.61	550.70	1248.76	2496.94	1729.13
C2	1067.31	0.00	1220.69	1660.70	687.37	2136.54	547.20	610.54	1878.95	1530.09
C3	905.39	1220.69	0.00	620.08	620.08	1138.52	1056.52	908.35	1789.90	883.60
C4	1524.52	1660.70	620.08	0.00	979.76	522.05	1631.79	1158.34	1476.05	503.05
C5	1052.48	687.37	620.08	979.76	0.00	1449.23	829.67	260.24	1455.17	867.92
C6	2038.61	2136.54	1138.52	522.05	1449.23	0.00	2149.62	1584.17	1462.14	697.46
C7	550.70	547.20	1056.52	1631.79	829.67	2149.62	0.00	924.19	2250.11	1670.89
C8	1248.76	610.54	908.35	1158.34	260.24	1584.17	924.19	0.00	1327.48	936.26
C9	2496.94	1878.95	1789.90	1476.05	1455.17	1462.14	2250.11	1327.48	0.00	979.20
C10	1729.13	1530.09	883.60	503.05	867.92	697.46	1670.89	936.26	979.20	0.00

通过已知其经纬度坐标, 可以得出两两城市之间的距离。随后通过算法 1, 获取其经纬度坐标的极差, 进而去确定主城市的选取。如果经度极差与纬度极差的比值接近 1, 说明东西方向和南北方向的空间范围相近, 研究区域在空间上接近一个正方形或圆形, 而不是狭长的长方形或不规则形状。这种对称性

表明空间分布较为均匀, 所以当经度的极差与纬度的极差相除, 结果处于 0.9~1.1 之间时, 则说明空间性比较均匀。然后将经纬度分别从小到大排列, 在经度、纬度两个因素中分别选择居于中间的两个城市(例题为 10 个城市则选择第五个, 第六个城市, 并可以避开极端值, 受异常值的影响较小, 确保选择的 城市更能代表整体的空间分布特征)。若有城市在经纬度中都被选中, 说明其均居于中间位置, 则直接将其认定为主城市, 反之, 没有共同的城市, 则将主城市数量为一个改为考虑多个的主城市。随后再计算经度、纬度的极差相对较大的作为判定依据划分左子集、左子集。若经度的极差与纬度的极差相除的结果大于 1.1 或者小于 0.9, 则说明区域可能呈东西狭长的形状或者区域可能呈南北狭长的形状, 进而直接选取极差大的因素作为划分左子集、右子集的依据。选取划分依据之后, 根据划分依据的因素, 主城市分别与其他城市对比起因素的具体大小值。比主城市小的进入到左子集中, 反之进入右子集。

例 1 经过算法 1 的计算, 选取的主城市为: C5; 根据经度继续进行划分, 划分的左子集为: {C3, C4, C6, C9, C10}; 右子集为: {C1, C2, C7, C8}。随后, 在邻接关系模型的基础上, 计算左子集的城市 的路线, 将 C3, C4, C6, C9, C10 在模型中记为 3, 4, 6, 9, 10, 其路线及其响应值见表 4。

Table 4. Left subset trial order
表 4. 左子集试验顺序

试验次数	路线	响应值	试验次数	路线	响应值
1	3-4-6-9-10	3583.47	7	3-6-4-9-10	4115.82
2	3-4-6-10-9	2818.78	8	3-6-4-10-9	3142.82
3	3-4-9-6-10	4255.72	9	3-6-9-4-10	4579.76
4	3-4-9-10-6	3772.78	10	3-6-10-4-9	3815.07
5	3-4-10-6-9	3282.72	11	3-9-4-6-10	4485.44
6	3-4-10-9-6	3564.46	12	3-9-6-4-10	4277.13

由表 4 可知 3-4-6-10-9 的距离是最短的, 进而考虑将 C5 加入此循环序列, 故有如下 6 种情况: 5-3-4-6-10-9, 3-5-4-6-10-9, 3-4-5-6-10-9, 3-4-6-5-10-9, 3-4-6-10-5-9, 3-4-6-10-9-5 可轻松得到 5-3-4-6-10-9 和 3-4-6-10-9-5 的路线都可到达一种最优路线, 并且要考虑主城市位于左子集后面, 则最终左子集最优路线为: 9-10-6-4-3-5, 其响应值为 3471.96。

随后将右子集的城市计算其路线。将 C1, C2, C7, C8 在模型中记为 1, 2, 7, 8, 则路线及其响应值见表 5。

Table 5. Right subset trial order
表 5. 右子集试验顺序

试验次数	路线	响应值
1	1-2-7-8	2538.71
2	1-2-8-7	2602.05
3	1-7-2-8	1708.44

由表可知 1-7-2-8 的距离是最短的, 进而考虑将 C5 加入此循环序列, 故有如下 5 种情况: 5-1-7-2-8, 1-5-7-2-8, 1-7-5-2-8, 1-7-2-5-8, 1-7-2-5-8 可轻松得到 5-1-7-2-8 和 1-7-2-5-8 的路线都可到达一种最优路线: 5-8-2-7-1, 其响应值为 1968.69。现在将其最优路线进行组合: 9-10-6-4-3-5-8-2-7-1。

6. 总结

试验顺序添加在各个领域中扮演着至关重要的角色, 无论是在精密的化学实验, 还是在关键的生物医学研究中都是如此。本文深入探讨了基于旅行商问题的路线最优化, 在试验设计的顺序添加方面, 本研究着眼于试验安排顺序的优化。特别地, 本文针对复杂的多因子顺序添加问题, 提出了一种既简洁又高效的算法, 以确定最优的试验安排顺序。这些理论成果和算法的构建, 不仅为实际工作中的试验者提供了坚实的依据和宝贵的参考, 也为试验的顺利实施带来了便利。特别当研究其他因子特性时如温度、浓度等, 可将设置成相应范围构建一个封闭的区间, 再根据算法 1 来进行最优顺序安排通过这种方法, 试验的成本得以降低, 从而极大地提升了试验的效率和效用。

参考文献

- [1] Cox, D.R. (1951) Some Systematic Experimental Designs. *Biometrika*, **38**, 312-323. <https://doi.org/10.1093/biomet/38.3-4.312>
- [2] Draper, N.R. and Stoneman, D.M. (1968) Factor Changes and Linear Trends in Eight-Run Two-Level Factorial Designs. *Technometrics*, **10**, 301-311. <https://doi.org/10.1080/00401706.1968.10490562>
- [3] Joiner, B.L. and Campbell, C. (1976) Designing Experiments When Run Order Is Important. *Technometrics*, **18**, 249-259. <https://doi.org/10.1080/00401706.1976.10489445>
- [4] Cheng, C. and Jacroux, M. (1988) The Construction of Trend-Free Run Orders of Two-Level Factorial Designs. *Journal of the American Statistical Association*, **83**, 1152-1158. <https://doi.org/10.1080/01621459.1988.10478713>
- [5] Coster, D.C. and Cheng, C. (1988) Minimum Cost Trend-Free Run Orders of Fractional Factorial Designs. *The Annals of Statistics*, **16**, 1188-1205. <https://doi.org/10.1214/aos/1176350955>
- [6] Cheng, C., Martin, R.J. and Tang, B. (1998) Two-Level Factorial Designs with Extreme Numbers of Level Changes. *The Annals of Statistics*, **26**, 1522-1539. <https://doi.org/10.1214/aos/1024691252>
- [7] Lin, D.K.J. and Peng, J. (2019) Order-of-Addition Experiments: A Review and Some New Thoughts. *Quality Engineering*, **31**, 49-59. <https://doi.org/10.1080/08982112.2018.1548021>
- [8] Peng, J., Mukerjee, R. and Lin, D.K.J. (2019) Design of Order-of-Addition Experiments. *Biometrika*, **106**, 683-694. <https://doi.org/10.1093/biomet/asz025>
- [9] Chen, J., Mukerjee, R. and Lin, D.K.J. (2020) Construction of Optimal Fractional Order-of-Addition Designs via Block Designs. *Statistics & Probability Letters*, **161**, Article 108728. <https://doi.org/10.1016/j.spl.2020.108728>
- [10] Winker, P., Chen, J. and Lin, D.K.J. (2020) The Construction of Optimal Design for Order-of-Addition Experiment via Threshold Accepting. In: *Contemporary Experimental Design, Multivariate Analysis and Data Mining*, Springer, 93-109. https://doi.org/10.1007/978-3-030-46161-4_6
- [11] Zhao, Y., Lin, D.K.J. and Liu, M. (2020) Designs for Order-of-Addition Experiments. *Journal of Applied Statistics*, **48**, 1475-1495. <https://doi.org/10.1080/02664763.2020.1801607>
- [12] Mee, R.W. (2020) Order-of-addition Modeling. *Statistica Sinica*, **30**, 1543-1559. <https://doi.org/10.5705/ss.202018.0210>