

运筹学业界应用：选址问题

2023年11月8日

黄一潇，顺丰速运

大纲

- 选址模型
 - 经典选址模型
 - P-中值模型案例
- 选址问题实践
 - 建模的挑战
 - 应用实践的挑战
- 总结

关于顺丰

顺丰与佐治亚理工学院签署校企合作协议，加速产学研合作进程

发布时间：2017-11-14 09:44:47 | 来源：中国网 | 作者： | 责任编辑：科学频道



深圳大学 土木与交通工程学院

College of Civil and Transportation Engineering, Shenzhen University

我院与顺丰集团签署战略合作框架协议

编辑：土木高级账号 发布时间：2022-12-08 08:16 浏览次数：1695



关于本人

- 教育经历

- 2018年清华大学管理科学与工程博士毕业，研究领域：城市物流
- 2014年1月于荷兰埃因霍温理工大学访学
- 2016年9月至2017年3月于美国佐治亚理工学院工业与系统工程系访学
- 博士期间于Transportation Science发表论文1篇，Transportation Research Part B: Methodological发表论文2篇

- 工作经历

- 2018年至今，顺丰工作，历任运筹优化算法工程师、运筹优化算法高级工程师、运筹优化算法资深工程师，现任运营模式算法组负责人
- 2020年至今，担任东北财经大学管理科学与工程学院硕士兼职导师
- 参与项目包括：城市内物流网络规划模型、城市间物流网络规划模型、航空网络规划模型、路由规划模型、时效评估模型
- 工作期间于期刊Transportation Science、Omega、Transport Policy各发表论文1篇

选址模型

经典选址模型

- Set covering (集覆盖)
 - 最小化选址数, 全覆盖
- Max covering (最大覆盖)
 - 固定选址数量, 最大化覆盖点数
- P-center (P-中心)
 - 固定选址数量, 最小化最大距离
- P-median (P-中值)
 - 固定选址数量, 最小化平均距离

Source: Daskin, M. S. (2008). *What you should know about location modeling*. *Naval Research Logistics (NRL)*, 55(4), 283-294.

经典选址模型

- P-中值模型
 - 选P个点，最小化货量 * 距离的加权和
- 数据
 - N : 需求点集合
 - M : 候选点集合
 - d_{ij} : 需求点到候选点的距离
 - q_i : 需求点的货量
 - P : 候选点选址数量
- 决策
 - $y_j \in \{0, 1\}$: 候选点是否选择
 - $x_{ij} \in \{0, 1\}$: 需求点 $i \in N$ 是否被候选点 $j \in M$ 覆盖

经典选址模型

- P-中值模型

- 选P个点，最小化货量 * 距离的加权和

- 数据

- N : 需求点集合
- M : 候选点集合
- d_{ij} : 需求点到候选点的距离
- q_i : 需求点的货量
- P : 候选点选址数量

$$\min \sum_{i \in N} \sum_{j \in M} q_i d_{ij} x_{ij}$$

$$\text{s.t.} \quad \sum_{j \in M} x_{ij} = 1, \quad \forall i \in N,$$

$$\sum_{j \in M} y_j = P,$$

$$x_{ij} \leq y_j, \quad \forall i \in N, j \in M,$$

$$x_{ij} \in \{0, 1\}, \quad \forall i \in N, j \in M,$$

$$y_j \in \{0, 1\}, \quad \forall j \in M.$$

- 决策

- $y_j \in \{0, 1\}$: 候选点是否选择
- $x_{ij} \in \{0, 1\}$: 需求点 $i \in N$ 是否被候选点 $j \in M$ 覆盖

经典选址模型（和一些小改动）

- P-中值模型 + 容量约束
- 对结果有什么样的影响？

$$\min \sum_{i \in N} \sum_{j \in M} q_i d_{ij} x_{ij}$$

$$\text{s.t.} \quad \sum_{j \in M} x_{ij} = 1, \quad \forall i \in N,$$

$$\sum_{j \in M} y_j = P,$$

$$x_{ij} \leq y_j, \quad \forall i \in N, j \in M,$$

$$\sum_{i \in N} q_i x_{ij} \leq \overline{Q} y_j, \quad \forall j \in M,$$

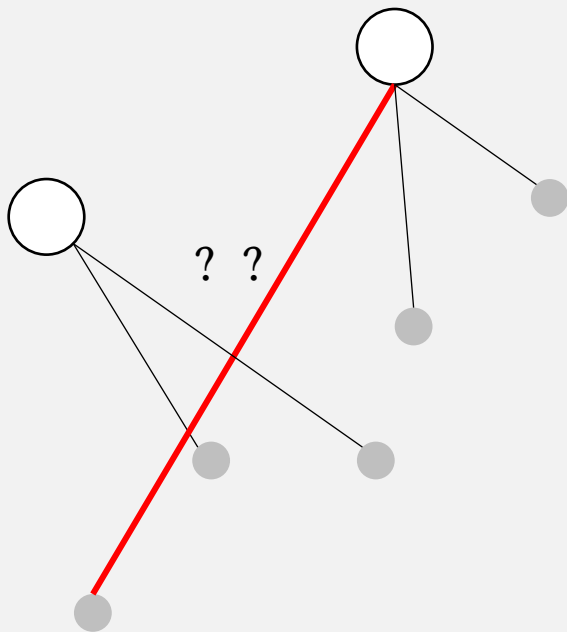
$$\sum_{i \in N} q_i x_{ij} \geq \underline{Q} y_j, \quad \forall j \in M,$$

$$x_{ij} \in \{0, 1\}, \quad \forall i \in N, j \in M,$$

$$y_j \in \{0, 1\}, \quad \forall j \in M.$$

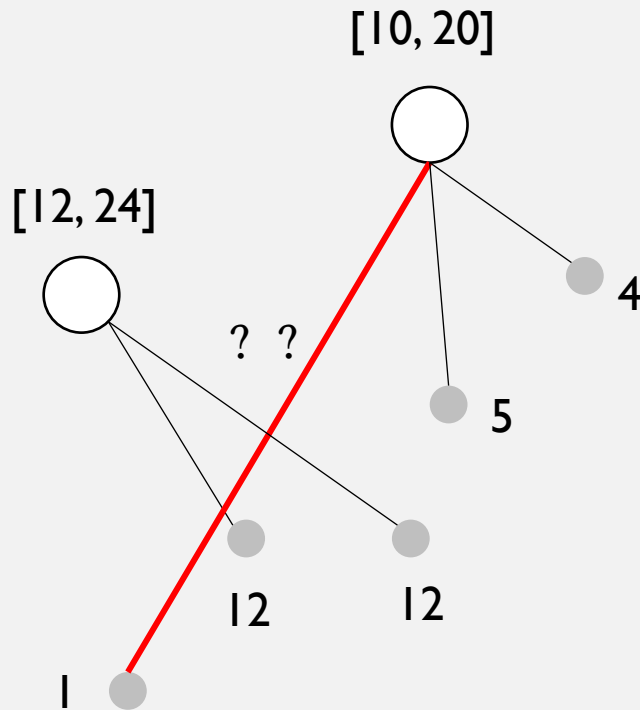
经典选址模型（和一些小改动）

- P-中值模型 + 容量约束
 - 对结果有什么样的影响？
 - 示例



经典选址模型（和一些小改动）

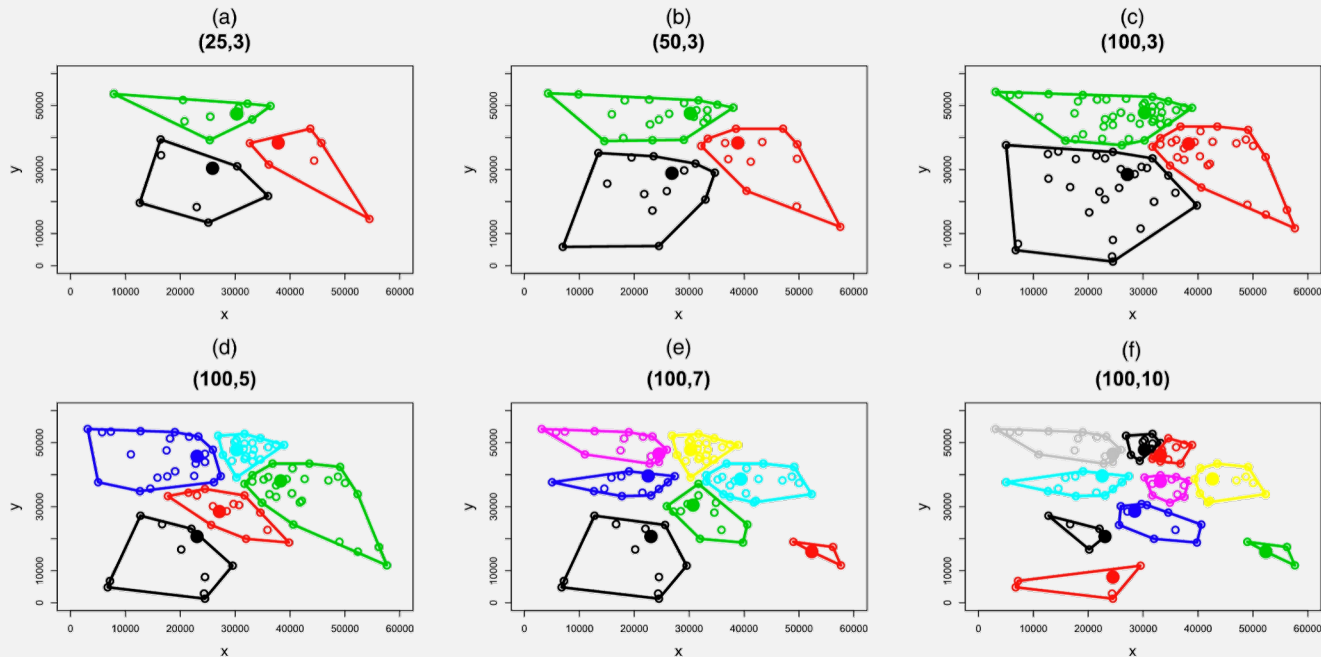
- P-中值模型 + 容量约束
 - 对结果有什么样的影响？
 - 示例



- 参数设置不合理？
- 加就近覆盖的约束？
- 上下限约束变成软约束？
- ...

经典选址模型（和一些小改动）

- P-中值模型 + 容量约束
- 覆盖关系“规整”问题：Spatial separability (Khaniyev et al., 2020)



Source: Khaniyev, T., Elhedhli, S., & Erenay, F. S. (2020). Spatial separability in hub location problems with an application to brain connectivity networks. *INFORMS Journal on Optimization*, 2(4), 320-346.

经典选址模型（和一些小改动）

- P-中值模型 + 容量约束 + **覆盖距离限制**

- 定义集合

$$A = \{(i, j) \mid i \in N, j \in M, d_{ij} \leq \tau\},$$

$$N_j = \{i \mid i \in N, (i, j) \in A\},$$

$$M_i = \{j \mid j \in M, (i, j) \in A\}.$$

$$\begin{aligned} \min \quad & \sum_{(i,j) \in A} q_i d_{ij} x_{ij} \\ \text{s.t.} \quad & \sum_{j \in M_i} x_{ij} = 1, & \forall i \in N, \\ & \sum_{j \in M} y_j = P, \end{aligned}$$

- 这样做有啥好处?

$$\begin{aligned} & x_{ij} \leq y_j, & \forall (i, j) \in A, \\ & \sum_{i \in N_j} q_i x_{ij} \leq \bar{Q} y_j, & \forall j \in M, \end{aligned}$$

- 这样做对结果有啥影响?

$$\begin{aligned} & \sum_{i \in N_j} q_i x_{ij} \geq \underline{Q} y_j, & \forall j \in M, \\ & x_{ij} \in \{0, 1\}, & \forall (i, j) \in A, \\ & y_j \in \{0, 1\}, & \forall j \in M. \end{aligned}$$

经典选址模型（和一些小改动）

- P-中值模型 + 容量约束 + 覆盖距离限制
 - 对结果有什么样的影响？
 - 可能无解
 - 某个需求点在距离范围内没有候选点可以覆盖
 - 候选点覆盖范围内的所有需求点违反容量约束
 - 怎么处理？
 - 返回“不可行/无解”
 - 无覆盖的需求点删除？
 - 无覆盖的需求点强制归集到最近的点（即使超出了距离限制）？

经典选址模型（和一些小改动）

- 小结
 - George E. P. Box: All models are wrong, but some are useful.
 - 数学模型并不是越细致越好
 - 越细致的数学模型对数据质量的要求越高
 - 越细致的数学模型的求解结果越不稳健
 - 越细致的数学模型越难以排查错误
 - 越细致的数学模型越难以分析总结规律
- 如何合理建模，是一门艺术

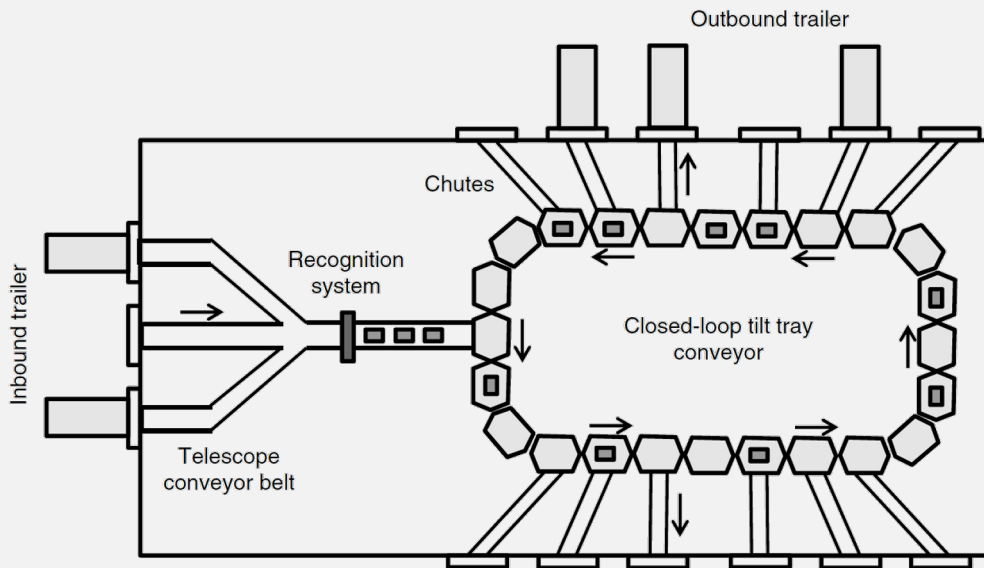
选址问题实践

选址问题应用

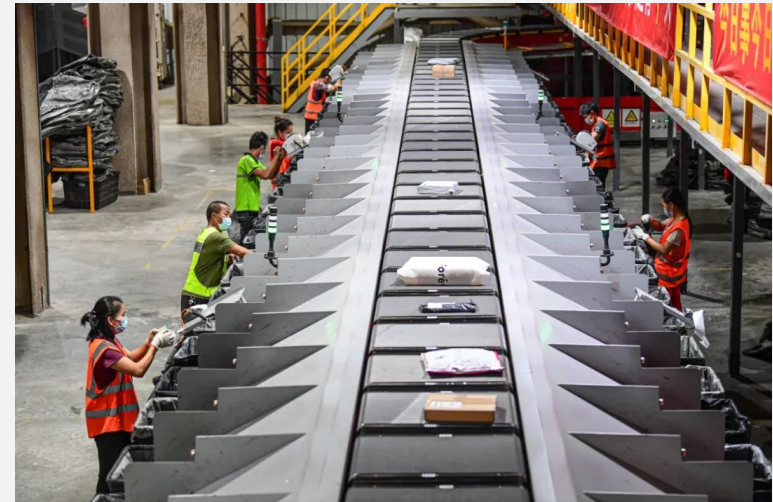
- 应用场景
 - 便利店选址
 - 房屋中介门店选址
 - 超市选址
 - 中转场选址
 - 快递网点选址
 - 银行网点选址
 - 核酸点选址
 -
- 决策维度
 - 战略规划
 - 年度预算
 - 考核监控
 - 落地执行

中转场选址场景

- 什么是中转场？
 - 不恰当的比喻：把乱序的扑克，按照花色、数字分堆
 - 占地面积较大，一般处于郊区
- 中转场选址，可能类似一个P-中值模型就可以，会有哪些挑战？



Source: Fedtke and Boysen (2014)



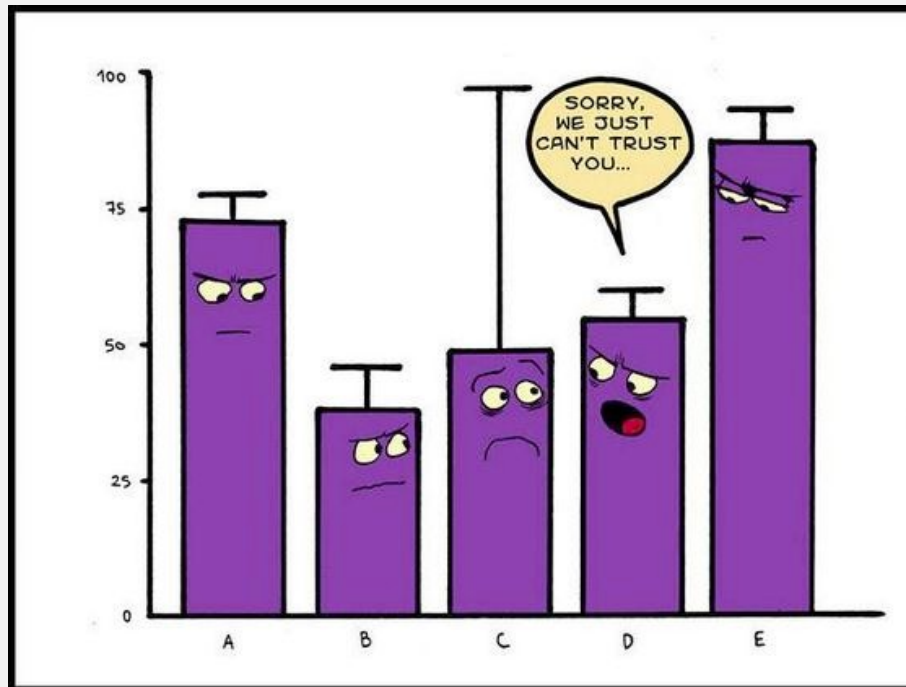
数学模型四要素

- 数据：我们知道什么
- 决策：我们可以改变什么
- 目标：我们怎么知道做得好不好
- 约束：我们受到哪些限制

$$\begin{aligned} \min \quad & \sum_{(i,j) \in A} q_i d_{ij} x_{ij} \\ \text{s.t.} \quad & \sum_{j \in M_i} x_{ij} = 1, & \forall i \in N, \\ & \sum_{j \in M} y_j = P, \\ & x_{ij} \leq y_j, & \forall (i,j) \in A, \\ & \sum_{i \in N_j} q_i x_{ij} \leq \overline{Q} y_j, & \forall j \in M, \\ & \sum_{i \in N_j} q_i x_{ij} \geq \underline{Q} y_j, & \forall j \in M, \\ & x_{ij} \in \{0, 1\}, & \forall (i,j) \in A, \\ & y_j \in \{0, 1\}, & \forall j \in M. \end{aligned}$$

挑战一：数据

- 数据永远是个问题。信息化程度高 \neq 优化模型输入质量高。
 - 信息化程度高：现状方案的数据完备
 - 优化模型输入质量高：所有潜在方案的数据完备



挑战一：数据

- 数据：我们知道什么
 - 候选场地集合
 - 网点集合
 - 货量数据
 - 导航距离
 - 场地数量
 - 场地处理能力上下限

挑战一：数据

- 数据：我们知道什么
 - 候选场地集合（候选点在哪里？哪些地方有物流园区、物流用地可用？）
 - 网点集合
 - 货量数据（未来几年的预测数据，存在偏差）
 - 导航距离
 - 场地数量
 - 场地处理能力上下限（和候选点在哪里有关系）

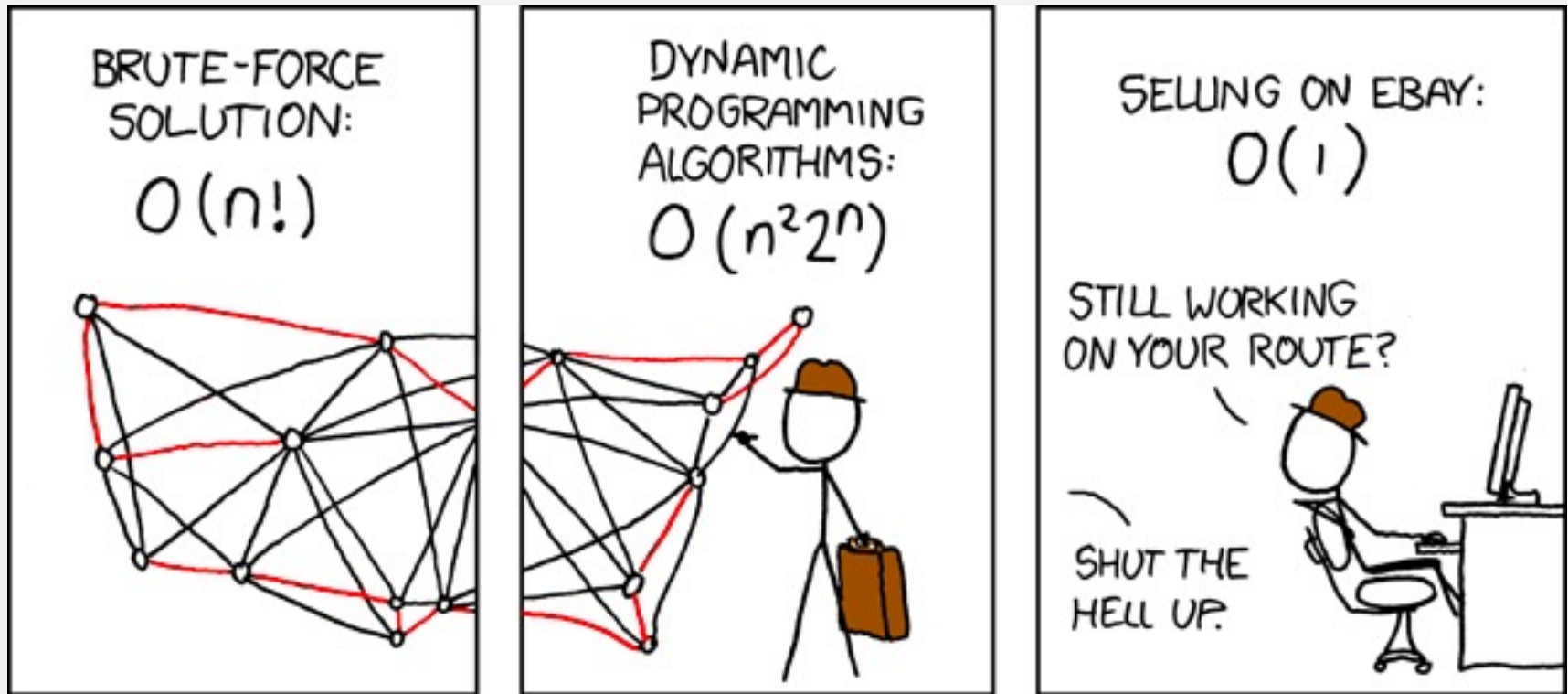
挑战一：数据

- 数据：我们知道什么
 - 候选场地集合（候选点在哪里？哪些地方有物流园区、物流用地可用？）
 - 网点集合
 - 货量数据（未来几年的预测数据，存在偏差）
 - 导航距离（是否满足三角不等式，数据来源）
 - 场地数量
 - 场地处理能力上下限（和候选点在哪里有关系）

| 始发 | 目的 | 耗时 |
|----|----|----|
| 南昌 | 拉萨 | 12 |
| 成都 | 拉萨 | 24 |

挑战二：决策

- 优化问题范围界定变化，导致决策变化。



挑战二：决策

- 决策：我们可以改变什么
 - 哪些候选点被选中了
 - 每个需求点被谁覆盖

挑战二：决策

- 决策：我们可以改变什么
 - 哪些候选点被选中了
 - 每个需求点被谁覆盖
 - 中转场要多大面积，要多少人（分工种），要多少设备（分功能）
 - 中转场的干线、支线怎么设计（问题边界不断外拓）

挑战三：目标

- 涉及成本的目标函数，如何近似核算。
- 不追求成本最优，但是隐含成本的约束。



挑战三：目标

- 目标：我们怎么知道做得好不好
 - 最小化运输成本（近似的）

挑战三：目标

- 目标：我们怎么知道做得好不好
 - 最小化运输成本
 - 不同场地位置、不同面积的成本不一样，要不要考虑
 - 人员成本、设备成本、干支线成本，要不要考虑
- 多个目标，如何权衡
- 如果存在软约束，约束违反的惩罚值如何加入目标函数

挑战四：约束

- 层出不穷的约束条件。

$$\begin{bmatrix} \cos 90^\circ & \sin 90^\circ \\ -\sin 90^\circ & \cos 90^\circ \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

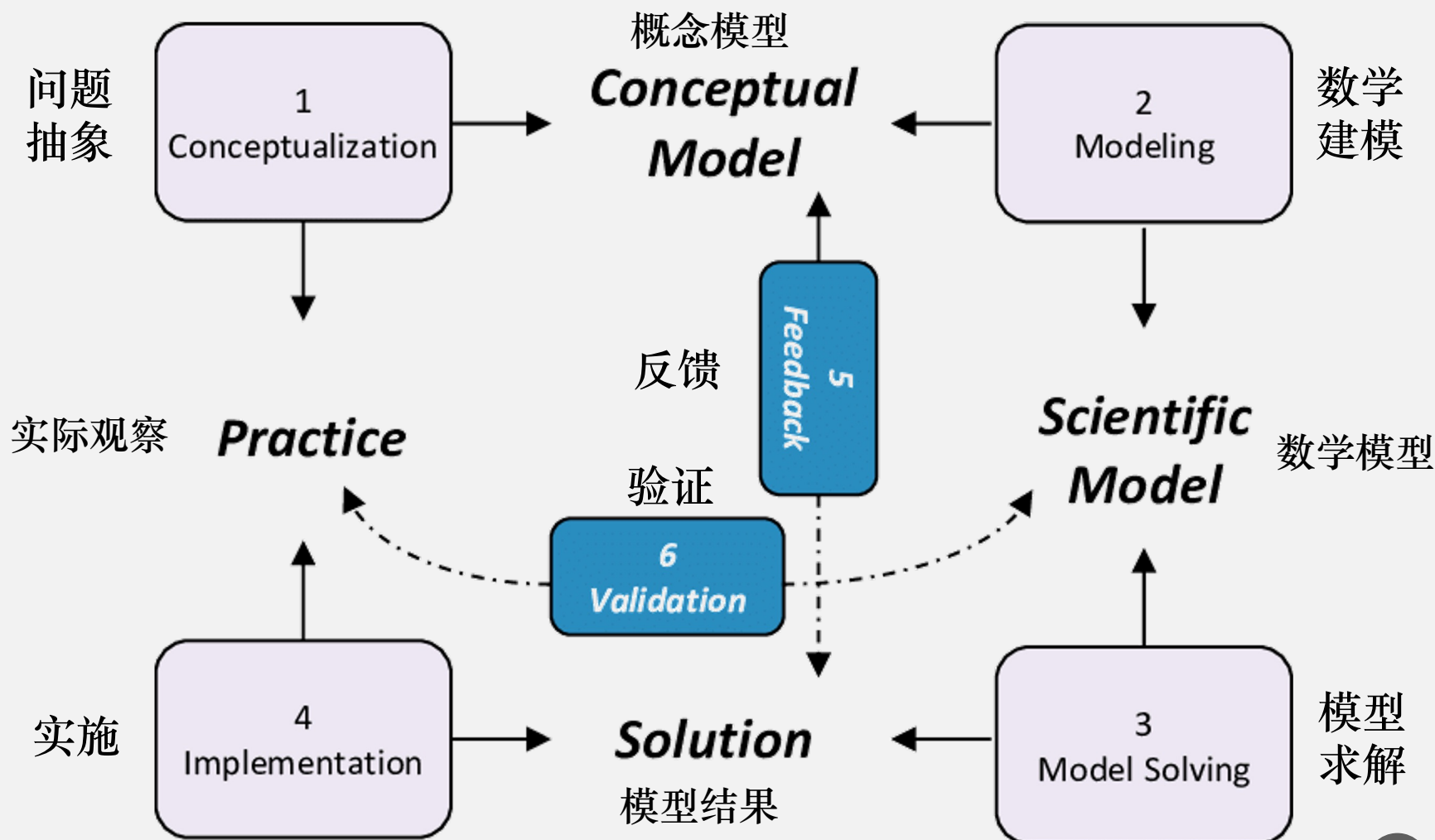
挑战四：约束

- 约束：我们受到哪些限制
 - 每个需求点必须被覆盖
 - 容量上下限约束

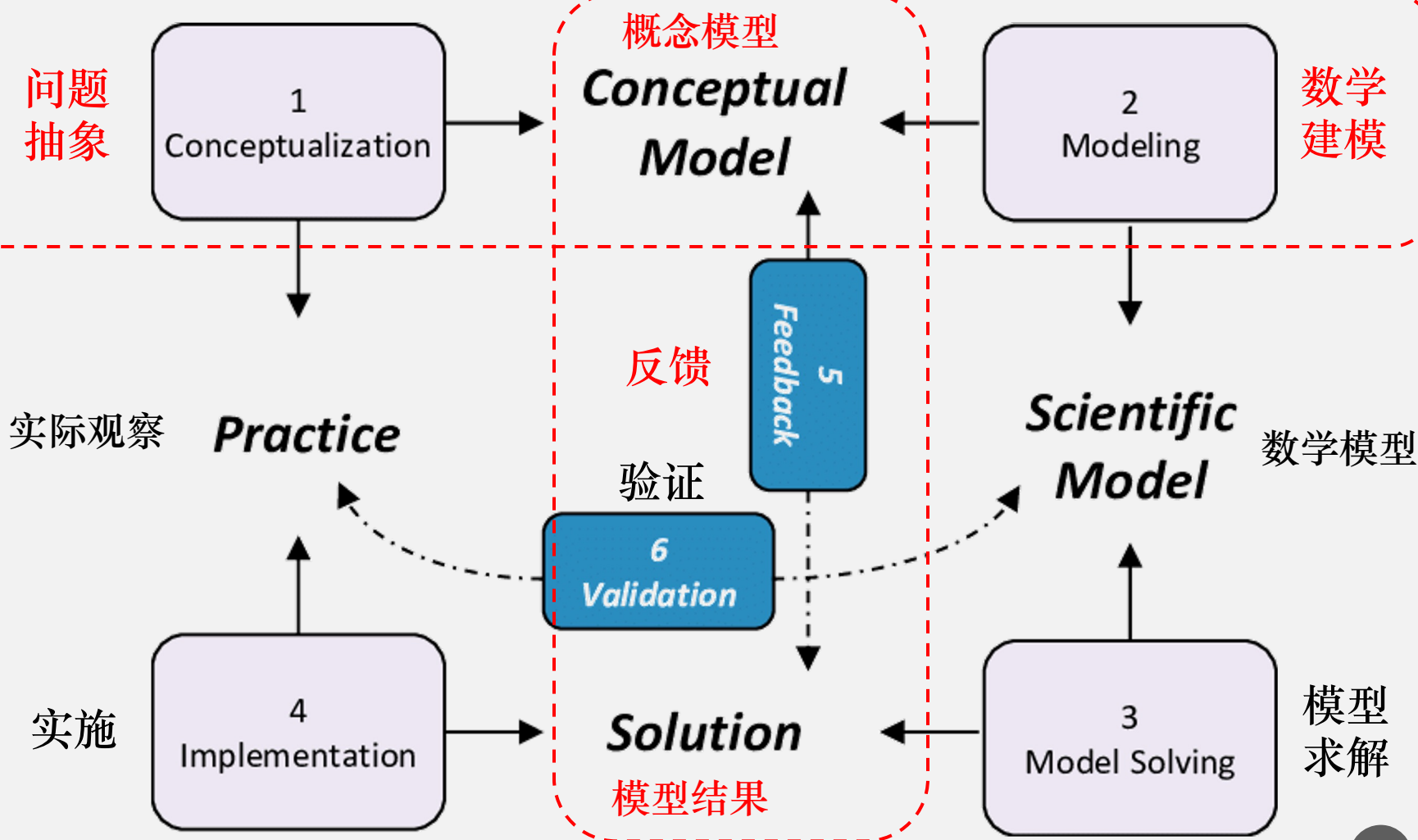
挑战四：约束

- 约束：我们受到哪些限制
 - 每个需求点必须被覆盖
 - 容量上下限约束
 - 面积约束
 - 平峰期产能约束
 - 高峰期产能约束
 - 覆盖关系规整
 - 复用现有场地（或者考虑租赁合同到期年限）
 - ...

数学建模流程

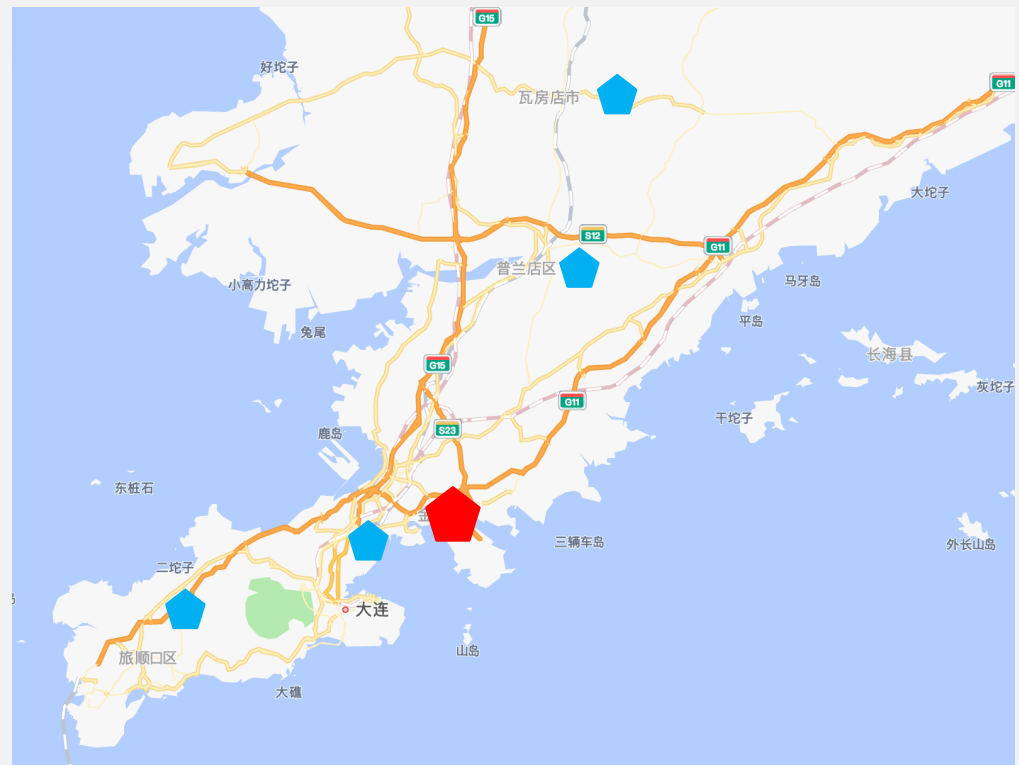


数学建模流程



选址问题实践挑战

- 结果解读
 - 算出来大连需要建4个中转场，合理么？
- 分析思路：对比现状
- 是不是运输成本太高了？
- 是不是没有考虑干线运输？



选址问题实践挑战

- 结果解读
 - 全国建仓库的案例
 - 算出来南昌的仓库覆盖拉萨，合理么？

| 始发 | 目的 | 耗时 |
|----|----|----|
| 南昌 | 拉萨 | 12 |
| 成都 | 拉萨 | 24 |

选址问题应用（CALLBACK）

- 应用场景
 - 便利店选址
 - 房屋中介门店选址
 - 超市选址
 - 中转场选址
 - 快递网点选址
 - 银行网点选址
 - 核酸点选址
 -
- 决策维度
 - 战略规划
 - 年度预算
 - 考核监控
 - 落地执行
- 个人作业：挑选一个选址应用场景，（1）描述一下这个场景，（2）描述一下这个场景中，应用经典模型（比如P-中值问题）有哪些挑战（数据、决策、目标、约束维度）。

总结

TAKEAWAY

- 选址问题在业界的应用场景丰富
 - 但是不是书本上的经典模型直接套用就能解决问题
 - 学术研究追求普适科学规律，业界应用关注具体细分场景
- 对于数学模型批判性思维（critical thinking）的萌芽
 - 实际场景是什么样子的？做决策的角度来说，权衡（trade-off）是什么？
 - 数学模型是否刻画了实际场景中的关键因素？
 - 数学模型是否和实际场景有偏差？这个偏差影响有多大？
 - 模型输出的结果是否合理？是否有管理启示（managerial insights）？
- 在实际运筹项目中，“题目出的对不对”比“题目解的对不对”更重要

Q&A