

上海交通大学MCM竞赛培训

2019年12月13日培训记录

(组织: 高晓泓@CS.SJTU, 记录人: 47-李欣杨)

MCM2020第二次培训 (线下)			
第13周: 2019年12月13日周五、地点: SY200报告厅、主持: 李之尧, Log记录人: (MCM2019D-M)			
时间	板块	内容	主讲人
18:15-19:15	模型讲解	排队论	冯润康 (CMO2015-银牌、MCM2018B-M、MCM2019B-M)
19:15-20:15	写作讲解	Latex使用概述	陈澈 (MCM2019B-M)
20:15-20:25	中场修整		
20:25-21:40	真题经验谈 2019D Time to leave the Louvre (逃离卢浮宫)	O/F/M/H/S 5种试卷比较	刘国航、刘陈正轶、赵经纬 (MCM2019D-M)
		历届D题概览	

模型讲解: 排队论与随机过程

主讲人: 冯润康

排队论用到的知识

需要高等数学、线性代数、概率统计的知识。实际上, 排队论会在大三的数学专业课中学到。

排队论概述

排队论是指从概率论与随机过程的角度, 对排队系统模型给出严格的理论推导。

看到跟排队有关的词汇 (包括实际的和虚拟的), 例如游乐场、红绿灯, 就可以考虑到使用排队论模型。

排队论模型过于大势所趋, 大家都喜欢用。所以不能仅仅套用排队论模型, 要全面考虑具体问题。例如: 食堂排队中, 学生的口味影响。

数学基础

马尔科夫过程: 这一时刻的状态只依赖于上一时刻的状态。每一种状态转移模式都有一定的概率, 同一个状态的出概率和、入概率和永远各为 100%。

泊松分布: 适用于描述单位时间内随机事件发生的次数。

泊松过程: 长度为 t 的任意时间区间的室间隔数服从泊松分布。

排队论模型

经典排队论类型包括 $M/M/1$ 、 $M/M/k$ 、 $M/G/1$ 等模型。

其中， M 表示指数分布， G 表示一般分布，第三个数表示服务台个数。

价格方程：设定某种付费规则，对处于某种状态的客户进行收费。价格方程对于所有排队模型都成立。

稳态概率：时间趋于无穷，系统达到稳态时，系统中有 n 个客户的概率

MM1 模型

最简单的模型，系统状态为系统内的客户数量，状态转移是客户进入、离开的过程。

由于这个模型过于抽象，它对美赛没什么帮助，必须对基本模型进行一些变形让它更接近实际。

一些常用的变形

系统最多容纳 n 人（例如排队时，看到队伍太长就直接离开了）

生灭排队模型（没讲）

MMK 型（增加服务台）

其他变形

MG1（离开概率为一般分布）

GM1（到达概率为一般分布）

批量到达

优先级队列（考虑到 vip、残障人士优先）

中断服务

服务串联（客户会连续受到多项服务，必须完成前一项才能开始后一项）

K 条服务线（更为复杂的服务串联情景，接近真实情况）

排队论固定的分析思路

①确定系统可能存在的所有状态，分析各状态之间的转化关系

②列出方程

③求解方程

④计算性能指标

第二步是最困难的，因为转化关系会受到多方面影响，很难找到合适的转化方程。

第三第四步比较简单，因为排队论列出的方程不会出现解不出来的情况，各性能指标直接套公式即可。

完成了这四个流程后，还需要带入具体数值进行结果分析比较。

使用了排队论的 o 奖论文的创新点

2003ICM C 题（行李安检）

将顾客按照航班起飞时间分到不同时间段，得到不同时间段的排队情况

2005MCM B 题（收费站）

考虑到司机如果发现一个地方比较堵，会加入旁边的队伍

模型延伸

①如果需要确定性能指标中的未知参数，可以转化为规划问题。注意性能指标是在平均意义下求得的。

②如果系统较为复杂，可以使用元胞自动机进行模拟排队。误差较大，需要多次试验进行误差分析。

写作讲解：LaTeX 入门

主讲人：陈澈

概述：

Word：直接打字，输出的内容是打字打出来的

LaTeX：打命令，通过编辑器的编辑才能得到你想要的东西

为什么要用 LaTeX？对文件支持，公式编辑，图片表格引用（尤其是公式编辑）

MacTeX：苹果用户专属

MikTeX：文件小，但是美赛所需的很多工具包都没有

TeXlive：推荐

TeXStudio：有点丑

Texmaker：好看一些，推荐

Sublime：程序员特供

sharelatex/overleaf：不稳定，不建议用

LaTeX 使用建议：

提前练习使用模板

最好大家都会用 LaTeX

各组员尽量用统一的编译软件（除非系统不同）

统一代码规范

少用新指令

提前配置好编译环境

注意页边距之类的隐形错误，美赛对格式有具体要求，别到最后发现了来不及改

LaTeX 和 word 一样，要自己多多使用才能熟练掌握。

在生成的文件中，“.Tex” “.cls” “.bib”，是需要手动编辑的。“pdf” 为生成的论文。

PPT 中没有写到的注意事项

所有命令以斜杠开头，否则作为文本处理

最好不要改 `usepackage` 指令里的东西，除非你看得懂里面写了什么，并且需要装新的宏包

`Minipage` 用于分栏（只用于图片排列，否则可能会被判为格式错误，导致 `usp`）

希腊字母要用命令输入，命令中第一个字母的大小写和输出的希腊字母相同（命令中第一个字母小写，希腊字母就小写，反之亦然）。

上下标可能会导致行距不均匀。

数学字体常用的就一个，不用记其他的。

“.”“,” 对应好多种指令，注意区分。

不能直接用斜杠作为除法，但可以打出分数线。

在 <https://mathpix.com/> 将截图直接转化为公式。每个账号每个月只能免费使用 50 次，但每个邮箱都可以注册一个账号，可以轻松大量白嫖。

但是截图其实也不好找，最好的方法还是提前打好公式模板

不要用图片自带的参数调整图片位置，经常不听话。应该用 float 宏包设置，虽然也可能不听话，但概率比较低。

斜表格就是用斜线把一个格子分成好几块，在里面填入不同的文字，一般来说最左上角的格子可能会出现这种情况。

三线表用于符号表，很常用，建议提前做好模板。表格的指令比较复杂，提前弄好的话可以节约很多时间。

代码放到附录里，正文里只要加入伪代码即可。

\tag 的编号是内部的，一般直接使用 \notag 即可。

Latex 里可以画图，代码跟 matlab 有点像，但是比较丑，建议用 matlab 或者 python 画图。

脚注不要加，否则会破坏页眉页脚的固定格式。

建议少使用各种缩写，因为不太正式。

题型讲解：D 题

主讲人： 刘国航

D 题简介：

题目类型为网络科学/运筹学。模型与网络有关，方法与运筹学有关。

一、网络科学

图论基本概念

图像用边的集合、定点的集合、相邻定点集合表示。

度：与某个顶点相连的边的个数。

握手定理：所有点的度加起来等于边的数量的两倍

按照变是否有方向区分有向图、无向图，按照变是否有权重区分

子图：在原图上删除某些顶点和边后形成的图。

路径：点和边都不能重复的路径。环就是首尾相连的路径。有路径的两点有连通性。

简单图：没有重边（两个相邻点中间有多条边）和自环（某个点自己连到自己）。

树：没有环且所有点相连。

完全图：用 K_n 来表示（ n 表示顶点数）。每两个点都相邻。

二分图：顶点能分为两组，所有边都在两组之间。

r -正则图：每个点的度都是 r 。

复杂网络衡量指标

度分布：顶点的度的概率分布，在随机图中会出现。

度相关性：顶点的度与相邻点的相关性。

直径：任意两点间的最长距离。

平均路径长度：最短路径的平均值。

集聚系数：顶点集成团（即构成完全图）的程度。

中心性：顶点的重要程度。有很多种指标。

特殊的网络

星型耦合网络：中间一个点，与外围的所有点相连。

ER 随机图：每两个点之间都有 p 的概率相连，每条边的出现都是独立的。通常关注 n 很大的时候的性质，通过更改每条边出现的概率来研究规律。

WS 小世界网络：也是随机产生，但是并不是完全随机。先生成一个最近邻耦合网络，再以概率 p 随机重连每一条边。特点是集聚系数大、平均路径长度小。说人话就是它的边很少，但点很近。

BA 无标度网络：由于 WS 小世界网络的度分布不符合实际，BA 无标度网络修正了这一问题。它由空图开始增长，通过择优连接机制保证度分布规律较为符合实际。

还有 NW 小世界网络/无标度聚集网络/无标度同配网络/无标度社区网络/超图……都是对于其他模型的修订。

运筹学

运筹学的目标是找到最优方案，它涵盖了很多领域。

图算法

图的表示法：

①邻接矩阵：直接记录相邻的点对。

②邻接表：记录每个顶点相邻的点。

图搜索算法：

①DFS（深度优先搜索）：沿一条路走到头后，再回头换别的路

②BFS（广度优先搜索）：先把一个点周围的点都走一遍，再走周围点周围的点

强连通分量：

Kosaraju 算法。

最小生成树：

删除一个图中的部分边使其变为一棵树，使剩下的边权重和最小。

生成方法有 Kruskal 算法和 Prim 算法，它们都是贪心算法。

最短路径：

单源：求出某两个点之间的最短路径

①Dijkstra 算法（只适用于正权图）

②Bellman-Ford 算法（SPFA）

多源：求出每两个点之间的最短路径

①Floyd-Warshall 算法

②Johnson's 算法

网络最大流：

Ford-Fulkerson 算法

规划论

线性规划：在满足一定的线性约束的情况下，使某个变量达到极值。初中就学过线性规

划最简单的形式。

整数规划：决策变量取值为整数，一般比线性规划难做。0-1 线性规划应用较为广泛，它的决策变量取值只能为 0 或 1。

目标规划：存在多个目标，有些为硬性目标，有些为软性目标。引入偏差变量、优先因子和权系数的概念，给目标分轻重缓急，尽量达成每一个目标。

动态规划：多阶段决策，通过状态转移方程、最优值函数、边界条件建立一个递归模型。根据 bellman 最优化原理，要使当前的决策保证剩余部分最优。

动态规划是最难理解的方法，而相应地功能也最强大，还能解决前面三种规划中的问题。

决策论

决策论是在不同决策环境下的方案择优方法，通常采取期望报酬最大准则。它适用的问题通常与现实问题相关联，受很多因素影响，难以得出准确结论。

决策树：将不同的决策方案放在一棵树里进行分析。影响决策的因素通常有一定的概率，决策树通过期望报酬得出最佳决策。

博弈论

博弈论适用于竞技性现象。它通常有个假设：玩家是理性的，每个玩家都想尽可能取得最大收益。

博弈论是决策论的一种，它需要将双方的决策都纳入考虑范围。

纳什均衡：所有玩家由于担心改变策略后收益降低，都不愿意主动改变现状的局势。绝大多数博弈都会发生纳什均衡。纳什均衡点可以用线性规划的方式算出来。

真题经验谈：D 题

主讲人： 赵经纬

2016 年 D 题

题目要求：

- ①筛选出有效信息，对有效信息的流动进行建模
- ②将预测结果与现实情况进行对比，验证模型有效性
- ③预测 2050 年信息情况
- ④探究信息如何改变公众兴趣和观点
- ⑤探究人的偏见、消息来源、信息网络结构对信息的影响

o 奖论文常用模型：

图算法、传染病模型

o 奖论文案例：

在媒体之间、个人之间分别建立信息传播模型。

建立基于信息流动的网络叠加模型。

建立“生命周期”模型

17 年 d 题

题目要求：

- ①对机场安检乘客流建模，分析排队中的瓶颈
- ②针对安检提出改进建议
- ③考虑地域因素对模型的影响
- ④针对相关政策和程序提出改进建议

o 奖论文常用模型：

排队论

o 奖论文案例：

将安检的过程抽象为三次连续排队的过程
分别进行分析寻找瓶颈

18 年 d 题

题目要求：

- ①在汽油车全部转为电动车的过程中，确定最终的充电站分布
- ②设计影响车辆转化进程的政策的时间轴

o 奖论文常用模型：

logistic 模型、传染病模型

o 奖论文案例：

用无向图寻找潜在的建立充电器位置
用 SPFA 算法计算各节点的最短路径
用线性规划模型得到充电器配置方案
用博弈论设计时间轴
用 Logistic 回归模型评估影响电动车产业的因素

19 年 d 题

题目要求：

- ①开发紧急疏散模型，撤离访客、让应急人员进入
- ②找出可能限制出口移动的潜在瓶颈

o 奖论文常用模型：

dijkstra 算法（最短路径搜索）

o 奖论文案例：

将逃跑过程分为四类：室内（非拥堵），进出口处，楼梯处，意外情况（例如摔倒、混乱），将四类逃跑过程所需的时间相加。

用分流的方式减缓出口处、楼梯处拥堵

对各种意外情况设定参数，增加疏散总时间

学长对自己论文的反思：

- ①公式要做得花里胡哨一些，多加点求和符号之类的高级符号
- ②可视化要做得更多更精美
- ③实验没有落地，要用真实结果对模型进行改进