Начало работ по методу Шелейховского в ВЦ ЛГУ в 1960-е гг.

(И.В. Романовский, вместо предисловия)

Меня попросили рассказать о том, как в НИИММе ЛГУ начинались работы, связанные с применениями математических методов и компьютеров в градостроительных расчетах. Время было в компьютерном отношении таким первобытным, что даже страшно вспоминать. Но попробуем.

В 1959 г. в связи с появлением в Ленинградском университете новых ставок я был переведен из очной аспирантуры матмеха ЛГУ в заочную и принят в НИИММ ЛГУ на должность младшего научного сотрудника, это была группа из пяти студентов кафедры теории вероятностей (кроме меня в ней былиО.Н. Бондарева, О.М. Калинин, Т.П. Красулина и Т.М. Товстик). Вскоре директор ВЦ ЛГУ Г.П. Самосюк предложил нашей группе перейти в Вычислительный центр, и из нее и нескольких сотрудников ВЦ (среди них был С.С. Сурин, имевший уже совместные работы по обобщенной транспортной задаче с М.К. Гавуриным и Г.Ш. Губинштейном) была создана группа исследования операций и теории вероятностей.

Я в это время уже преподавал: читал студентам-математикам курс экстремальных задач (в счет часов военной кафедры, которая заказала такой курс), а на кафедре вычислительной математики руководил созданной внутри кафедры специализацией исследования операций (25 человек ежегодно). Студенты были сильные (некоторые из них попали в ваш институт), и мы могли привлекать их к работе.

Среди задач, которыми нам предложили заниматься, была задача, предложенная сотрудником ЛенЗНИИЭП[1] Г.Д. Платоновым, – выбор перспективных планов жилищного строительства в Ленинграде, исходя из выдвинутого КПСС лозунга «каждой семье – отдельную квартиру».[2] Лозунг

---

[1] Ленинградский зональный НИИ типового и экспериментального проектирования жилых и общественных зданий и сооружений.

[2] Хотелось это добиться к 1960 г., но ясно было, что не получается. Но именно этот лозунг стоял за массовым развитием жилищного строительства во времена И. С. Хрущева. Было бы интересно оценить, сколько ленинградцев обязаны своим рождением появлению этого нового жилья.

был очень удобен для нашего исследования, он сразу решал (на бумаге) не решенную в нашем городе до сих пор проблему использования коммунальных квартир.

Я предложил нашим заказчикам использовать схему расчетов, в которой жители города в разные периоды времени как бы «выселяются» из своих квартир и заселяются снова с учетом существующих на данный момент нормативов. Заселение осуществлялось выбором для каждой квартиры (типов квартир оказалось не очень много) варианта заселения. Набор вариантов был свой для каждого из этапов, и допускались варианты с коммунальным заселением. В отдаленной (и так и не наступившей) перспективе при таком повторном заселении коммунальное использование квартир исключалось[3].

Мы в несколько итераций подготовили данные (наши трудности были поучительны, но в этой заметке речь не о них) нормально все сосчитали и вместе с заказчиками доложились на каком-то совещании в Киеве [1].

Для меня с этой работой психологически связывался все время еще один «градостроительный заказ», который у нас шел мучительно и имел гораздо большие последствия. От этого заказа сохранилась еще одна публикация [8], авторами которой были Э.Г. Мовчан, участвовавшая в работе [1], и инженер Ленпроекта А.Г. Дынкин. Нужно было рассчитать некоторую матрицу пассажирских перевозок по методу ленинградского архитектора Г.В. Шелейховского. Об этом втором заказе и пойдет речь.

А.Г. Дынкин принес мне какую-то затрепанную работу Шелейховского, скорее всего, это было довоенное литографированное издание [2].

Я читал эту работу и восхищался смелостью моделирования и тем, с какой легкостью автор находит нужные для моделирования числовые данные, комбинируя (скажем так, – я точно не помню) данные о занятости

[3] Тогда-то я и узнал от коллег-архитекторов, что в нашем городе существуют такие коммунальные кошмары, как 4-комнатные квартиры площадью больше 200 кв. м. Квартира, в которой прошло мое детство, принадлежала высокооплачиваемому чиновнику, в ней достаточно свободно размещалось 9 семейств. Помню, как я был поражен, когда коллега (М. Е.Розина) объясняла мне типичное устройство большой квартиры, и я впервые «понял» нашу квартиру. Для этой квартиры светлое будущее так еще и не наступило.

лондонских таксистов и затраты на бензин в Берлине. Здесь мы тогда нашим градостроителям ничем помочь не могли, данные нам предоставил А.Г. Дынкин (методика подготовки данных очень детально описана в [8]), а программу для расчетов составила сотрудница лаборатории исследования операций молодая выпускница матмеха Э.Г. Мовчан. Программа была написана в кодах ЭВМ М-20, трансляторов у нас тогда не было[4]. Алгоритм был очень простым:

Пусть $I = 1 : m$, $J = 1 : n$, задана матрица положительных весов $W = \|w_{i,j}\|_{IJ}$, т. е. $i \in I, j \in J$, и положительные векторы $A=(a_i)_I$ и $B= b_j)_J$, причем $\sum_{i \in I} a_i = \sum_{j \in J} b_j$.

Требуется найти положительную матрицу $X=\|x_{i,j}\|_{IJ}$, удовлетворяющую «транспортным ограничениям» $\sum_{j \in J} x_{ij} = a_i$, $\sum_{i \in I} x_{ij} = b_j$,

используя следующий алгоритм Шелейховского (потом его стали называть алгоритмом Шацкого-Шелейховского):

1. Положить $X:=W$;
2. Для каждого $i \in I$ перенормировать i-ю строку матрицы X:
найти $S = \sum_{j \in J} x_{ij}/a_i$, а затем положить $x_{ij} = x_{ij} /S$, $j \in J$.
3. Для каждого $j \in J$ перенормировать j-й столбец матрицы X:
найти $S = \sum_{i \in I} x_{ij}/b_j$, а затем положить $x_{ij} = x_{ij} /S$, $i \in I$.
4. Если при каком-либо j значение S будет существенно отличаться от 1, вернуться к шагу 2.

Было интересно выяснить, что же дает этот странный, но быстро сходящийся алгоритм. Прежде всего, было ясно, что в расчетах все время выдерживается такая связь матриц W и X: $x_{ij}=w_{ij}\times\alpha_i\times\beta_j$; фактически пересчитываются только коэффициенты α_i и β_j.

Отсюда следует, что если взять произвольные индексы $i < m$ и $j < n$, то

[4] Но в статью [8] программа на Алголе-60 была вставлена

всегда выполняются соотношения : $\dfrac{x_{ij} / w_{ij}}{x_{in} / w_{in}} = \dfrac{x_{mj} / w_{mj}}{x_{mn} / w_{mn}}$

Это равенство было мне чем-то знакомо. И действительно, похожие равенства (немного более простые) встречаются, когда вы рассматриваете задачу о максимуме энтропии. Вскоре стало понятно, что если рассмотреть задачу о максимизации при транспортных ограничениях «взвешенной энтропии» $\sum\limits_{i,j} x_{ij} \times \ln (w_{ij}/x_{ij})$,

и алгоритм Шелейховского сходится, то он сходится к оптимальному решению.

Я попросил своего аспиранта, Льва Мееровича Брэгмана, заняться вопросом о сходимости алгоритма Шелейховского, и очень быстро он эту сходимость доказал. Его доказательство было опубликовано в [36], а затем более подробно в [10]. Вскоре Брэгман обобщил метод на задачу поиска максимума взвешенной энтропии для произвольных линейных ограничений [11]. Брэгман нашел формулу для пересчета текущего решения $X=\{x_i\}_\mathrm{N}$ в случае ограничения общего вида. Скажем, если требуется изменить решение так, чтобы оно удовлетворяло ограничению $\sum\limits_{k\in N} a_{ik}x_k = b_i$

то новое приближение X^* ищется в виде $x^*_к = \exp(\lambda\times a_{ik})x_k$,
где λ – единственный корень уравнения $\sum\limits_{k\in N} a_{ik} \exp(\lambda\times a_{ik})x_k = b_i$.

За прошедшее с тех пор время метод Брэгмана стал всемирно признан,[5] причем особое внимание уделяется тому способу, которым была доказана сходимость исходного метода Шелейховского. Из многочисленных ссылок выберем недавнюю статью [37], в которой вводятся диаграммы Брэгмана-Вороного.

А я сам пытался понять, почему в градостроительных задачах может возникать распределение с максимальной энтропией, встречавшееся нам ранее разве лишь в моделях статистической физики. Аналогии города и

[5] Каждый может убедиться в этом в Интернете

модели с совокупностью молекул были достаточно убедительны для кулуарных разговоров, но не более.

Мне запомнилась беседа с профессором физического факультета Г.Ф. Друкаревым, замечательно читавшим у нас на матмехе курс физики. Я подошел к нему, рассказал о своих раздумьях и высказал опасение, что семья в модели города мало похожа на частицу из статфизической модели. Он мне ответил вопросам «А почему вы думаете, что реальная частица больше похожа на частицу из модели?». Возразить было нечего.

Рассказывал я об этой задаче и А.Н. Колмогорову, приезжавшему на матмех для чтения лекций. Никаких советов не последовало, и это было вполне естественно – ситуация была очень непривычная.

Примерно в то же время я рассказал об этом круге задач Б.Г. Питтелю, который занимался такими «молекулярными» моделями в связи с моделированием игр [13]. Бориса Герсоновича модель заинтересовала, и вскоре появилась его статья [14] с простейшей моделью «коллективного поведения». Со статей [14-16] начался огромный цикл исследований, в которых математические методы успешно применялись к задачам градостроительства. Но не мне вам об этом рассказывать.

Мне очень жаль, что до сих пор не нашел достаточного развития вопрос об использовании метода Брэгмана для приближенного решения задач линейного программирования.

Пусть мы хотим решить, например, линейную задачу – максимизировать доход $CX = \sum_{i,j} x_{ij} \times c_{ij}$ при балансовых ограничениях (как в задаче оценки территории). Положим $w_{ij} = \exp(\gamma \times c_{ij})$ и будем решать задачу максимизации «взвешенной энтропии» $\sum_{i,j} x_{ij} \times \ln (w_{ij} / x_{ij})$ при указанных ограничениях.

Учитывая выражение для w_{ij}, имеем

$$\max \Rightarrow \sum_{i,j} x_{ij} \times \ln (w_{ij} / x_{ij}) = \sum_{i,j} x_{ij} \times (\gamma \times c_{ij} - \ln x_{ij}) = \gamma \times CX + H(X),$$

или $\max \Rightarrow CX + (1/\gamma) \times H(X)$. ($H(X) = \sum_{i,j} x_{ij} \times \ln x_{ij}$ как в молекулярной физике).

Для транспортной задачи w_{ij}=exp(-$\gamma\times t_{ij}$), и соответственно,
max =>-TX+($1/\gamma$)×$H(X)$ (т.е. минимизируем TX). Значит, мы максимизируем линейную функцию с энтропийной добавкой – штрафом.

При малом коэффициенте в энтропийном слагаемом (т.е. при больших значениях γ) решение этой сглаженной задачи по значению исходной целевой функции CX будет мало отличаться от решения исходной задачи линейного программирования (в данном случае, транспортной задачи). Но у него будет замечательная особенность, – если в задаче есть несколько оптимальных решений, то в приближенном решении они все будут представлены в оптимальном решении, в отличие от традиционной базисности метода потенциалов и других «симплекс-методов». Это дает основание говорить о «мягких» решениях и сравнивать их с известным «разверсточным» распределением примитивного управления.[6] Так что мы не знаем решения линейной задачи, но знаем, что чем больше значение γ, тем ближе решение энтропийной задачи к максимуму CX.

Конечно, здесь дело не только в нескольких оптимальных решениях, а в том, что есть много решений, близких к оптимальному. Решение энтропийной оптимизационной задачи является центральной точкой огромного множества разных решений, и можно ожидать, что чем меньше коэффициент при энтропийном слагаемом, тем больший вес в совокупности имеют решения, близкие по значению целевой функции к оптимальному решению линейной задачи. При этом нужно иметь в виду, что как правило, показатели типа c_{ij} или t_{ij} известны приблизительно, и детерминизм традиционных методов в этой ситуации играет негативную, роль, а иногда затрудняет дальнейшее приспособление решения задачи линейного программирования, приближающего исходную ситуацию, к деталям, не вошедшим в модель.

[6] Неправильно думать, что разверстка – это подход к задачам распределения, характерный только для эпохи военного коммунизма. По словарю Даля «разверстать — делить по-равну или по скольку придется». Соответствующий английский термин *apportionment* относится к середине XVI в