19373545 彭兴宇

19373084 倪云昊

19373433 张必浪

19373418 王卓

1 背景

现在有一个大作业交由一个小组来完成，在进行小组分工时，我们可能遇到以下问题：

①应该按什么原则分工？

②大作业由多个环节组成，当有两个或多个人都想要相同任务环节时，这个任务环节分配给谁？被分配到这个任务环节的人如何对未分配到的人做出补偿？

③每个人对小组工作的贡献不同，如何设置奖励分配规则才公平合理？

2 建模

2.1 基础模型

模型

①一个大作业有n个选题，小组内有n位学生。

②每个学生只能完成大作业中的一个选题。

③有一张$n\times n$的得分表R，R_{ij}表示学生i完成选题j能得到的分数（保证非负）。

④若大作业的所有选题都完成了，则组里每个学生获得其完成选题的相应分数；否则每位学生的得分均为0。

⑤为保证诚实性，将对④的结果应用VCG机制，算出每个学生的最终收益。

例

一项任务有A，B，C，D四个环节，有甲、乙、丙、丁四个人，每个人只能从四个环节中选一个来完成，且必须四个环节都完成了，四个人才能得到收益，否则四个人的收益都为0.

由于四个人各自的能力与擅长的环节不同，对于同一个环节，不同的人来完成会获得不同的分数。现在，甲乙丙丁四个小伙伴凑在一起，报出自己对于完成每个环节所得分数的预期（如下表），然后决定每个人的分工。（我们假设每个人对自己的能力估计是完全准确的，例如若甲对自己完成环节A的估值为6，且甲真的被分配去完成环节A了，那么甲最后得到的收益就一定是6）。

	A	B	C	D
甲	2	5	4	1
乙	3	3	2	2
丙	5	1	2	3
丁	1	1	2	4

一个分工的理所当然的目标是最大化社会福利，即最大化四人的收益之和，那么在本例中，一个最大化社会福利的方案为甲B，乙C，丙A，丁D。

然而，最大化社会福利的前提是我们要保证四人报出的值是诚实的——用一用VCG吧！

假设甲不参与，则其他三人无法完成任务，总收益为0；而甲加入后，任务可以完成，且其他三人的收益之和为2+5+4=11.

根据VCG机制，甲的最终收益为5-(0-11)=16.

假设乙不参与，则其他三人无法完成任务，总收益为0；而乙加入后，任务可以完成，且其他三人的收益之和为2+5+4=14.

根据VCG机制，乙的最终收益为2-(0-14)=16.

事实上，在这种机制下，每个人的得分都会等于小组的社会福利——每个成员都获得自己完成环节的得分，并且由于这个成员的存在，让其他三个成员能够得到分数，于是其他三人完成环节的得分也会在VCG的作用下加给这个成员。

小组成员每个人的最终收益都相同，这忽视了成员对小组贡献的不同。于是我们提出下面的倍乘机制。

2.2 倍乘机制

模型

①一个大作业有n个选题，共有m位学生要完成大作业。

②每个学生只能完成大作业中的一个选题。

③有一张$m \times n$的得分表R，R_{ij}表示学生i完成选题j能得到的分数（保证非负）。

④学生可以自发地组队，但要求属于同一队的学生的选题各不相同。

⑤每个学生的初始得分=$2^{\text{小组人数}} \times \text{个人选题得分}$。

⑥为保证诚实性，将对⑤的结果应用VCG机制，算出每个学生的最终收益。

例

我们仍然使用这张表格，最佳分工仍是甲B，乙C，丙A，丁D。

	A	B	C	D
甲	2	5	4	1
乙	3	3	2	2
丙	5	1	2	3
丁	1	1	2	4

甲的初始得分为$2^4 \times 5 = 80$。把甲剔除后，最优方案是乙B、丙A、丁D，所以甲的最终收益为

$$80 - [2^3 \times (3+5+4) - 2^4 \times (2+5+4)] = 160$$

乙的初始得分为$2^4 \times 2 = 32$。把乙剔除后，最优方案是甲B、丙A、丁D，所以乙的最终收益为

$$32 - [2^3 \times (5+5+4) - 2^4 \times (5+5+4)] = 144$$

丙的初始得分为$2^4 \times 5 = 80$。把丙剔除后，最优方案是甲B、乙A、丁D，所以丙的最终收益为

$$80 - [2^3 \times (5+3+4) - 2^4 \times (5+2+4)] = 160$$

丁的初始得分为$2^4\times 4=64$。把丁剔除后，最优方案是甲B、乙C、丙A，所以丙的最终收益为

$$64-[2^3\times(5+2+5)-2^4\times(5+2+5)]=160$$

可以做点直觉上的认识：

①甲、丙、丁都是把自己最擅长的技能施展出来，且每个人所负责的选题都是全组成员负责这个选题中得分最高的，所以三个人都对小组做出了最大的贡献，三个人得到了相同的最终收益（尽管丁的选题分数是4而甲和丙是5）；

②整体来看乙能力稍弱，其负责任何一个选题都不是小组成员中得分最高的，于是最终收益不及其他三人；但乙为了让甲和丙能去负责其得分最高的选题，放弃了个人得分为3的A和B，而选择了得分只有2的C——于是乎，尽管乙的初始得分只有32，但它这一“谦让”的行为也让它获得了极大的奖励，得到了一个与甲和丙相差不那么大的最终收益。（具体而言，只看初始得分，乙比甲和丙少80-32=48；但看最终收益，乙只比丙和甲少160-144=16）。

3 分析

3.1 诚实性

用VCG，这合理吗？

这其实是在问我们的倍乘机制是不是一个多参数机制设计环境。

要是一个多参数机制设计环境，需要满足以下的条件：

Q：参与者有限吗？

A：n个参与者，是有限的。

Q：结果集合有限吗？

A：n个参与者只有n！种分工结果（相当于全排列），是有限的。

Q：有估值函数和报价函数吗？

A：其实R表就给出了这个函数。对于每一个分工结果，我们都可以通过查R表，再乘以2^n映射到估值（或报价）。

Q：支付函数怎样确定的？

A：利用VCG。

显然，我们的倍乘机制是一个多参数机制的设计环境，那么对其应用VCG机制就能够实现诚实性。

撒个谎试试

我们来看看在应用VCG下的倍乘机制里欺骗会发生什么。

这是我们前面用过的那张表格：

	A	B	C	D
甲	2	5	4	1
乙	3	3	2	2
丙	5	1	2	3
丁	1	1	2	4

甲、丙、丁都拿到了自己心目中最好的选题了，没欺骗的必要。对于乙而言，假设他特别想要选题B，于是把自己完成选题B的得分谎报为5，于是表格变成了下面这样：

	A	B	C	D
甲	2	5	4	1
乙	3	5	2	2
丙	5	1	2	3
丁	1	1	2	4

现在的最佳分工为甲C、乙B、丙A、丁D。我们只需关注乙的最终得分会如何变化就行了。

别忘了5分是谎报的，乙的初始得分只有$3 \times 2^4 = 48$。

把乙剔除后，最佳分工为甲B、丙A、丁D，于是乙的最终得分为：

$$48 - [2^3 \times (5+5+4) - 2^4 \times (4+5+4)] = 144$$

在前面我们就计算过，乙诚实报价时的收益就是144，因此撒谎并没有让乙得到更高的收益。

3.2 分配原理

推导最终收益的公式

倍乘机制下，小组成员的最终收益是怎样的呢？我们可以做理论上的推导。

假如小组里只有一个成员，那么该成员的最终收益就是他完成各选题得分中的最大值。

假如小组有n位成员（$n \geq 2$），那么对任意成员i，我们假设其完成被分配的选题后的得分为w_i，于是成员i的初始得分为：

$$2^n w_i$$

设小组的最大选题得分之和为W，在把i剔除后小组剩余成员最大选题得分之和为W_i'（此处的W和W_i'都还没乘倍乘因子，就只是表格里的数加起来得到的）。于是根据VCG机制，成员i的最终收益为

$$u_i = 2^n w_i - [2^{n-1} W_i' - 2^n (W - w_i)] = 2^n W - 2^{n-1} W_i'$$

这个公式有何意义？

得到结果了！我们先来来分析这个结果有何意义：

①是满足个体理性的，即不会出现收益为负的情况。这是因为$W \geq W_i'$必然成立（在W'的分工基础上加入i时，保持原有成员的分工都不变，然后随便给i一个没人选的选题，由于w_i非负，则必有$W_i' + w_i \geq W_i'$，而W是i加入后的小组最大选题得分之和，故有$W \geq W_i' + w_i \geq W_i'$），于是$u_i = 2^n W - 2^{n-1} W_i' = 2^{n-1}(2W - W_i') \geq 0$。

②解决了基础模型组内所有人收益都相同的的问题。对每一位成员而言，W是相同的，但W_i'未必是相同的，这就能够使得成员的最终收益产生不同。

③直观来看，W反映了整个团队的总实力，每个成员的个人能力越强，就能使得W越大。而W_i'表示把i剔除后剩余成员能够取得的最大收益——若W_i'很大，意味着把i剔除了也没造成多大损失，意味着i对于小组没那么重要，所以i的u_i就没有那么大；反之，若W_i'很小，意味着把i剔除会导致小组的得分之和大幅减少，这意味着i对于小组而言是很重要的，所以i的u_i会比较大。

④我们的倍乘机制还挺人文关怀的，体现在W的倍数是2^n而W_i'的倍数是2^{n-1}。W反映了全体成员的共同努力，而这个值也被全体成员所共享；W_i'用来体现差异，用以反映不同成员对于小组的重要性（也即对小组贡献的一种反映）。但W_i'的倍数只有W的一半，因此这种差异也不会让那些相对处于劣势的组员收益太差（例如前面乙得144而甲乙丙得160）。又体现了差异，又不至于让这个差异太过悬殊，就这点而言我们的倍乘机制还做得蛮好的。

⑤最终收益是与集体相关的变量决定的，因此在满足小组选题得分之和最大的分工安排有多种时，不会产生纷争。

3.3 积极性

多多益善

在基础模型中，我们要求每个组的人员数一定要凑满，不然就无法获得收益——但在倍乘机制中，我们对这点进行了放松，学生们可以自由决定组内人员数。然而，接下来我们会证明，倍乘机制是多多益善的，学生们都愿意组成更大的小组——小组人数越多，小组内每个人的最终收益越高。我们把倍乘机制的这种属性称为积极性。

积极性定理

设大作业一共有n个选题。任取有$m-1$个成员的小组G（$2 \le m \le n$），任取独立的学生i，在i加入G后，不会使i和G内所有原有成员的最终收益变差。

证明

先考虑i的最终收益，假设i一人成组时，所能得到的最大收益为w_i^*，在其加入G后完成被分配选题的得分为w_i。

假设i尚未加入G时，G的最大选题得分之和为W_i'；而在i加入G后，G的最大选题得分之和为W。

那我们的目标就是证明

$$2^m W - 2^{m-1} W_i' \ge w_i^*$$

即证

$$2^{m-1}(2W - W_i') \ge w_i^*$$

假设让i选择得分w_i^*对应的选题时，其余成员的最大选题得分之和为$W_i'^*$，因为W是i加入后的小组最大选题得分之和，所以

$$W \ge w_i^* + W_i'^* \ge w_i^*$$

又上一节中已经得到$W \ge W_i'$，于是

$$2^{m-1}(2W - W_i') \ge 2W - W_i' \ge w_i^* + W_i' - W_i' \ge w_i^*$$

故证得i加入G后，其最终收益不会比原来差。

接着考察G中任意的原有成员j。在i未加入G时，j的最终收益为

$$2^{m-1} W_i' - 2^{m-2} W_{ij}''$$

其中W_{ij}''是在没有i的G中把j剔除后，剩余成员的最大选题得分之和。

而在i加入G后，j的最终收益为

$$2^m W - 2^{m-1} W_j'$$

其中W_j'是在i加入后的G中把j剔除后，剩余成员的最大选题得分之和。

那我们的目标就是证明

$$2^mW-2^{m-1}W_j'\geq 2^{m-1}W_i'-2^{m-2}W_{ij}''$$

即证明

$$4W-2W_j'\geq 2W_i'-W_{ij}''$$

因为$W\geq W_i',W\geq W_j'$，所以

$$4W-2W_j'\geq 2W_i'+2W_j'-2W_j'=2W_i'\geq 2W_i'-W_{ij}''$$

证毕。

意义

积极性定理告诉我们，如果一个大作业有n个选题，那么小组一定会凑满n个人。这使得我们的倍乘机制在解决基础模型所有人的最终收益都相同的问题时，仍然保留了“大家想要把小组人数凑满”的性质。当一项大作业的各部分关系紧密、必须要每个部分都完成才有意义时，这样的性质就尤为重要。

3.4 最优分工求解

假设有n个组员对应n个任务，其中第i个组员对第j个任务的评估值为w_{ij}，且分配方案为$X=\{x_{ij}\}_{n\times n}$，其中$x_{ij}=1$表示将第j个任务分配给第i个组员，否则$x_{ij}=0$。要想求得最佳分配方案，等价于则求解如下的一个规划(P)：

$$\begin{aligned}
Max\qquad & f(X)=\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}x_{ij}\\
s.t.\qquad & \sum_{i=1}^{n}x_{ij}=1,j=1,\cdots,n\\
& \sum_{j=1}^{n}x_{ij}=1,i=1,\cdots,n\\
& x_{ij}\in\{0,1\},i,j=1,\cdots,n
\end{aligned}$$

规划(P)的可行域是离散的，我们考虑一个可行域增广的规划(AP)：

$$\begin{aligned}
Max\qquad & f(X)=\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}x_{ij}\\
s.t.\qquad & \sum_{i=1}^{n}x_{ij}=1,j=1,\cdots,n\\
& \sum_{j=1}^{n}x_{ij}=1,i=1,\cdots,n\\
& x_{ij}\geq 0,i,j=1,\cdots,n
\end{aligned}$$

下证，规划(P)的解可以由规划(AP)的解转换得到。

Step 1：运用单纯形法或表上作业法得到(AP)的一个最优解X^*

Step 2：若X^*含非$0,1$元素，则找到由非$0,1$元素围成的交错闭回路

下面证明交错闭回路的存在性。

显然，若可行解X的某行或某列有非$0,1$元素，则在该行或该列至少存在两个非$0,1$元素（引理1）。

定义 交错闭回路：如果图C是一个环，且其所有边是水平或竖直的，并且任意相邻两条边相互垂直，那么称图C为交错闭回路。

首先，将X中的所有非$0,1$元素视作为图G的顶点，若两顶点在同一行或同一列，则连上一条边。

根据引理1，图G中每一个顶点的度至少为2——与该顶点邻接的边，至少有一条边连接同一行的另一个顶点，以及另一条边连接同一列的另一个顶点，且这两条边相互垂直。

任选一顶点v_1沿某条边e_1出发搜索至下一个顶点v_2，必定能够从v_2找到一条边e_2继续搜索，且e_2与e_1相互垂直，依次递推下去，均保证e_k与e_{k+1}相互垂直。

由于顶点与边的数目有限，搜索过程中一定会找到之前搜索到过的顶点v，从而找到闭回路$v_k e_k v_{k+1} e_{k+1} \cdots e_{m-1} v_m e_m v_k$，我们此时只用考虑$e_k$与$e_m$是否相互垂直即可。

若e_k与e_m相互垂直，则该闭回路即为交错闭回路；

否则，e_k与e_m共线，因此v_m, v_k, v_{k+1}共线，则v_m, v_{k+1}之间存在边e'且一定与e_{k+1}, e_{m-1}分别垂直，那么用$v_m e' v_{k+1}$替换掉$v_m e_m v_k e_k v_{k+1}$得到的闭回路为交错闭回路。

Step 3：含非$0,1$元素的最优解X^*，沿交错闭回路进行以下方式转化为另一个最优解X'，且使得其非$0,1$元素数目能够严格减少

引理 2 交错闭回路的顶点数为偶数

在该交错闭回路中取子链$v_1 e_1 v_2 \cdots v_m e_m v_1$以及$v_m e_m v_1 e_1, v_2$

若$m = 2k+1$，则在子链$v_1 e_1 v_2 \cdots v_m e_m v_1$中，可根据递推关系知$e_1, e_3, \cdots, e_m$方向相同

而在子链$v_m e_m v_1 e_1, v_2$中，有e_m与e_1相互垂直，产生矛盾

因此m只能为偶数，故该回路的顶点数也为偶数

根据**引理 2**，在最优解X^*中，存在一条由非$0,1$元素围成的顶点数为偶数的交错闭回路$v_1 e_1 v_2 \cdots v_{2k} e_{2k} v_1$，不妨设其对应的决策变量为$x_1, \cdots, x_{2k}$，以及价值评估系数为$w_1, \cdots, w_{2k}$

并取：

$$\Delta x = \min_{1 \le i \le 2k} \{x_i\} > 0$$

假设存在$x_k, x_k + \Delta x > 1$，则$x_k > 1 - \Delta x, x_{k-1} \le 1 - x_k < \Delta x$，与$\Delta x$的定义矛盾。

因此$\forall j \le 2k, x_j - \Delta x \ge 0, x_j + \Delta x \le 1$，定义

$$x^+ = (x_1 + \Delta x, x_2 - \Delta x, \cdots, x_{2k-1} + \Delta x, x_{2k} - \Delta x), 0 \le x^+ \le 1$$
$$x^- = (x_1 - \Delta x, x_2 + \Delta x, \cdots, x_{2k-1} - \Delta x, x_{2k} + \Delta x), 0 \le x^- \le 1$$

考虑用x^+或x^-替代最优解X^*中的$(x_1, x_2, \cdots, x_{2k-1}, x_{2k})$得到的$X_+^*, X_-^*$，依然满足不等式约束

不妨设x_1, x_2在同一行，则$\{x_1 + x_2, x_3 + x_4, \cdots, x_{2k-1} + x_{2k}\}$以及$\{x_{2k} + x_1, x_2 + x_3, \cdots, x_{2k-2} + x_{2k-1}\}$分别影响行等式约束和列等式约束

由于回路的交错性，满足$x_i^+ + x_{i+1}^+ = x_i + x_{i+1} = x_i^- + x_{i+1}^-$

所以x^+, x^-依然满足等式约束

于是X_+^*, X_-^*依然是规划的可行解，代入可以分别得到新的目标函数值：

$$f(X_+^*) = f(X^*) - \sum_{i=1}^{2k} (-1)^i w_i \Delta x$$
$$f(X_-^*) = f(X^*) + \sum_{i=1}^{2k} (-1)^i w_i \Delta x$$

由于X^*是最优解，所以$f(X^*)$是最优值，满足

$$f(X^*) \geq f(X^*_+) = f(X^*) - \sum_{i=1}^{2k}(-1)^i w_i \Delta x$$

$$f(X^*) \geq f(X^*_-) = f(X^*) + \sum_{i=1}^{2k}(-1)^i w_i \Delta x$$

从而得到

$$\sum_{i=1}^{2k}(-1)^i w_i \Delta x = 0$$

$$f(X^*_-) = f(X^*_+) = f(X^*)$$

因此X^*_-, X^*_+也是最优解

若$\Delta x = x_{2m}$，则取$X' = X^*_+$，使得$x^+_{2m} = 0$

若$\Delta x = x_{2m-1}$，则取$X' = X^*_-$，使得$x^-_{2m-1} = 0$

从而使得最优解X'中的非$0, 1$元素比X^*至少减少一个

Step 4：令$X^* = X'$，重复Step 2-3，直至最优解中的非$0, 1$元素被全部消除

由于X^*中的非$0, 1$元素是有限的，所以在有限次迭代后，最优解X^*可以在维持最优的情况下消除非$0, 1$元素，转化成Y^*的形式。

4 应用

考虑最终得分范围的实例

实际应用时，我们要考虑每位学生的最终收益，也即大作业的最终得分，要维持在一个比较合理的区间。假设大作业有4个选题，如果我们希望最后学生们的得分大致分布在25~35这个区间上，那么单个选题得分设定在0.5~1上就比较合理（这当然是靠尝试与估算得来的结果）。

给一个实例。下表中的数据是随机生成的。

	A	B	C	D
甲	0.9	0.5	0.8	0.6
乙	0.9	0.5	0.7	0.8
丙	0.8	0.6	0.5	0.5
丁	0.9	0.8	0.9	0.5

最优分工为甲A，乙D，丙B，丁C，最大选题得分之和为3.2。

把甲剔除后，最优分工为乙D，丙A，丁C，选题得分之和为2.5，所以甲的最终得分为

$$2^4 \times 3.2 - 2^3 \times 2.5 = 31.2$$

把乙剔除后，最优分工为甲A，丙B，丁C，选题得分之和为2.4，所以乙的最终得分为

$$2^4 \times 3.2 - 2^3 \times 2.4 = 32.0$$

把丙剔除后，最优分工为甲A，乙D，丁C，选题得分之和为2.6，所以丙的最终得分为

$$2^4 \times 3.2 - 2^3 \times 2.6 = 30.4$$

把丁剔除后，最优分工为甲C，乙D，丙A，选题得分之和为2.4，所以丁的最终得分为

$$2^4 \times 3.2 - 2^3 \times 2.4 = 32.0$$

简单讨论一下：

①丁在A、B、C上都是最强的，而乙在D上是最强的，两人刚好优势互补，都是对小组贡献最大的，于是最终得分均为小组内最高的32.0。

②丙只有A上得分较好，但A选题上四人表现均十分好，且乙的其余三项得分均不太高，所以乙是相对处于劣势的，而其最终得分也是最低的。

③猜想由于单个选题分数的范围0.5~1太窄了，导致最后四人得分差异没那么明显。扩大单个选题分数范围至0~1或许能让最后得分的差异化更明显一些。限于篇幅这里就不再展开了……

存在的问题

每位学生对个人能力的估计未必准确。当完成选题和的实际得分与当初声明的不同时该如何处理，是一个问题。