USP

ZJU ACEE

Assignment 1

for math model

zzw4257

2024

HOMEWORK N° #

Introduction

Problem1

k 名专家对 n 件作品按从优到劣的顺序进行排序，用 $\sigma_j^i = l$ 表示专家 i 认为作品 j 位于第 l 位。记 $\sigma_i = (\sigma_1^i, \sigma_2^i, \cdots, \sigma_n^i)$ 为专家 i 的排序向量，$i = 1, \cdots, k$，$\Sigma = \{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \cdots, \boldsymbol{\sigma}_k\}$ 为 k 名专家的排序集合。

对两个 n 维向量 $\mathbf{u} = (u_1, u_2, \cdots, u_n)$ 和 $\mathbf{v} = (\nu_1, \nu_2, \cdots, \nu_n)$，定义它们之间的距离为 $L_1(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^n |u_i - v_i|$。排序 $\boldsymbol{\sigma}$ 与一组排序 $\Sigma = \{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \cdots, \boldsymbol{\sigma}_k\}$ 的综合距离定义为 $d(\sigma, \Sigma) = \sum_{i=1}^k L_1(\sigma, \sigma_i)$。对给定的 Σ，与 Σ 综合距离最小的排序记为 σ^*。

(1) 给定 $\Sigma = \{\sigma_1, \sigma_2, \cdots, \sigma_k\}$，求 n 维向量 $\mu = (\mu_1, \mu_2, \cdots, \mu_n)$，使得 $\sum_{i=1}^k L_1(\mu, \sigma_i)$ 最小。你能否根据 μ 给出 n 件作品的一种综合排序 σ'，σ' 是否也能从 μ 得到，为什么?

(2) 有人提议用 Borda 计分法给出综合排序。首先计算作品 j 的平均得分

$$\beta_j = \frac{1}{k} \sum_{i=1}^k \sigma_j^i,$$

然后按得分从小到大的顺序对作品进行排序（得分相同的作品之间的顺序可任意确定），由此给出一种综合排序 σ''。证明：对任意 j

$$\sum_{i=1}^k |\beta_j - \sigma_j^i| \le 2 \sum_{i=1}^k |\mu_j - \sigma_j^i|.$$

(3) 证明：$d(\sigma', \Sigma) \le 3d(\sigma^*, \Sigma)$ 且 $d(\sigma'', \Sigma) \le 5d(\sigma^*, \Sigma)$。

Solution:

首先重申两个定义

$$d(\boldsymbol{\sigma},\boldsymbol{\Sigma})=\sum_{i=1}^{k}L_1(\boldsymbol{\sigma},\boldsymbol{\sigma_i}),L_1(\boldsymbol{\sigma},\boldsymbol{\sigma_i})=\sum_{j=1}^{n}|\sigma_j^i-\sigma| \tag{1.1}$$

(1) 由 $\min_\mu\sum_{i=1}^k\sum_{j=1}^n|\mu_j-\sigma_j^i|$ 出发我们考虑对每一个维度实现最小化对给定的 j，问题转为求

$$\min_{\{x_1,\cdots,x_k\},\delta}\sum_{i=1}^{k}|\sigma-x_i| \tag{1.2}$$

$(\mu_j=\delta,\sigma_j^i=x_i)$ 经典结论为 μ_j 取中位数最优，证明省略

对于 μ 的每一维按得分从小到大的顺序对作品进行排序（得分相同的作品之间的顺序可任意确定），由此给出一种综合排序 σ'

但是我们无法获取 $\sigma*$ ，在存在 $\mu_i=\mu_j(i\neq j)$ 时，我们无法找到最优的排序方式。

(2)

[法一]

一个一步到位的方法

$$\begin{aligned}
d(\beta,\Sigma)-d(\mu,\Sigma)&=\sum_{i=1}^{k}(|\beta_j-\sigma_j^i|-|\mu_j-\sigma_j^i|) &(1.3)\\
&\leqslant\sum_{i=1}^{k}|(\beta_j-\sigma_j^i)-(\mu_j-\sigma_j^i)| &(1.4)\\
&=\sum_{i=1}^{k}|\mu_j-\beta_j| &(1.5)\\
&=|k\mu_j-\sum_{i=1}^{k}\sigma_j^i| &(1.6)\\
& &(1.7)\\
d(\mu,\Sigma)&=\sum_{i=1}^{k}|\mu_j-\sigma_j^i| &(1.8)\\
&\geqslant|\sum_{i=1}^{k}(\mu_j-\sigma_j^i)| &(1.9)\\
&=|k\mu_j-\sum_{i=1}^{k}\sigma_j^i| &(1.10)
\end{aligned}$$

由上式

$$d(\beta,\Sigma)-d(\mu,\Sigma)\leqslant|k\mu_j-\sum_{i=1}^{k}\sigma_j^i|\leqslant d(\mu,\Sigma)$$

则

$$\boxed{d(\beta,\Sigma) \leqslant 2d(\mu,\Sigma)} \tag{1.11}$$

成立

[**法二**]

PS：在本题中采取向量的大小关系为所有维的大小关系，为偏序关系而非全序关系！

当 $\mu \geqslant \beta$ 即中位数大于等于平均数时

不失一般性，我们记 $\sigma_j^1 \geqslant \cdots \sigma_j^k$

则 $\exists m \in [1,k], s.t. \beta_j - \sigma_j^1 \geqslant \beta_j - \sigma_j^2 \cdots 0 \geqslant \cdots \beta_j - \sigma_j^k$ 故

$$d(\beta,\Sigma) = \sum_{i=1}^{k} |\beta_j - \sigma_j^i| \tag{1.12}$$

$$= 2\sum_{i=1}^{m} (\beta_j - \sigma_j^i) \tag{1.13}$$

$$\leqslant 2\sum_{i=1}^{m} (\mu_j - \sigma_j^i) \tag{1.14}$$

$$\leqslant 2\sum_{i=1}^{k} |\mu_j - \sigma_j^i| = 2d(\mu,\Sigma) \tag{1.15}$$

剩下的情况是 $\mu < \beta$ 即中位数小于平均数：

首先我们考虑证明对给定的 j，

$$\boxed{\sum_{i=1}^{k} |\mu - \beta| \leqslant \sum_{i=1}^{k} |\mu - \sigma_j^i|} \tag{1.16}$$

，

则

$$\begin{aligned}
\text{LHS} &= \sum_{i=1}^{k} (\beta - \mu) \\
&= k \cdot \beta - k \cdot \mu \\
&= \sum_{i=1}^{k} \sigma_i - \sum_{i=1}^{k} \mu \\
&= \sum_{i=1}^{k} (\sigma_i - \mu) \\
&\leqslant \sum_{i=1}^{k} |\sigma_i - \mu| \\
&= \text{RHS}
\end{aligned}$$

故在 $\mu < \beta$ 情况下

$$\forall j \in [1,n], \sum_{i=1}^{k} |\mu - \beta| \leqslant \sum_{i=1}^{k} |\mu - \sigma_j^i|$$

扩展到多维定义下则

$$\boxed{\sum_{i=1}^{k} L_1(\mu,\beta) \leqslant \sum_{i=1}^{k} L_1(\mu,\sigma_i)} \tag{1.17}$$

因此对于所证目标，我们考虑对左侧利用三角不等式放缩

$$\begin{aligned}\sum_{i=1}^{k}|\beta-\sigma_j^i| &= \sum_{i=1}^{k}|\sigma_j^i-\beta| &(1.18)\\ &\leqslant \sum_{i=1}^{k}|\sigma_j^i-\mu|+|\mu-\beta| &(1.19)\end{aligned}$$

带入 $\sum_{i=1}^{k}|\mu-\beta| \leqslant \sum_{i=1}^{k}|\mu-\sigma_j^i|$
我们有

$$\boxed{\sum_{i=1}^{k}|\beta_j-\sigma_j^i| \leq 2\sum_{i=1}^{k}|\mu_j-\sigma_j^i|.} \tag{1.20}$$

由于对于所有 j 均成立，我们可以扩展

$$\boxed{d(\beta,\Sigma) \leqslant 2\cdot d(\mu,\Sigma)} \tag{1.21}$$

证毕
(3) 一个连续放缩的过程

$$\begin{aligned}d(\sigma',\Sigma) &= \sum_{i=1}^{k} L_1(\sigma',\sigma_i) &(1.22)\\ &\leqslant \sum_{i=1}^{k} L_1(\sigma',\mu)+L_1(\mu,\sigma_i) &(1.23)\\ &\leqslant \sum_{i=1}^{k} L_1(\sigma*,\mu)+L_1(\mu,\sigma_i) &(1.24)\\ &\leqslant \sum_{i=1}^{k} L_1(\sigma*,\sigma_i)+L_1(\mu,\sigma_i)+L_1(\mu,\sigma_i) &(1.25)\\ &\leqslant 3\sum_{i=1}^{k} L_1(\sigma*,\sigma_i) &(1.26)\end{aligned}$$

即

$$\boxed{d(\sigma',\Sigma) \leq 3d(\sigma^*,\Sigma)} \tag{1.27}$$

(1.14) 是定义导出的

从

$$d(\beta,\Sigma) \leqslant 2d(\mu,\Sigma) \tag{1.28}$$

出发

我们有

$$\sum_{i=1}^{k} L_1(\beta,\sigma_i) \leqslant 2\sum_{i=1}^{k} L_1(\mu,\sigma_i) \tag{1.29}$$

$$\Rightarrow d(\sigma'',\Sigma) = \sum_{i=1}^{k} L_1(\sigma'',\sigma_i) \tag{1.30}$$

$$\sum_{i=1}^{k} L_1(\sigma'',\beta) + L_1(\beta,\sigma_i) \tag{1.31}$$

$$\leqslant \sum_{i=1}^{k} L_1(\sigma*,\beta) + L_1(\beta,\sigma_i) \tag{1.32}$$

$$\leqslant \sum_{i=1}^{k} L_1(\sigma*,\sigma_i) + L_1(\sigma_i,\beta) + L_1(\beta,\sigma_i) \tag{1.33}$$

$$\leqslant 5\sum_{i=1}^{k} L_1(\sigma*,\sigma_i) \tag{1.34}$$

即

$$d(\sigma'',\Sigma) \leq 5d(\sigma^*,\Sigma) \tag{1.35}$$

Problem2

n 支球队进行比赛，每场比赛在两支球队之间进行，任意两支球队之间至多进行一场比赛，每支球队参与比赛的场数相同。记队 i 与队 j 比赛中，队 i 的得分为 p_{ij} ，队 j 的得分为 p_{ji}，队 i 的分差为 $q_{ij} = p_{ij} - p_{ji}$。与队 i 进行过比赛的球队集合记为 T_i。约定 $i \in T_i$，且 $q_{ii} = 0$。记 $|T_1|{=}T_2| = \cdots {=} |T_n| = l$

(1) 记 s_i 为队 i 在各场比赛中分差之和，即 $s_i = \sum_{j\in T_i} q_{ij}$ ，$\mathbf{S} = (s_1, s_2, \cdots, s_n)^T$ 称为分差向量，可用来衡量各球队的实力。若四支球队之间的比赛结果如表所示，求向量 $\mathbf{S}$: 里 $\mathbf{0}$;

(2) 对任意 $j \in T_i$，若 $k \in T_j$，则称队 i 与队 k 之间进行了一场“二级比赛”，且在该场比赛中队 i 的分差为 $q_{ij} + q_{jk}$。(队 i 可与自身进行二级比赛，队 i 与队 j 之间可以进行多场二级比赛)。记 $s_i^{(2)}$ 为队 i 在所有可能的 l^2 场二级比赛中的分差之和，$\mathbf{S}^{(2)} = (s_1^{(2)}, s_2^{(2)}, \cdots, s_n^{(2)})^T$ 称为二级分差向量。对表中所示的比赛结果，求向量 $\mathbf{S}^{(2)}$;

(3) 定义矩阵 $\mathbf{M} = (m_{ij})_{n\times n}$，其中 $m_{ij} = \begin{cases} 1, & 若 j \in T_i \\ 0, & 其他 \end{cases}$，试给出由 M 和 S 计算 $\mathbf{S}^{(2)}$ 的公式，并说明

A-B	5-10
A-D	57-45
B-C	10-7
C-D	3-10

Table 1.1: Caption

$\mathbf{M}^2$ 中各元素的含义。

(4) 类似地，对任意整数 r，可定义 r 级比赛和 r 级分差向量 $\mathbf{S}^{(}r)$，试给出由 M 和 $\mathbf{S}$ 计算 $\mathbf{S}^{(}r)$ 的公式。

Solution:

(1)

$$\begin{cases} p_{AB} = 5 \\ p_{BA} = 10 \end{cases} \rightarrow q_{AB} = -5, q_{BA} = 5$$

于是我们有

$$s_A = q_{AB} + q_{AD} = 7 \tag{1.36}$$
$$s_B = q_{BC} + q_{BA} = 8 \tag{1.37}$$
$$s_C = q_{CB} + q_{CD} = -10 \tag{1.38}$$
$$s_D = q_{DA} + q_{DC} = -5 \tag{1.39}$$

然后我们有

$$\boldsymbol{S} = \begin{bmatrix} 7 \\ 8 \\ -10 \\ -5 \end{bmatrix}$$

(2)

关系图为

$$\begin{tikzcd} A \arrow[r, dash] \arrow[d, dash] & B \arrow[d, dash] \\ D \arrow[r, dash] & C \end{tikzcd}$$

从定义出发

$$s_A^{(2)} = (q_{AB} + q_{BC}) + (q_{AD} + q_{DC}) + (q_{AA} + q_{AB}) + (q_{AA} + q_{AD}) + (q_{AB} + q_{BB}) + (q_{AD} + q_{DD}) = 31 \tag{1.40}$$

$$\dots \tag{1.41}$$

按照上述方式，我们可以获取一个更加简洁美观的形式

$$s_A^{(2)}=(l+1)s_a+s_b+s_d=31 \tag{1.42}$$
$$s_B^{(2)}=(l+1)s_b+s_a+s_c=29 \tag{1.43}$$
$$s_C^{(2)}=(l+1)s_c+s_b+s_d=-37 \tag{1.44}$$
$$s_D^{(2)}=(l+1)s_d+s_a+s_c=-23 \tag{1.45}$$

解释一下

A 出发第二阶比赛的结果可以视作 A,B,D 出发的结果，加上 A 出发到 A,B,D 的补充共贡献。

其余类似

$$\boldsymbol{S}=\begin{bmatrix}31\\29\\-37\\-23\end{bmatrix}$$

(3)

$$\mathbf{M}=\begin{bmatrix}1&1&0&1\\1&1&1&0\\0&1&1&1\\1&0&1&1\end{bmatrix}$$

结合实际含义，我们有

$$\boldsymbol{S}^{\mathbf{(2)}}=(\mathbf{M}+l\cdot\mathbf{E})\boldsymbol{S}$$

计算得到

$$\mathbf{M^2}=\begin{bmatrix}3&2&2&2\\2&3&3&2\\2&2&3&2\\2&2&2&3\end{bmatrix}$$

其中 $m_{i,j}$ 是两个队 2 级比赛轮数

更一般的 $(\mathbf{M}^r)_{i,j}$ 表示在 r 级比赛意义下 i 与 j 比赛次数

(4) 首先我们可以分析得到, 一个位置出发的 r 级比赛可以由其一次到达的位置的 $r-1$ 级比赛，接上足够数量的一级比赛获取。

容易证明

$$\boldsymbol{S}^{(r+1)}=\mathbf{M}\boldsymbol{S}^{(r)}+l^r\boldsymbol{S} \tag{1.46}$$

可以解得

$$\boldsymbol{S}^{(r+1)}=\left(\mathbf{M}^r+l\cdot\mathbf{M}^{r-1}+\cdots+l^{r-1}\cdot\mathbf{M}+l^r\cdot\mathbf{E}\right)\boldsymbol{S} \tag{1.47}$$

注意到 $\mathbf{M}-l\cdot\mathbf{E}$ 不可逆，这个等比数列求和结果不能化简

即

$$\boldsymbol{S}^{(r)}=\left(\mathbf{M}^{r-1}+l\cdot\mathbf{M}^{r-2}+\cdots+l^{r-2}\cdot\mathbf{M}+l^{r-1}\cdot\mathbf{E}\right)\boldsymbol{S} \tag{1.48}$$