

的研究当 I-O 特性中负轴部分具有截断型曲线的影响;也有答卷讨论 n 个输入信号的交调分析等等。

在稳定性分析中比较多的是讨论多项式拟合系数 ($b_0 = 0$ 或 $\neq 0$), 阶数(引入高于 3 阶)对解的影响。也有的研究幅度 A_i 对选频的影响, 有的答卷指出本题若 $A_i \equiv A$ (工程上是有意义的)则 $A > A_0$ (某个值)下本题无解, 等等。

以上各点反映出大学生在建模知识、能力上的进步。

最后本文对答卷中反映出来的一些问题提几点建议:

1. 本竞赛属数学建模竞赛, 因而在充分发挥计算机能力的同时, 仍应加强数学能力上的训练, 应注意引导学生将学过的数学知识多方面地运用于解决实际问题。随着计算软件的普及, 有可能存在削弱数学分析能力的训练而依赖于高级软件, 这对于学生的培养是不利的。例如本题的答卷中就有不加数学分析对数百种可能的选择和上万次的判断(频率约束)完全交付计算机去做; 答卷中也出现 $10\log_{10} x = x$ 计算 SNR 的错误。此外; 许多用规划解题, 目标函数没有给出很好的数学表达等等。

2. 应鼓励多学科学生组队。有许多工程类的学生对本题背景的理解和模型的改进非常好, 但数学能力比较弱; 而一些理科(尤其数学)组解答虽正确但答卷比较平淡, 有工程意义的改进推广不太多, 如有的讨论改动 $[f_{\min}, f_{\max}]$, 这在工程设计上一般是不许任意改动的。此外, 这种结合也对培养学生多科协作的思维和训练十分有益。

一个给足球队排名次的方法——B 题

戚立峰 毛 威 马 斌

指导教师 樊启洪

(北京大学数学系, 100871)

摘要 本文利用层次分析法建立了一个为足球队排名次的数学模型。它首先对用来排名次的数据是否充分作出判断, 在能够排名次时对数据的可依赖程度作出估计, 然后给出名次。文中证明了这个名次正是比赛成绩所体现的各队实力的顺序。

文中将看到此模型充分考虑了排名结果对各场比赛成绩的重要性的反馈影响, 基本上消除了由于比赛对手的强弱不同造成的不公平现象。文中还证明了模型的稳定性, 这保证了各队在发挥水平上的小的波动不会对排名顺序造成大的变动。本模型比较完满地解决了足球队排名次问题, 而且经过简单修改, 它可以适用于任何一种对抗型比赛的排名。

§ 1. 问题的提出及分析

表 1 (见第 73 页的表)给出的是我国 12 支足球队在 1988—1989 年全国甲级队联赛中的成绩, 要求通过建立数学模型, 对各队进行排名次。

按照通常的理解, 排名的目的是根据比赛成绩排出反映各队真实实力状况的一个顺序。为达到这一点, 一个好的排名算法应满足下面一些基本要求:

(1) 保序性; (2) 稳定性; (3) 能够处理不同场比赛的权重; (4) 能够判断成绩表的可约性; (5) 能够准确地进行补残; (6) 容忍不一致现象; (7) 对数据可依赖程度给出较为精确的描述。

可以想象,各队的真实实力水平在成绩表中反映出来(见§3假定II),所以根据排名目的,我们要求排名顺序与成绩表所反映的各队实力水平的顺序是一致的,这就是要求(1)。

也就是说,如果 a 比 b 表现出色, a 的名次就应排在 b 前面。但 a 比 b 出色不能只由 a 对 b 的这一场比赛所决定,必须参考 a, b 相对于其他队的成绩,象 a 平 c, c 胜 d, d 平 b 这组比赛对 a, b 的相对表现是有影响的。为使一个算法满足保序性,就必须充分考虑到将 a, b 连结起来的所有场比赛。下面的例子表明积分法不满足保序性。

例1. a 平 c, c 胜 d, d 平 b, a 平 b 。

在上述比赛中 a 表现应比 b 出色,但按积分法计算 a, b 都积2分。其原因就在于积分法没有把 a 平 c, c 胜 d, d 平 b 这组比赛中所体现的 a, b 实力对比情况考虑进去;

要求(2)是说成绩表的变动不会对排名结果造成巨大影响。这是由于球队发挥水平存在正常波动而必须提出的,如果这种正常的小波动引起名次的巨大变化,那么排名就不令人信服;

要求(3)使得不同场比赛在排名中的地位不同,这是因为在实际比赛中,往往会有有的队不幸遇到较强的队而输掉。为了避免由于对手的强弱不同造成的不公平,要求(3)是必须的。但现行的排名制度大都满足不了要求(3),以致于许多时候“运气”对名次起了重要作用;

要求(4)一(7)是为适应实际比赛中可能会出现的一些复杂情况而提出的。

首先是可能某两个队之间没有打比赛,我们称之为数据(成绩)残缺。对于两队成绩残缺,只能通过它们同其他队的比赛成绩来判断它们的实力对比。如果残缺元素过多,就有可能导致参赛队分成两组,组与组之间没有比赛,称这种情况为成绩表可约,这时显然是不应该排名次的。这样就有要求(4),(5);

其次是前后比赛成绩矛盾,比如说 a 胜 b, b 胜 c, c 平 a ,称这种情况为数据不一致。如果不一致情况过于严重,说明比赛偶然因素太大,数据的可依赖程度太低,应该考虑放弃比赛成绩。所以排名算法还应满足(6),(7)。

本文使用的层次分析法的特征根方法已满足了上述要求,下面将在§2中给出具体算法,§3中给出算法满足上述要求的解释和论证。

§2. 模型设计及其算法

一、基本假设和名词约定

假设 I. 参赛各队存在客观的真实实力(见名词约定1)。这是任何一种排名算法的基础。

假设 II. 在每场比赛中体现出来的强队对弱队的表面实力对比是以它们的真实实力对比为中心的互相独立的正态分布。(见名词约定2)

这条假设保证了我们可以以比赛成绩为依据对球队的真实实力进行排名,另外它在很大程度上反映了球队水平发挥的不稳定性。

名词约定

1. 称 $w = (w_1, w_2, \dots, w_n)$ 为真实实力向量,如果 w_i 的大小表现了 T_i 的实力强

弱. 当 w_i 的大小表现了 T_i 在比赛中出色程度时, 称 w 为排名向量. 由假设 II, 两者应是近似相同的, 以后就把它们当成同一个.

2. 称 T_i 对 T_j 这场比赛中体现出来的 T_i 对 T_j 的相对强弱程度为 T_i 对 T_j 的表面实力对比, 一般记作 a_{ij} , 当 T_i 与 T_j 成绩残缺时约定 $a_{ij} = 0$. 显然地有

$$(i) a_{ij} \geq 0, (ii) a_{ji} = 1/a_{ij}, (iii) a_{ii} = 1. \quad (2.1)$$

矩阵 $A = (a_{ij})_{n \times n}$ 就称为比赛成绩的判断矩阵, 它是可以通过各种方法(见 §5)从比赛成绩中求出来的.

由假设 II, 若 T_i 对 T_j 成绩不残缺且 $w_i/w_j \geq 1$ 时有

$$a_{ij} \sim N(w_i/w_j, \sigma_{ij}^2) \quad (2.2)$$

这里 w 是真实实力向量.

3. 称方阵 $A_{n \times n}$ 为正互反对称的, 若 (1) $a_{ij} > 0$, (2) $a_{ji} = \frac{1}{a_{ij}}, 1 \leq i, j \leq n$. 显然一个无残缺的比赛成绩的判断矩阵是正互反对称的.

4. 称矩阵 $A_{n \times n}$ 是可约的, 若 A 能用行列同时调换化为 $\begin{pmatrix} A_1 & 0 \\ A_2 & A_4 \end{pmatrix}$, 这里 A_1, A_4 都是方阵. 在[1]的227页证明了一个判断矩阵可约当且仅当成绩表可约.

5. 称判断矩阵 A 是一致的, 若对任意 $1 \leq i, k, j \leq n$ 满足 $a_{ij} \cdot a_{jk} = a_{ik}$. 显然地, A 一致则存在 w , 使得

$$A = (w_i/w_j)_{n \times n} \quad (2.3)$$

6. 称矩阵 A 的最大正特征根 λ_{\max} 为主特征根; 对应于 λ_{\max} 的右特征向量 w 称为主特征向量, 若 $\sum_{i=1}^n w_i = 1$ 且 $w_i > 0$.

由非负矩阵的 Perron-Frobenius 定理, 一个判断矩阵 A 的 λ_{\max} 存在唯一且可以让对应于 λ_{\max} 的特征向量 $w^{(1)}$ 的每个分量都大于零, 令 $w = w^{(1)} / \sum_{i=1}^n w_i^{(1)}$ 即得主特征向量.

二、模型的设计与算法

我们的模型的主要部分是一个算法, 模型的输入是一张成绩表, 输出是关于是否可约的判断、数据可依赖程度值和排名次的结果.

算法

(一) 根据比赛成绩表构造判断矩阵 A .

i 从 1 到 n , j 从 1 到 n 循环.

1) 若 T_i 与 T_j 互胜场次相等, 则

1° 净胜球 = 0 时令 $a_{ij} = a_{ji} = 1$; 跳出作下一步循环;

2° T_i 净胜球多时以 T_i 净胜 T_j 一场作后续处理.

2) 若 T_i 净胜 T_j k 场且 $k > 0$, 则

1°
$$b_{ij} = \begin{cases} 2k, & 1 \leq k \leq 4; \\ 9, & k > 4. \end{cases}$$

2° $m_{ij} = T_i$ 胜 T_j 平均每场净胜球数;

$$d_{ij} = \begin{cases} 1, & m_{ij} > 2; \\ 0, & 0 \leq m_{ij} \leq 2; \\ -1, & m_{ij} < 0. \end{cases}$$

3° $a_{ij} = b_{ij} + d_{ij}, a_{ji} = 1/a_{ij}$.

3) 若 T_i 与 T_j 无比赛成绩, 则 $a_{ij} = a_{ji} = 0$.

(二) 检测 A 的可约性, 如果可约则输出可约信息后退出。

(三) 构造辅助矩阵 \tilde{A}

i 从 1 到 n , j 从 1 到 n 循环

$$\tilde{a}_{ij} = \begin{cases} a_{ij}, & i \neq j \text{ 且 } a_{ij} \neq 0; \\ m_i + 1, & i = j, \text{ 其中 } m_i \text{ 为 } A \text{ 的第 } i \text{ 行 } 0 \text{ 的个数}; \\ 0, & a_{ij} = 0. \end{cases}$$

(四) 计算 \tilde{A} 的主特征根 λ_{\max} 和主特征向量 w

1) 允许误差 ε , 任取初始正向量 $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T$, 令 $k = 0$, 计算

$$m_0 = \max_{1 \leq i \leq n} \{x_i^{(0)}\};$$

$$y^{(0)} = (y_1^{(0)}, \dots, y_n^{(0)})^T = \frac{1}{m_0} x^{(0)}.$$

2) 迭代计算

$$x^{(k+1)} = \tilde{A} y^{(k)};$$

$$m_{k+1} = \max_{1 \leq i \leq n} \{x_i^{(k+1)}\};$$

$$y^{(k+1)} = \frac{1}{m_{k+1}} x^{(k+1)};$$

$$k = k + 1;$$

$$\text{直到 } |m_{k+1} - m_k| < \varepsilon.$$

3) $\lambda_{\max} = m_k; w = y^{(k)} / \sum_{i=1}^n y_i^{(k)}.$

(五) 按 w 各分量由大到小的顺序对参赛各队排名次。

(六) 计算 $h = \sum_{\substack{w_i > w_j \\ a_{ij} \neq 0}} \left(\frac{a_{ij}}{w_i / w_j} - 1 \right)^2 + \sum_{\substack{w_i = w_j \\ a_{ij} \neq 0 \\ i > j}} \left(\frac{a_{ij}}{w_i / w_j} - 1 \right)^2;$

$$Y = \frac{n(n-1)}{2} - \sum_{i=1}^n \frac{m_i}{2}, \text{ 其中 } m_i \text{ 为 } A \text{ 的第 } i \text{ 行 } 0 \text{ 的个数}.$$

根据 $2h$ 查 χ^2 表得到可依赖程度 $\alpha = P(\chi^2 > 2h)$.

关于算法的几点说明

算法的第(一)步可以有多种不同的方法, 这在 § 5 还将讨论。

第(二)步实际上是把 A 看作有向图的邻接矩阵表示求图是否连通, 算法是标准的, 可参阅任何一本关于算法的书, 这里省略。它在可约时作的退出处理保证了以后各步处理的是一个不可约阵。

第(四)步使用的是幂法, 其整个算法收敛性和正确性的证明可参阅[1]的 103 页。

第(五)步是一个排序, 可参阅任何一本关于算法的书。

第(六)步我们举一个例子,若算出 $2h = 47.56, r = 48$,则在 χ^2 表的自由度为 48 一行找到 47.56,它所在的列的 α 值为 65% 左右。

§ 3. 算法的理论分析

一、排名的合理性和保序性要求

关于为什么无残缺的判断矩阵 A 的主特征向量就是排名向量是层次分析法中特征根法的基础,可以在[1]的 211 页找到详细证明,这里只作简单说明。先假定比赛无残缺,此时算法中 $\tilde{A} = A$ 。

先看一下 A 为一致矩阵时,由(2.3)式存在 w 使得 $A = (w_i/w_j)_{n \times n}$, 显然向量 w 就是排名向量。

$$\text{而我们有} \quad \sum_{j=1}^n (w_i/w_j) \cdot w_j = n \cdot w_i, \quad i = 1, 2, \dots, n,$$

即

$$Aw = nw \quad (3.1)$$

在[1]的 109 页证明了下述定理:

定理. n 阶正互反矩阵是一致的,当且仅当 $\lambda_{\max} = n$ 。

再由(3.1)可见 w 还是 A 的主特征向量,这样,对于一个一致矩阵 A ,求排名向量就是求 A 的主特征向量。

对于一个不一致的判断矩阵 A (注意: 无残缺),令

$$\|A\| = \sum_{1 \leq i, j \leq n} a_{ij} \quad (3.2)$$

$$w_i = \sum_{j=1}^n a_{ij} / \|A\|, 1 \leq i \leq n, \quad (3.3)$$

由于 w_i 是 A 的第 i 列元素(即 T_i 与其他队的表面实力对比)的和被 $\|A\|$ 除,可以猜测它给出了 T_i 的排序权重。

但正如问题分析中所提到的, T_i 与 T_j 的实力对比必须考虑到将 T_i 与 T_j 连结起来的所有场比赛,反应到判断矩阵 A 上就是所有 $a_{ii_1} a_{i_1 i_2} \cdots a_{i_{k-1} j}$ 都要考虑进去。

令 $a_{ij}^{(k)}$ 是 A^k 的第 i 行 j 列元素,不难看出

$$a_{ij}^{(k)} = \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_{k-1}=1}^n a_{ii_1} a_{i_1 i_2} \cdots a_{i_{k-1} j} \quad (3.4)$$

而 $a_{ij}^{(k)}$ 就是考虑了所有经过 k 场比赛将 T_i, T_j 连结起来的路径后反映的 T_i, T_j 的相对强弱,称其为 T_i 对 T_j 的 k 步优势。

当 $i_{k-1} = j$ 时 $a_{i_{k-1} j} = 1$,所以(3.4)式成为

$$a_{ij}^{(k)} = \sum_{i_1=1}^n \cdots \sum_{\substack{i_{k-1}=1 \\ i_{k-1} \neq j}}^n a_{ii_1} \cdots a_{i_{k-1} j} + \sum_{i_1=1}^n \cdots \sum_{i_{k-2}=1}^n a_{ii_1} \cdots a_{i_{k-2} j},$$

注意到等式右端后一项正是 $a_{ij}^{(k-1)}$,所以 k 步优势就隐含了 $k-1$ 步以及 $k-2, \dots, 1$ 。

$$\text{同(3.3)式, 令 } w^{(k)} = \sum_{j=1}^n a_{ij}^{(k)} / \|A^k\|, i = 1, \dots, n,$$

再令 $w^{(k)} = (w_1^{(k)}, \dots, w_n^{(k)})^T$, 可以想象, 当 k 足够大时, $w^{(k)}$ 就给出了 A 所反映的排名向量. 在[1]的 104 页证明了等式

$$\lim_{k \rightarrow \infty} \frac{A^k e}{e^T A^k e} = w, \text{ 其中 } e = (1, 1, \dots, 1)^T, w$$

是 A 的主特征向量.

即
$$\lim_{k \rightarrow \infty} w^{(k)} = w.$$

所以在充分考虑了足够多步优势后得到的排名向量 $w^{(\infty)}$ 就是 A 的主特征向量 w .

上面的讨论表明在比赛无残缺时, 我们的排名是合理的和保序的, 下面来看残缺的情况.

二、残缺的处理

对于一个残缺的判断矩阵 A , 可以通过下述方法转化成一中讨论的情形

$$c_{ij} = \begin{cases} a_{ij}, & a_{ij} \neq 0, \\ d_{ij}, & a_{ij} = 0, \end{cases} \text{ 其中 } d_{ij} \text{ 为正数,}$$

如果这样得到的矩阵 $C = (c_{ij})_{n \times n}$ 的主特征向量为 w , 那么当 $d_{ij} = w_i / w_j$ 时, 我们认为补残是准确的. 如果令

$$c_{ij} = \begin{cases} a_{ij}, & a_{ij} \neq 0; \\ w_i / w_j, & a_{ij} = 0; \end{cases}$$

$$\tilde{a}_{ij} = \begin{cases} a_{ij}, & a_{ij} \neq 0, i \neq j; \\ 0, & a_{ij} = 0, i \neq j; \\ m_i + 1, & i = j, m_i \text{ 是 } A \text{ 的第 } i \text{ 行 } 0 \text{ 的个数}; \end{cases}$$

$$C = (c_{ij})_{n \times n};$$

$$\tilde{A} = (\tilde{a}_{ij})_{n \times n}.$$

则有下面命题成立:

命题. $Cw = \lambda w$ 等价于 $\tilde{A}w = \lambda w$.

证. $\sum_{j=1}^n c_{ij} w_j = \lambda w_i, i = 1, \dots, n.$

$$\Leftrightarrow \sum_{\substack{j=1 \\ a_{ij} \neq 0, i \neq j}}^n a_{ij} w_j + \sum_{\substack{j=1 \\ a_{ij} = 0}}^n (w_i / w_j) \cdot w_j + w_i = \lambda w_i, i = 1, \dots, n.$$

$$\Leftrightarrow \sum_{\substack{j=1 \\ i \neq j}}^n a_{ij} w_j + (m_i + 1) w_i = \lambda w_i, i = 1, \dots, n.$$

$$\Leftrightarrow \sum_{j=1}^n \tilde{a}_{ij} w_j = \lambda w_i, i = 1, \dots, n.$$

证毕.

由上述命题还可知, C 的最大正特征根也是 \tilde{A} 的主特征根, C 的主特征向量也是 A 的主特征向量. 这样, 我们只需解 $\tilde{A}w = \lambda_{\max} w$ 即可, 这正是算法(三)、(四)步作的工作.

从上面讨论可知, 本模型对于残缺的处理是非常准确的, 满足了要求(1),(5). 另外算法的第(二)步对成绩表的可约性作出了判断, 这也满足了因为残缺而提出的要求(4).

下面继续讨论其余四个要求

三、 对手的强弱对自己名次的影响

排名向量满足 $\tilde{A}w = \lambda_{\max}w$, 即

$$w_i = \frac{1}{\lambda_{\max}} \sum_{j=1}^n \tilde{a}_{ij} w_j, i = 1, 2, \dots, n.$$

如果 T_i 对 T_k 成绩不残缺, 则 $\tilde{a}_{ik} = a_{ik} > 0$, 固定 a_{ik} , 令 w_k 变大, 则 $\tilde{a}_{ik}w_k$ 就会变大, 从而引起 w_i 变大. 这实际上是排名结果对每场比赛权重的反馈影响.

这样的话, 若 T_i 对 T_j 战绩固定, T_j 排名靠前, T_i 也会因此受益. 这就满足了要求(3).

四、 模型稳定性的分析

不加证明地引用下面定理 ([1]103 页)

定理. 设 A 为 $n \times n$ 复矩阵, λ_1 是 A 的单特征根, B 是 $n \times n$ 矩阵, 则一定可以从 $A + B_\epsilon$ (其中 $|\epsilon|$ 足够小) 的特征根中找到一个特征根 $\tilde{\lambda}$ 满足 $\tilde{\lambda} = \lambda_1 + O(\epsilon)$.

由名词约定 6 中解释 \tilde{A} 的最大正特征根是单的, 由上述定理可知, 只要判断矩阵的变动微小, 主特征根的变动就是微小的, 进一步容易证明线性方程组 $(\tilde{A} - \lambda_{\max}E)w = 0$ 的满足 $\sum_{i=1}^n w_i = 1$ 的解的变动是微小的, 即主特征向量的变动是微小的, 排名是稳定的, 满足了要求(2).

五、 关于可依赖程度的分析

很明显本模型是容忍不一致现象的, 即满足要求(6).

当 A 是一个残缺的不一致矩阵时, 由它得到的排名向量设为 w , 由名词约定(1)我们认为这就是真实实力向量, 令

$$\delta_{ij} = \frac{a_{ij}}{w_i/w_j} - 1, i, j = 1, \dots, n. \quad (3.5)$$

则由(2.2)式知 $w_i/w_j \geq 1$ 时,

$$\delta_{ij} = \frac{a_{ij} - w_i/w_j}{w_i/w_j} \sim N\left(0, \frac{\sigma_{ij}^2}{w_i/w_j}\right). \quad (3.6)$$

为计算方便, 我们进一步假定 $w_i/w_j \geq 1$ 时,

$$\frac{\sigma_{ij}^2}{w_i/w_j} = \sigma^2 \text{ 为常数,} \quad (3.7)$$

令

$$h = \sum_{\substack{w_i/w_j > 1 \\ a_{ij} \neq 0}} \delta_{ij}^2 + \sum_{\substack{w_i/w_j = 1 \\ a_{ij} \neq 0, i > j}} \delta_{ij}^2. \quad (3.8)$$

则 h 可看作 A 的前后矛盾程度, 再由(3.6), (3.7)可知

$$h/\sigma^2 \sim \chi_\gamma^2, \quad (3.9)$$

其中

$$\gamma = \frac{n(n-1)}{2} - \sum_{i=1}^n \frac{m_i}{2}, \quad (3.10)$$

m_i 为第 i 行零的个数.

那么对某个固定 A_0 , 可以通过(3.10)求出 r_0 , 通过(3.8)求出 h_0 , 设随机变量 $h/\sigma^2 \sim \chi^2_{r_0}$, 则查 χ^2 表可得到

$$\alpha = P\left(\frac{h}{\sigma^2} > \frac{h_0}{\sigma^2}\right) \quad (3.11)$$

称 α 为 A_0 的可依赖程度. 则一个判断矩阵 A_0 的可依赖程度为 α 就表示, 如果与 A_0 相同的几个队在同样的比赛程序(队编号相同, 残缺元素相同)下踢大量赛季的比赛(假定各队水平不长进), 判断矩阵为 A_0 的这次的前后矛盾程度 h_0 比大约 $\alpha \times 100\%$ 的赛季的比赛前后矛盾程度 h 要小.

σ^2 的值可以用统计的方法估出, 在本模型中我们只是简单地取 $\sigma^2 = \frac{1}{2}$.

α 临界值的确定可以很灵活地由比赛组织者决定, 也可以通过大量好的和坏的比赛成绩比较给出一个值.

这样, 我们的模型就满足了要求(7).

§ 4. 模型运行结果的分析

我们在计算机上实现了上述模型, 并对表 1 中的数据进行了排名, 结果是令人满意的, 运算时间小于 1 秒, 得到的结果是:

排名顺序(由强到弱): $T_7, T_3, T_1, T_9, T_2, T_{10}, T_8, T_{12}, T_6, T_5, T_{11}, T_4$.

数据可依赖程度为 65%;

T_7 踢了 9 场比赛, 全部获胜, T_4 踢了 9 场比赛全部输掉, 所以 T_7 第一而 T_4 最末是显然的. 下面考虑一对水平接近的队 T_1 和 T_3 .

在 T_1, T_3 与其他队的比赛中, 只有与 T_9, T_4, T_5 的比赛中, T_1 成绩比 T_3 稍好, 而在与其余 6 个队的比赛中, T_3 成绩都优于 T_1 , 而且在 T_1 与 T_3 比赛时 T_3 在净胜球方面占了上风, 因此, 将 T_3 排在 T_1 前面是合适的.

数据可依赖程度为 65% 说明表 1 中所给数据还是不错的, 当然由于算法中取 $\sigma^2 = \frac{1}{2}$ 是先验的, 这个指标暂时还不是准确的.

§ 5. 模型优缺点及改进方向

通过与现行的一些排名方法的比较, 上述模型的优势是很明显的:

- 1) 它存在反馈机制, 并且具有稳定性, 保证了排名的公平和令人信服;
- 2) 能较准确地处理残缺, 不一致等性质很差的数据, 对比赛程序没有严格的要求;
- 3) 灵活机动, 这包括了它提供了对比赛成绩表进行取舍的参考指标, 以及它适合任意 N 个队任何对抗型比赛的排名;
- 4) 满足保序性.

模型主要的一个缺点就是算法复杂, 必须用到计算机, 而且对指导教师制定战略造成了困难, 这是无法改进的, 但这同时也使球队的战术水平在比赛中的地位上升, 有利于刺

激竞争。另外我们还基于另一种思路建立了一个便于手算的模型,由于算法简单,效果没有本模型好,本文中省略。

在从成绩表构造判断矩阵时用到的方法也不是最好的,它只是为了简单和较合乎常识,这一步在整个模型里引入的误差最大。稍微复杂一点的方法是根据成绩通过查表或专家咨询获得实力对比的值。

另外一个不足之处是在某些残缺元素过多的情况下排名的稳定性和可靠性较低,而可依赖程度这个指标并没有考虑这些情况。如比较下面两个判断矩阵,它们的差别就不大。

$$\begin{pmatrix} 1 & 1 & 0 & \frac{1}{2} \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix} \text{ 与 } \begin{pmatrix} 1 & 1 & 0 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ \frac{1}{2} & 0 & 1 & 1 \end{pmatrix}.$$

但排名结果分别为 T_4, T_3, T_2, T_1 和 T_2, T_1, T_3, T_4 . 结构变化很大. 这种情况可以也只能对比赛程序作一些要求,以避免这种几乎可约的情形,本模型并没有作这种工作。

还有就是象 § 4 所说的,可依赖程度的计算中取 $\sigma^2 = \frac{1}{2}$ 是没有多少道理的,这可以通过用统计的方法估出 σ^2 来解决。

不基于本模型的不足,模型的改进余地也是很大的。它只使用了层次分析法中单一准则一个层次的排序方法,可以考虑使用多个准则和递阶层次,比如将净胜局数,净胜球数,射门次数,犯规次数作成四个准则,两个层次。甚至能将观众反应等许多细小因素考虑在内,使排名更加反应球队实力。

参 考 文 献

- [1] 王莲芬,许树柏,层次分析法引论,中国人民大学出版社,北京,1990.
- [2] 叶其孝等编,大学生数学建模竞赛辅导教材,湖南教育出版社,1993.
- [3] 许卓群等编,数据结构,高等教育出版社,1987.
- [4] 杜荣骞编,生物统计学,高等教育出版社,1985.