详细解答

STEP III 1996，第 12 题

概率与统计 • 马可夫链

STEP III 1996，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/96-s3-q12

本题在 *Beyond STEP* 第 12.16 节*顺向与逆向的马可夫链*中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch12/markov.html

关键思路

这是一条马可夫链（Markov chain）：下一条定理的类型只取决于当前定理的类型。第 (i) 部分是全概率公式（law of total probability），以当前定理的类型作为条件。第 (ii) 部分要求找出一个在这条链作用下保持不变的分布（即*平稳分布*，stationary distribution），这归结为一个小型线性方程组。第 (iii) 部分把这条链倒过来走，而把条件概率（conditional probability）反转，正是贝叶斯定理（Bayes' theorem）的作用。

解答

把定理依次编号。设 X 为某一条定理的类型，Y 为紧接其后那条定理的类型，于是

$$\mathrm{P}(X = i) = a_i, \qquad \mathrm{P}(Y = j \mid X = i) = p_{ij}, \qquad \mathrm{P}(Y = j) = b_j.$$

第 (i) 部分

事件 $\{X = 1\}$、$\{X = 2\}$、$\{X = 3\}$ 互斥（mutually exclusive），而且其中必有一个发生，因为每条定理恰好属于三种类型之一。因此它们构成样本空间（sample space）的一个划分，而事件 $\{Y = j\}$ 可分拆成三个互不相交的部分：

$$\{Y = j\} = \{X = 1,\, Y = j\} \cup \{X = 2,\, Y = j\} \cup \{X = 3,\, Y = j\}.$$

把各部分的概率相加，并把每一项写成 $\mathrm{P}(X=i)\,\mathrm{P}(Y=j \mid X=i)$，得

$$b_j = \sum_{i=1}^{3} \mathrm{P}(X=i)\,\mathrm{P}(Y=j \mid X=i) = a_1 p_{1j} + a_2 p_{2j} + a_3 p_{3j},$$

这就是全概率公式。题目给出的观察在这里用上了：不论当前定理之前发生过甚么，$\mathrm{P}(Y=j \mid X=i)$ 都等于 p_{ij}。

第 (ii) 部分

我们需要对每个 j 都有 $b_j = a_j$，同时 $a_1 + a_2 + a_3 = 1$，因为各 a_i 构成一个概率分布。沿着数表的各列由上而下读，得到三条方程

$$\begin{aligned} a_1 &= 0.3a_1 + 0.2a_2 + 0.1a_3, \\ a_2 &= 0.3a_1 + 0.4a_2 + 0.3a_3, \\ a_3 &= 0.4a_1 + 0.4a_2 + 0.6a_3. \end{aligned}$$

> **想一想** 三个未知数却有三条方程，再加上 $a_1 + a_2 + a_3 = 1$：这不是多了一条方程吗？

并没有多，因为这三条方程并非互相独立。由于数表每一行之和都是 1，把三条方程相加便得 $a_1 + a_2 + a_3 = a_1 + a_2 + a_3$。所以其中任意两条都蕴含第三条，而真正把解确定下来的，是归一化条件（normalising condition）。

中间那条方程最容易处理，因为 a_1 和 a_3 在其中的系数相同。利用 $a_1 + a_3 = 1 - a_2$，

$$a_2 = 0.3(1-a_2) + 0.4a_2 \quad\Longrightarrow\quad 0.9a_2 = 0.3 \quad\Longrightarrow\quad a_2 = \tfrac{1}{3}.$$

于是 $a_1 + a_3 = \frac{2}{3}$。把第一条方程乘以 10，得 $7a_1 = 2a_2 + a_3 = \frac{2}{3} + a_3$，再代入 $a_3 = \frac{2}{3} - a_1$，

$$7a_1 = \tfrac{4}{3} - a_1 \quad\Longrightarrow\quad a_1 = \tfrac{1}{6}, \qquad a_3 = \tfrac{2}{3} - \tfrac{1}{6} = \tfrac{1}{2}.$$

验算第三条方程：$0.4 \cdot \frac{1}{6} + 0.4 \cdot \frac{1}{3} + 0.6 \cdot \frac{1}{2} = \frac{1}{15} + \frac{2}{15} + \frac{3}{10} = \frac{1}{2} = a_3$。因此

$$a_1 = \tfrac{1}{6}, \qquad a_2 = \tfrac{1}{3}, \qquad a_3 = \tfrac{1}{2}.$$

由此可见，长远而言，Ecks 教授的定理有一半是错的。

第 (iii) 部分

现在改设 X 为前一条定理的类型，Y 为指定那条定理的类型，而 X 的分布如第 (ii) 部分所得。我们要求的是 $q_{ij} = \mathrm{P}(X=i \mid Y=j)$。根据贝叶斯定理，

$$q_{ij} = \frac{\mathrm{P}(X=i)\,\mathrm{P}(Y=j \mid X=i)}{\mathrm{P}(Y=j)} = \frac{a_i p_{ij}}{b_j} = \frac{a_i p_{ij}}{a_j},$$

其中最后一步用了第 (ii) 部分的 $b_j = a_j$。这正是题目指定 a_i 取这些值的原因：这样一来，指定那条定理的分布便与它前一条定理的分布相同。

逐列计算：

- $j = 1$（$a_1 = \frac{1}{6}$）：$q_{11} = \dfrac{\frac{1}{6}\cdot 0.3}{\frac{1}{6}} = \frac{3}{10}$，$q_{21} = \dfrac{\frac{1}{3}\cdot 0.2}{\frac{1}{6}} = \frac{2}{5}$，$q_{31} = \dfrac{\frac{1}{2}\cdot 0.1}{\frac{1}{6}} = \frac{3}{10}$。
- $j = 2$（$a_2 = \frac{1}{3}$）：$q_{12} = \dfrac{\frac{1}{6}\cdot 0.3}{\frac{1}{3}} = \frac{3}{20}$，$q_{22} = 0.4 = \frac{2}{5}$，$q_{32} = \dfrac{\frac{1}{2}\cdot 0.3}{\frac{1}{3}} = \frac{9}{20}$。
- $j = 3$（$a_3 = \frac{1}{2}$）：$q_{13} = \dfrac{\frac{1}{6}\cdot 0.4}{\frac{1}{2}} = \frac{2}{15}$，$q_{23} = \dfrac{\frac{1}{3}\cdot 0.4}{\frac{1}{2}} = \frac{4}{15}$，$q_{33} = 0.6 = \frac{3}{5}$。

按原来数表的方式排列（第 i 行、第 j 列）：

$$(q_{ij}) = \begin{pmatrix} \frac{3}{10} & \frac{3}{20} & \frac{2}{15} \\ \frac{2}{5} & \frac{2}{5} & \frac{4}{15} \\ \frac{3}{10} & \frac{9}{20} & \frac{3}{5} \end{pmatrix} = \begin{pmatrix} 0.3 & 0.15 & 0.1\dot{3} \\ 0.4 & 0.4 & 0.2\dot{6} \\ 0.3 & 0.45 & 0.6 \end{pmatrix}.$$

这里和为 1 的是各列：指定那条定理的类型 j 固定之后，前一条定理必定属于某一种类型。

验算

- (q_{ij}) 每一列之和都是 1；例如 $\frac{2}{15} + \frac{4}{15} + \frac{3}{5} = 1$。
- 对角线上的元素满足 $q_{ii} = p_{ii}$，与公式 $q_{ii} = a_i p_{ii} / a_i$ 所预示的一致。
- 把这条链模拟 10^6 步，得到的长期频率为 0.167、0.333、0.501，而倒推所得的比例与上面每个 q_{ij} 的差距都在 0.001 之内。

常见错误

- 把数表乘以由各 a_i 组成的列向量（即用了行而不是列）。方程是 $b_j = \sum_i a_i p_{ij}$，所以求和是沿着第 j 列由上而下进行的。
- 忘记了 $a_1 + a_2 + a_3 = 1$，然后对三条平衡方程（balance equations）无法确定各 a_i 感到困惑。
- 以 p_{ji} 作为第 (iii) 部分的答案。时间上往回走，并不等于把数表反过来读；这条链并不可逆（reversible），因为 $q_{12} = 0.15$，而 $p_{12} = 0.3$。

STEP 之外

向量 (a_1, a_2, a_3) 是这条链的平稳分布：它是转移矩阵（transition matrix）对应特征值（eigenvalue）1 的左特征向量（left eigenvector）。由于矩阵的每个元素都是正数，不论第一条定理属于哪种类型，第 n 条定理的类型的分布都会收敛至这个平稳分布。数表 (q_{ij}) 描述的是时间反转（time-reversed）后的链；若一条链反转后与原来的看起来一样（即满足 $a_i p_{ij} = a_j p_{ji}$，称为细致平衡条件，detailed balance condition），这类链在随机漫步（random walk）和蒙地卡罗方法（Monte Carlo methods）中占有核心地位。Grimmett 与 Welsh 合著的 *Probability: An Introduction* 中有一章讲马可夫链，深入浅出。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。