詳細解答

STEP III 1996，第 12 題

概率與統計 • 馬可夫鏈

> **STEP III 1996，第 12 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/96-s3-q12

本題在 *Beyond STEP* 第 12.16 節順向與逆向的馬可夫鏈中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/markov.html

關鍵思路

這是一條馬可夫鏈（Markov chain）：下一條定理的類型只取決於當前定理的類型。第 (i) 部分是全概率公式（law of total probability），以當前定理的類型作為條件。第 (ii) 部分要求找出一個在這條鏈作用下保持不變的分佈（即平穩分佈，stationary distribution），這歸結為一個小型線性方程組。第 (iii) 部分把這條鏈倒過來走，而把條件概率（conditional probability）反轉，正是貝葉斯定理（Bayes' theorem）的作用。

解答

把定理依次編號。設 X 為某一條定理的類型，Y 為緊接其後那條定理的類型，於是

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

第 (i) 部分

事件 $\{X = 1\}$、$\{X = 2\}$、$\{X = 3\}$ 互斥（mutually exclusive），而且其中必有一個發生，因為每條定理恰好屬於三種類型之一。因此它們構成樣本空間（sample space）的一個劃分，而事件 $\{Y = j\}$ 可分拆成三個互不相交的部分：

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

把各部分的概率相加，並把每一項寫成 $\mathrm{P}(X=i)\,\mathrm{P}(Y=j \mid X=i)$，得

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

這就是全概率公式。題目給出的觀察在這裏用上了：不論當前定理之前發生過甚麼，$\mathrm{P}(Y=j \mid X=i)$ 都等於 p_{ij}。

第 (ii) 部分

我們需要對每個 j 都有 $b_j = a_j$，同時 $a_1 + a_2 + a_3 = 1$，因為各 a_i 構成一個概率分佈。沿着數表的各列由上而下讀，得到三條方程

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

> ? **想一想** 三個未知數卻有三條方程，再加上 $a_1 + a_2 + a_3 = 1$：這不是多了一條方程嗎？

並沒有多，因為這三條方程並非互相獨立。由於數表每一行之和都是 1，把三條方程相加便得 $a_1 + a_2 + a_3 = a_1 + a_2 + a_3$。所以其中任意兩條都蘊含第三條，而真正把解確定下來的，是歸一化條件（normalising condition）。

中間那條方程最容易處理，因為 a_1 和 a_3 在其中的係數相同。利用 $a_1 + a_3 = 1 - a_2$，

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

於是 $a_1 + a_3 = \frac{2}{3}$。把第一條方程乘以 10，得 $7a_1 = 2a_2 + a_3 = \frac{2}{3} + a_3$，再代入 $a_3 = \frac{2}{3} - a_1$，

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

驗算第三條方程：$0.4 \cdot \frac{1}{6} + 0.4 \cdot \frac{1}{3} + 0.6 \cdot \frac{1}{2} = \frac{1}{15} + \frac{2}{15} + \frac{3}{10} = \frac{1}{2} = a_3$。因此

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

由此可見，長遠而言，Ecks 教授的定理有一半是錯的。

第 (iii) 部分

現在改設 X 為前一條定理的類型，Y 為指定那條定理的類型，而 X 的分佈如第 (ii) 部分所得。我們要求的是 $q_{ij} = \mathrm{P}(X=i \mid Y=j)$。根據貝葉斯定理，

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

其中最後一步用了第 (ii) 部分的 $b_j = a_j$。這正是題目指定 a_i 取這些值的原因：這樣一來，指定那條定理的分佈便與它前一條定理的分佈相同。

逐列計算：

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

按原來數表的方式排列（第 i 行、第 j 列）：

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

這裏和為 1 的是各列：指定那條定理的類型 j 固定之後，前一條定理必定屬於某一種類型。

驗算

- (q_{ij}) 每一列之和都是 1；例如 $\frac{2}{15}+\frac{4}{15}+\frac{3}{5}=1$。
- 對角線上的元素滿足 $q_{ii}=p_{ii}$，與公式 $q_{ii}=a_ip_{ii}/a_i$ 所預示的一致。
- 把這條鏈模擬 10^6 步，得到的長期頻率為 0.167、0.333、0.501，而倒推所得的比例與上面每個 q_{ij} 的差距都在 0.001 之內。

常見錯誤

- 把數表乘以由各 a_i 組成的列向量（即用了行而不是列）。方程是 $b_j=\sum_i a_ip_{ij}$，所以求和是沿着第 j 列由上而下進行的。
- 忘記了 $a_1+a_2+a_3=1$，然後對三條平衡方程（balance equations）無法確定各 a_i 感到困惑。
- 以 p_{ji} 作為第 (iii) 部分的答案。時間上往回走，並不等於把數表反過來讀；這條鏈並不可逆（reversible），因為 $q_{12}=0.15$，而 $p_{12}=0.3$。

STEP 之外

向量 (a_1, a_2, a_3) 是這條鏈的平穩分佈：它是轉移矩陣（transition matrix）對應特徵值（eigenvalue）1 的左特徵向量（left eigenvector）。由於矩陣的每個元素都是正數，不論第一條定理屬於哪種類型，第 n 條定理的類型的分佈都會收斂至這個平穩分佈。數表 (q_{ij}) 描述的是時間反轉（time-reversed）後的鏈；若一條鏈反轉後與原來的看起來一樣（即滿足 $a_i p_{ij} = a_j p_{ji}$，稱為細緻平衡條件，detailed balance condition），這類鏈在隨機漫步（random walk）和蒙地卡羅方法（Monte Carlo methods）中佔有核心地位。Grimmett 與 Welsh 合著的 *Probability: An Introduction* 中有一章講馬可夫鏈，深入淺出。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。