풀이

# STEP III 1996, 문제 12

확률과 통계 • 마르코프 연쇄

**STEP III 1996, 12번 문제**

더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:

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

이 문제는 *Beyond STEP*의 12.16 절 **앞으로 보고 거꾸로 보는 마르코프 연쇄**에서 단계별 안내와 함께 자세히 다룬다: https://beyondstep.keithmcnulty.org/ko/ch12/markov.html

## 핵심 아이디어

이것은 마르코프 연쇄(Markov chain)이다. 다음 정리의 유형은 현재 정리의 유형에만 의존한다. 소문항 (i)은 현재 정리의 유형을 조건으로 하는 전확률 법칙(law of total probability)이다. 소문항 (ii)는 연쇄가 바꾸지 않는 분포, 곧 **정상** 분포(stationary distribution)를 구하라는 것이며, 이는 작은 연립일차방정식이 된다. 소문항 (iii)은 연쇄를 거꾸로 돌리는 것인데, 조건부 확률(conditional probability)을 뒤집는 일이 바로 베이즈(Bayes) 정리(Bayes' theorem)가 하는 일이다.

## 풀이

정리들에 차례로 번호를 붙인다. 주어진 한 정리의 유형을 $X$, 그 다음 정리의 유형을 $Y$라 하면

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

### 소문항 (i)

사건 $\{X = 1\}$, $\{X = 2\}$, $\{X = 3\}$은 서로 배반(mutually exclusive)이고, 모든 정리는 세 유형 가운데 정확히 하나에 속하므로 이 중 하나는 반드시 일어난다. 따라서 이 사건들은 표본공간(sample space)을 분할(partition)하며, 사건 $\{Y = j\}$는 서로소인 세 조각으로 나뉜

다.

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

각 조각의 확률을 더하고, 각각을 $\mathrm{P}(X = i)\,\mathrm{P}(Y = j \mid X = i)$로 쓰면

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

이고, 이것이 전확률 법칙이다. 여기서 주어진 관찰 사실이 쓰인다. 현재 정리 이전에 무슨 일이 있었든 $\mathrm{P}(Y = j \mid X = i)$는 $p_{ij}$이다.

**소문항 (ii)**

각 $j$에 대하여 $b_j = a_j$이어야 하고, $a_i$들은 확률분포를 이루므로 $a_1 + a_2 + a_3 = 1$도 성립해야 한다. 배열의 열을 따라 내려 읽으면 세 방정식은 다음과 같다.

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

> ? **생각해 보기** 미지수 세 개에 방정식 세 개, 거기에 $a_1 + a_2 + a_3 = 1$까지 있다. 방정식이 하나 너무 많은 것 아닌가?

그렇지 않다. 세 방정식이 독립이 아니기 때문이다. 배열의 각 행의 합이 1이므로, 세 방정식을 더하면 $a_1 + a_2 + a_3 = a_1 + a_2 + a_3$이 된다. 따라서 어느 두 방정식이든 나머지 하나를 함의하며, 해를 하나로 정해 주는 것은 정규화 조건(normalising condition)이다.

가운데 방정식이 가장 쉽다. 거기서는 $a_1$과 $a_3$의 계수가 같기 때문이다. $a_1 + a_3 = 1 - a_2$를 이용하면

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

그러면 $a_1 + a_3 = \frac{2}{3}$이다. 첫째 방정식에 10을 곱하면 $7a_1 = 2a_2 + a_3 = \frac{2}{3} + a_3$이고, $a_3 = \frac{2}{3} - a_1$을 대입하면

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

셋째 방정식을 확인하자. $0.4 \cdot \frac{1}{6} + 0.4 \cdot \frac{1}{3} + 0.6 \cdot \frac{1}{2} = \frac{1}{15} + \frac{2}{15} + \frac{3}{10} = \frac{1}{2} = a_3$이다. 따라서

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

그러므로 장기적으로 보면 Ecks 교수의 정리 가운데 절반은 거짓이다.

### 소문항 (iii)

이제 $X$를 바로 앞 정리의 유형, $Y$를 특정한 정리의 유형으로 잡고, $X$는 소문항 (ii)에서와 같은 분포를 따른다고 하자. 구하려는 것은 $q_{ij} = \mathrm{P}(X = i \mid Y = j)$이다. 베이즈 정리에 의하여

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

이며, 마지막 단계에서는 소문항 (ii)의 $b_j = a_j$를 이용하였다. 문제가 $a_i$를 이 값들로 정해 둔 이유가 바로 이것이다. 그러면 특정한 정리가 바로 앞 정리와 같은 분포를 가진다.

열별로 계산하면 다음과 같다.

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

원래 배열처럼 ($i$행, $j$열로) 늘어놓으면

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

여기서는 합이 1이 되는 것이 **열**이다. 특정한 정리의 유형 $j$를 고정하면, 바로 앞 정리는 어떤 유형이든 하나의 유형이었어야 하기 때문이다.

## 검산

- $(q_{ij})$의 각 열의 합은 1이다. 예를 들어 $\frac{2}{15} + \frac{4}{15} + \frac{3}{5} = 1$이다.
- 대각성분은 공식 $q_{ii} = a_i p_{ii} / a_i$가 예측하는 대로 $q_{ii} = p_{ii}$를 만족한다.
- 연쇄를 $10^6$단계 시뮬레이션한 결과, 장기 빈도는 0.167, 0.333, 0.501이었고, 거꾸로 센 비율은 위의 모든 $q_{ij}$와 0.001 이내로 일치하였다.

## 흔한 실수

- 배열에 $a_i$들로 이루어진 열벡터를 곱하는 것(열 대신 행을 사용하는 것). 방정식은 $b_j = \sum_i a_i p_{ij}$이므로, 합은 $j$열을 따라 내려가며 취한다.

- $a_1 + a_2 + a_3 = 1$을 잊고서, 세 균형 방정식이 $a_i$를 결정하지 못한다는 사실에 당황하는 것.
- 소문항 (iii)의 답으로 $p_{ji}$를 쓰는 것. 시간을 거꾸로 거슬러 가는 것은 배열을 반대 방향으로 읽는 것과 같지 않다. $q_{12} = 0.15$이지만 $p_{12} = 0.3$이므로 이 연쇄는 가역(reversible)이 아니다.

## STEP 너머

벡터 $(a_1, a_2, a_3)$는 이 연쇄의 정상 분포이며, 전이 행렬(transition matrix)의 고윳값(eigenvalue) 1에 대한 왼쪽 고유벡터(left eigenvector)이다. 행렬의 모든 성분이 양수이므로, 첫 정리가 어떤 유형이었든 $n$번째 정리의 유형의 분포는 이 분포로 수렴(converge)한다. 배열 $(q_{ij})$는 **시간 역전**(time-reversed) 연쇄를 나타내며, 거꾸로 돌려도 원래와 똑같아 보이는 연쇄(세부 균형 조건(detailed balance condition)인 $a_i p_{ij} = a_j p_{ji}$를 만족하는 연쇄)는 무작위 행보(random walk)와 몬테카를로 방법(Monte Carlo methods)에서 핵심적인 역할을 한다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)에 마르코프 연쇄를 쉽게 다룬 장이 있다.

Beyond STEP 풀이: https://beyondstep.keithmcnulty.org/ko/. 풀이 글은 CC0 1.0이다. 문제 자체의 저작권은 © Cambridge University Press & Assessment에 있으며 여기에 싣지 않는다. 영어 원문을 기계 번역하였으며 아직 검토되지 않았다.