MATH4240: Stochastic Processes Tutorial 2

WONG, Wing Hong

The Chinese University of Hong Kong

whwong@math.cuhk.edu.hk

28 January, 2021

Matrix diagonalization

In linear algebra, we know that an $n \times n$ matrix P is said to be *diagonalizable* if there exists an invertible $n \times n$ matrix Q such that

$$Q^{-1}PQ = \Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{1}$$

is a diagonal matrix. Write Q as $Q = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, where α_i's are the column vectors of Q, then (1) can be written as

$$P\alpha_i = \lambda_i \alpha_i, \quad i = 1, 2, \cdots, n.$$

Hence each α_i is a right eigenvector of P, and each λ_i is a corresponding eigenvalue. By the invertibility of Q, we know that P is diagonalizable if and only if P has n linearly independent eigenvectors.

Matrix diagonalization

Example. Let $\{X_n\}_{n\geq 0}$ be the two-state Markov chain (page 2 in textbook) with the state space $\mathcal{S} = \{0, 1\}$ and the transition matrix

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix},$$

where $0 \leq p, q \leq 1$ and $0 < p + q < 2$ (that is, p and q are neither both equal to 0 nor both equal to 1).
As $\det(\lambda I - P) = (\lambda - 1)(\lambda - 1 + p + q)$, P has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 1 - p - q$.
Solving linear equation $P\alpha_i = \lambda_i \alpha_i$, $i = 1, 2$, we have

$$\alpha_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \alpha_2 = \begin{pmatrix} p \\ -q \end{pmatrix}.$$

Matrix diagonalization

Let $Q = (\alpha_1, \alpha_2) = \begin{pmatrix} 1 & p \\ 1 & -q \end{pmatrix}$. Then $Q^{-1} = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{1}{p+q} & -\frac{1}{p+q} \end{pmatrix}$ and

$$Q^{-1}PQ = \begin{pmatrix} 1 & 0 \\ 0 & 1-p-q \end{pmatrix} = \Lambda.$$

Matrix diagonalization

Hence

$$P^n = (Q\Lambda Q^{-1})^n = Q \begin{pmatrix} 1 & 0 \\ 0 & (1-p-q)^n \end{pmatrix} Q^{-1}$$
$$= \begin{pmatrix} \frac{q}{p+q} + \frac{p}{p+q}(1-p-q)^n & \frac{p}{p+q} - \frac{p}{p+q}(1-p-q)^n \\ \frac{q}{p+q} - \frac{q}{p+q}(1-p-q)^n & \frac{p}{p+q} + \frac{q}{p+q}(1-p-q)^n \end{pmatrix}$$

and

$$\lim_{k\to\infty} P^k = \lim_{k\to\infty} (Q\Lambda Q^{-1})^k = Q \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix}.$$

A remark on two-state Markov chains

Let $\pi_0 = (\pi_0(0), \pi_0(1))$ be the initial distribution. We have already concluded that the distribution of X_n is given by

$$P(X_n = 0) = \frac{q}{p+q} + (1-p-q)^n \left(\pi_0(0) - \frac{q}{p+q} \right),$$

$$P(X_n = 1) = \frac{p}{p+q} + (1-p-q)^n \left(\pi_0(1) - \frac{p}{p+q} \right).$$

As $0 < p+q < 2$, $|1-p-q| < 1$. Let $n \to \infty$ and conclude that

$$\lim_{n\to\infty} P(X_n = 0) = \frac{q}{p+q} \quad \text{and} \quad \lim_{n\to\infty} P(X_n = 1) = \frac{p}{p+q}.$$

The distribution $(\frac{q}{p+q}, \frac{p}{p+q})$ above is called the *limiting distribution*. It describes the long-term behavior of the chain approximately.

A remark on two-state Markov chains

If initially we choose

$$\pi_0(0) = \frac{q}{p+q} \quad \text{and} \quad \pi_0(1) = \frac{p}{p+q},$$

then for all $n \geq 0$,

$$P(X_n = 0) = \frac{q}{p+q} \quad \text{and} \quad P(X_n = 1) = \frac{p}{p+q}.$$

Hence the distribution of X_n is independent of n. We call such $(\frac{q}{p+q}, \frac{p}{p+q})$ the *stationary distribution*. Recall that

$$\lim_{k \to \infty} P^k = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix} = \begin{pmatrix} \pi_0 \\ \pi_0 \end{pmatrix}.$$

This coincidence will be studied in Chapter 2.

Exercise on Markov Chain

Let $X_n, n \geq 0$ be the two-state Markov chain. Find (a) $P(X_1 = 0 | X_0 = 0 \text{ and } X_2 = 0)$ and (b) $P(X_1 \neq X_2)$.

Solution

(a) $P(X_1 = 0 \mid X_0 = 0 \text{ and } X_2 = 0)$

$$= \frac{P(X_0 = 0, X_1 = 0, X_2 = 0)}{P(X_0 = 0, X_2 = 0)}$$

$$= \frac{P(X_0 = 0, X_1 = 0, X_2 = 0)}{P(X_0 = 0, X_1 = 0, X_2 = 0) + P(X_0 = 0, X_1 = 1, X_2 = 0)}$$

$$= \frac{\pi_0(0)(1-p)^2}{\pi_0(0)(1-p)^2 + \pi_0(0)pq}$$

$$= \frac{(1-p)^2}{(1-p)^2 + pq} .$$

Solution

(b) $P(X_1 \neq X_2)$

$$
\begin{aligned}
&= P(X_0 = 0, X_1 \neq X_2) + P(X_0 = 1, X_1 \neq X_2) \\
&= P(X_0 = 0, X_1 = 0, X_2 = 1) + P(X_0 = 0, X_1 = 1, X_2 = 0) + \\
&\quad P(X_0 = 1, X_1 = 0, X_2 = 1) + P(X_0 = 1, X_1 = 1, X_2 = 0) \\
&= \pi_0(0)(1-p)p + \pi_0(0)pq + (1-\pi_0(0))qp + (1-\pi_0(0))(1-q)q \\
&= pq + \pi_0(0)(1-p)p + (1-\pi_0(0))(1-q)q.
\end{aligned}
$$

Ehrenfest chain

(Also see Example 2, page 7 in textbook)

In relation with statistical mechanics, P. and T. Ehrenfest have proposed the following model in 1911. A box contains d molecules. Furthermore, the box is separated in two halves A and B by a "wall" with a small membrane. In each of the successive time instants, a single molecule chosen randomly among all d molecules crosses the membrane to the other half of the box.

Ehrenfest chain

Let X_n denote the number of molecules in A at time n. Then $\{X_n\}_{n\geq 0}$ is a Markov chain with state space $\mathcal{S} = \{0, 1, 2, \dots, d\}$. The transition function is given by

$$\begin{aligned} P(x, x-1) &= \frac{x}{d}, & 1 \leq x \leq d, \\ P(x, x+1) &= \frac{d-x}{d}, & 0 \leq x \leq d-1, \end{aligned}$$

and $P(x, y) = 0$ otherwise. Indeed, $P(x, x-1)$ is the probability, given x particles in side A, that the randomly chosen molecule belongs to side A. Analogously, $P(x, x+1)$ corresponds to the passage of a molecule form side B to side A.

Exercise on Ehrenfest Chain

Let $X_n, n \geq 0$ be the Ehrenfest chain. Find $P(X_0 = X_2)$ if there is some $j \in \{0, 1, 2, \ldots d\}$ such that $P(X_0 = j) = 1$.

Solution

If $j = 0$, i.e. $P(X_0 = 0) = 1$, then

$$\begin{aligned} P(X_0 = X_2) &= P(X_0 = X_2 \text{ and } X_0 = 0) \\ &= \sum_{k=0}^{d} P(X_0 = X_2 = 0 \text{ and } X_1 = k) \\ &= P(X_0 = 0, X_1 = 1, X_2 = 0) \\ &= P(X_0 = 0, X_1 = 1) P(X_1 = 1, X_2 = 2) \\ &= 1 \times \frac{1}{d} \\ &= \frac{1}{d} \end{aligned}$$

Similarly, if $j = d$, we also have $P(X_0 = X_2) = \frac{1}{d}$.

Solution

If $0 < j < d$, then

$$
\begin{aligned}
P(X_0 = X_2) &= P(X_0 = X_2 \text{ and } X_0 = j) \\
&= \sum_{k=0}^{d} P(X_0 = X_2 = j \text{ and } X_1 = k) \\
&= P(X_0 = j, X_1 = j-1, X_2 = j) \\
&\quad + P(X_0 = j, X_1 = j+1, X_2 = j) \\
&= \frac{j}{d} \times \frac{d-j+1}{d} + \frac{d-j}{d} \times \frac{j+1}{d} \\
&= \frac{2dj + d - 2j^2}{d^2}
\end{aligned}
$$

Genetics chain

(Also see Example 7, page 11 in textbook)
Consider a gene composed of d subunits, $d > 0$, and each subunit is either normal(=N) or mutant(=M) in form. Consider a cell with a gene composed of x M-subunits and $d - x$ N-subunits. The gene duplicates before the cell divides into two descendants. Each corresponding descendant gene is composed of d units chosen randomly from the $2x$ M-subunits and the $2(d - x)$ N-subunits.

Genetics chain

Suppose we follow a fixed descendant line from a given gene. Let X_0 be the number of M-subunits initially, and let X_n, $n \geq 1$, be the number of M-subunits in the nth descendant gene. Then $\{X_n\}_{n\geq 0}$ is a Markov chain with state space $\mathcal{S} = \{0, 1, 2, \ldots, d\}$. The transition function is given by

$$P(x, y) = \begin{cases} \dfrac{\binom{2x}{y}\binom{2d-2x}{d-y}}{\binom{2d}{d}}, & \text{if } 2x - d \leq y \leq 2x; \\ 0, & \text{otherwise.} \end{cases}$$

States 0 and d are absorbing states for this chain.