Markov chains invariant distributions
2024-9-17
BFN/bfn

Theme of the day

$$\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j$$

under connecitivity and mild additional assumptions of $\boldsymbol{P}$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P_{ij}^{(n)} \to \pi_j$$

under connecitivity assumptions of $\boldsymbol{P}$

leading to the fix point equation

$$\boldsymbol{\pi} \quad = \quad \boldsymbol{\pi}\boldsymbol{P}$$

Section 4.1 Regular chain

Section 4.1 treats a somewhat simple case with the concept of a regular Markov chain. The setting is a finite state space $S = \{0, 1, 2, \ldots, N\}$ and the regularity assumption is

$$\exists k \forall (i,j) \in S \times S : P_{ij}^{(k)} > 0$$

$\boldsymbol{P}^{N^2} > 0$ implies $\{X_n, n \geq 0\}$ regular. One only needs to keep track of whether entries are positive in successive squaring, $\boldsymbol{P}, \boldsymbol{P}^2, \boldsymbol{P}^4, \ldots, \boldsymbol{P}^{2^n}$ until $n \geq \frac{2\log(N)}{\log(2)}$.

It is claimed that for a regular chain

$$\lim_{n\to\infty} P_{ij}^{(n)} \quad = \quad \lim_{n\to\infty} \mathbb{P}\{X_n = j | X_0 = i\} = \pi_j$$

independent of the initial value $X_0 = i$.

Now Theorem 4.1 tells that the limit $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_N)$ can be found as the unique solution to the equation system

$$\boldsymbol{\pi} \quad = \quad \boldsymbol{\pi}\boldsymbol{P}$$

such that $\sum_{k=0}^{N} \pi_k = 1, (\boldsymbol{\pi e} = 1)$. $\boldsymbol{\pi}$ is indeed a solution

$$
\begin{aligned}
\boldsymbol{P}^{(n)} &= \boldsymbol{P}^{(n-1)}\boldsymbol{P} \\
P_{ij}^{(n)} &= \sum_{k=0}^{N} P_{ik}^{(n-1)} P_{kj} \\
\pi_j &= \lim_{n\to\infty} P_{ij}^{(n)} = \lim_{n\to\infty} \sum_{k=0}^{N} P_{ik}^{(n-1)} P_{kj} = \sum_{k=0}^{N} \left(\lim_{n\to\infty} P_{ik}^{(n-1)} \right) P_{kj} \\
&= \sum_{k=0}^{N} \pi_k P_{kj}
\end{aligned}
$$

So

$$
\pi_j = \sum_{k=0}^{N} \pi_k P_{kj}
$$

Uniqueness of $\boldsymbol{x} = \boldsymbol{xP}, \quad \boldsymbol{xe} = 1$. Elementwise we get

$$
x_j = \sum_{k=0}^{N} x_k P_{kj} = \sum_{k=0}^{N} \left(\sum_{\ell=0}^{N} x_\ell P_{\ell k} \right) P_{kj} = \sum_{\ell=0}^{N} x_\ell \sum_{k=0}^{N} P_{\ell k} P_{kj} = \sum_{\ell=0}^{N} x_\ell P_{\ell j}^{(2)} = \sum_{m=0}^{N} x_m P_{mj}^{(n)}
$$

by recursion. Now taking the limit on both sides we get

$$
\lim_{n\to\infty} x_j = \lim_{n\to\infty} \sum_{m=0}^{N} x_m P_{mj}^{(n)} = \sum_{m=0}^{N} x_m \lim_{n\to\infty} P_{mj}^{(n)} = \sum_{m=0}^{N} x_m \pi_j = \pi_j \sum_{m=0}^{N} x_m = \pi_j
$$

$\boldsymbol{P}$ is called double stochastic if $\boldsymbol{Pe} = \boldsymbol{e}$ and $\boldsymbol{P'e} = \boldsymbol{e}$ elementwise

$$
\sum_{j=0}^{N} P_{ij} = 1, \quad \sum_{i=0}^{N} P_{ij} = 1
$$

both rows and columns sum to 1. In this case $\pi_j = \frac{1}{N+1}$

$$
\sum_{j=0}^{N} \frac{1}{N+1} P_{ij} = \frac{1}{N+1} \sum_{i=0}^{N} P_{ij} = \frac{1}{N+1}
$$

V_i^{n-1} Time spent in state i in first n time steps

$$
\begin{aligned}
V_j^{n-1} &= \sum_{k=0}^{n-1} \mathbf{1}\{X_k = j\} \\
\mathbb{E}\left(V_j^{n-1} | X_0 = i\right) &= \mathbb{E}\left(\sum_{k=0}^{n-1} \mathbf{1}\{X_k = j\} \middle| X_0 = i\right) \\
\mathbb{E}\left(\frac{V_j^{n-1}}{n} \middle| X_0 = i\right) &= \mathbb{E}\left(\frac{1}{n}\sum_{k=0}^{n-1} \mathbf{1}\{X_k = j\} \middle| X_0 = i\right) = \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{E}\left(\mathbf{1}\{X_k = j\} | X_0 = i\right) = \frac{1}{n}\sum_{k=0}^{n-1} P_{ij}^{(k)}
\end{aligned}
$$

What happens in the limit?

$$
\text{if } \lim_{n\to\infty} a_n = a \Leftrightarrow \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} a_n = a
$$

We conclude

$$
\lim_{n\to\infty} \mathbb{E}\left(\frac{V_j^{n-1}}{n} \middle| X_0 = i\right) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} P_{ij}^{(k)} = \pi_j
$$

Section 4.3 The classification of states

We say that state j is accessible from state i if $\exists k \in \mathbb{N} : P_{ij}^{(k)} > 0$. If j is accessible from i and i is accessible from j we say that i and j communicate and write $i \leftrightarrow j$. This is an equivalence relateion

i) $i \leftrightarrow i$

ii) $i \leftrightarrow j \Leftrightarrow j \leftrightarrow i$

iii) $i \leftrightarrow j$ and $j \leftrightarrow k \Leftrightarrow i \leftrightarrow k$

Proof of iii)

$$
\begin{aligned}
i \leftrightarrow j \quad &\Leftrightarrow \exists n_1 : P_{ij}^{(n_1)} > 0 \\
j \leftrightarrow k \quad &\Leftrightarrow \exists n_2 : P_{jk}^{(n_2)} > 0
\end{aligned}
$$

Now

$$
P_{ik}^{n_1+n_2} = \sum_{\ell} P_{i\ell}^{(n_1)} P_{\ell k}^{(n_2)} \geq P_{ij}^{(n_1)} P_{jk}^{(n_2)} > 0
$$

and similarly for $P_{ki}^{n_3+n_4}$. So a state can be in at most one class and will be in exactly one class. A class that can not be left is called closed.

First passage and first return probabilities

$$
\begin{aligned}
f_{ij}^{(n)} &= \mathbb{P}\{X_1 \neq j, X_2 \neq j \dots X_{n-1} \neq j, X_n = j | X_0 = i\} \\
\mathbb{P}\{\exists k : X_k = j | X_0 = i\} &= \sum_{n=1}^{\infty} f_{ij}^{(n)} = f_{ij} \\
f_{ii}^{(n)} &= \mathbb{P}\{X_1 \neq j, X_2 \neq j \dots X_{n-1} \neq j, X_n = i | X_0 = i\}
\end{aligned}
$$

$f_{ii} < 1$ We say that state i is transient

$f_{ii} = 1$ We say that state i is recurrent

$$
\begin{aligned}
R_i &= \min\{n \in \mathbb{N} : X_n = i\} \\
f_{ii}^{(n)} &= \mathbb{P}\{R_i = n\}, \quad n = 1, 2, 3, \dots \\
m_i &= \mathbb{E}(R_i | X_0 = i) = \sum_{n=1}^{\infty} n f_{ii}^{(n)} \\
P_{ij}(n) &= \sum_{k=1}^{n} f_{ij}^{(k)} P_{jj}^{(n-k)}, \quad n \geq 1 \\
P_{ii}(n) &= \sum_{k=1}^{n} f_{ii}^{(k)} P_{ii}^{(n-k)}, \quad n \geq 1
\end{aligned}
$$

For a transient state i we define M as the number of returns to the state.

$$
\begin{aligned}
\mathbb{P}\{M \geq 1\} &= f_{ii} \\
\mathbb{P}\{M \geq n\} &= f_{ii}^n \\
M &\in \text{geo}(1 - f_{ii}) \\
\mathbb{E}(M) &= \frac{f_{ii}}{1 - f_{ii}}
\end{aligned}
$$

Theorem 4.3: A state is recurrent if and only if $\sum_{i=1}^{\infty} P_{ii} = \infty$, equivalently a state is transient if and only if $\sum_{i=1}^{\infty} P_{ii} < \infty$.

We prove Theorem 4.3 in the transient version. If state i is transient then the random variable M of number of returns has finite mean $\mathbb{E}(M) < \infty$ and we can write M as

$$
\begin{aligned}
M &= \sum_{n=1}^{\infty} \mathbf{1}\{X_n = i\} \\
\infty &> \mathbb{E}(M | X_0 = i) = \mathbb{E}\left(\sum_{n=1}^{\infty} \mathbf{1}\{X_n = i\} \middle| X_0 = i \right) = \sum_{n=1}^{\infty} \mathbb{E}\left(\mathbf{1}\{X_n = i\} \middle| X_0 = i \right) = \sum_{n=1}^{\infty} P_{ii}^{(n)}
\end{aligned}
$$

Now suppose $\sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty$

A recurrent state for which $\mathbb{E}(M_i|X_0 = i) = \infty$ is said to be null-recurrent. A recurrent state for which $\mathbb{E}(M_i|X_0 = i) < \infty$ is said to be non null-recurrent or positive recurrent. All states in a recurrent class will be either null-recurrent or non-null recurrent. So we can speak of class properties rather than just state properties. Absorbing states constitute isolated positive recurrent classes.

A class that can be left is necessarily transient. We can partition the state space in a set of closed communication classes and a set of transient states (the latter might consist of one or more communicating classes). A Markov chain that consists of only one communicating class is set to be irreducible.

Periodicity, states can only be visited at certain times (nd).

Section 4.4 The basic limit theorem of Markov chains

Recall

$$\begin{aligned} R_i &= \min\{n \in \mathbb{N} : X_n = i\} \\ m_i &= \mathbb{E}(R_i|X_0 = i) = \sum_{n=1}^{\infty} n f_{ii}^{(n)} \end{aligned}$$

For irreducible aperiodic $\{X_n; n \in \mathbb{N}\}$ Theorem 4.3 says

$$\lim_{n\to\infty} P_{ii}^{(n)} = \frac{1}{\sum_{n=1}^{\infty} n f_{ii}^{(n)}} = \frac{1}{m_i}$$

For a positive ergodic chain $\{X_n; n \in \mathbb{N}\}$ we have in addition

$$\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j, \quad \pi_j = \sum_{i=0}^{N} \pi_i P_{ij}, \quad \sum_{i=0}^{N} \pi_i = 1, \quad N \text{ finite or } \infty$$

with $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_N)$ being the unique solution to $\boldsymbol{\pi} = \boldsymbol{\pi P}$ with $\boldsymbol{\pi e} = 1$.

Suppose $\mathbb{P}\{X_0 = i\} = \pi_i, \forall i$, so $\boldsymbol{p}^{(0)} = \boldsymbol{\pi}$ then from

$$\begin{aligned} \boldsymbol{p}^{(n)} &= \boldsymbol{p}^{(n-1)} \boldsymbol{P} = \boldsymbol{p}^{(0)} \boldsymbol{P}^n \\ \boldsymbol{p}^{(n)} &= \boldsymbol{\pi P}^n = \boldsymbol{\pi P P}^{n-1} = \boldsymbol{\pi P}^{n-1} = \boldsymbol{\pi} \end{aligned}$$

The chain is called stationary. We call $\boldsymbol{\pi}$ the stationary or invariant probability distribution.

For a stationary chain we have

$$\mathbb{P}\{X_n = i, X_{n+1} = j\} \quad = \quad \mathbb{P}\{X_n = i\}\mathbb{P}\{X_{n+1} = j | X_n = i\} = \pi_i P_{ij}$$

$$\begin{aligned}
\boldsymbol{P} &= \left\| \begin{array}{ccccc} 0 & 1 & 0 & 0 & \dots \\ q_1 & 0 & p_1 & 0 & \dots \\ 0 & q_2 & 0 & p_2 & \dots \end{array} \right\| \\
\pi_0 &= q_1 \pi_1 \\
\pi_1 &= \pi_0 + \pi_2 q_2 \\
\pi_1 &= \pi_1 q_1 + \pi_2 q_2 \\
\pi_1 p_1 &= \pi_2 q_2 \\
\pi_i &= \pi_{i-1} p_{i-1} + \pi_{i+1} q_{i+1} \\
\pi_i &= \pi_i q_i + \pi_{i+1} q_{i+1} \\
\pi_i p_i &= \pi_{i+1} q_{i+1} \\
\pi_i &= \pi_0 \prod_{k=0}^{i-1} \frac{p_k}{q_k}
\end{aligned}$$

Assume $p_k = p$ and $q_k = q$

$$\begin{aligned}
X_i &= \left(\frac{p}{q}\right)^i x_0 \\
\sum_{i=0}^{\infty} \left(\frac{p}{q}\right)^i x_0 &= x_0 \sum_{i=0}^{\infty} \left(\frac{p}{q}\right)^i = x_0 \frac{1}{1 - \frac{p}{q}} = x_0 \frac{q}{q-p}, \quad p < 1
\end{aligned}$$

This is an example of a periodic chain

Interpretation/roles of limit probabilities