Linear Algebra II
Discrete-Time Markov Chain

Ming-Hsuan Kang

Probability Vectors and Transition Matrices

Consider a state space $\Omega = \{s_1, \cdots, s_n\}$. A vector in $\mathbb{R}^n$ is called a probability vector of Ω if it contains non-negative entries which sum to 1. Suppose the probability of going from the state s_j to state s_i in is equal to P_{ij} from time k to time $k+1$. Then the $n \times n$ matrix $P = (P_{ij})$ is called a transition matrix, which columns are probability vectors.

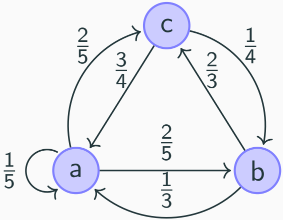
Definition

A matrix or a vector is called positive (resp. non-negative) if all of its entries are positive (resp. non-negative).

Example

The following is a transition matrix with 3 states and its diagram.

$$\begin{pmatrix} \frac{1}{5} & \frac{1}{3} & \frac{3}{4} \\ \frac{2}{5} & 0 & \frac{1}{4} \\ \frac{2}{5} & \frac{2}{3} & 0 \end{pmatrix}$$



Markov chain

Given the initial probability vector $\vec{\pi}^{(0)}$, define the probability vector at time k (which must be a positive integer) with respect to the transition matrix P using the following recursive formula

$$\vec{\pi}^{(k)} = P\vec{\pi}^{(k-1)}.$$

The above definition implies that

$$\vec{\pi}^{(k)} = P\vec{\pi}^{(k-1)} = \cdots = P^k\vec{\pi}^{(0)}.$$

The process given by the above description is called a discrete-time Markov chain. Our goal of this lecture is to study the asymptotic behavior of $\vec{\pi}^{(k)}$ when k goes to infinity.

Example

For the transition matrix $P = \begin{pmatrix} \frac{1}{5} & \frac{1}{3} & \frac{3}{4} \\ \frac{2}{5} & 0 & \frac{1}{4} \\ \frac{2}{5} & \frac{2}{3} & 0 \end{pmatrix}$, if $\vec{\pi}^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, we have

$$\vec{\pi}^{(1)} = \begin{pmatrix} \frac{1}{5} \\ \frac{2}{5} \\ \frac{2}{5} \end{pmatrix}, \vec{\pi}^{(5)} \approx \begin{pmatrix} 0.424853 \\ 0.25024 \\ 0.324907 \end{pmatrix}, \vec{\pi}^{(20)} \approx \begin{pmatrix} 0.416667 \\ 0.25 \\ 0.333333 \end{pmatrix}.$$

If $\vec{\pi}^{(0)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, we have

$$\vec{\pi}^{(1)} = \begin{pmatrix} \frac{1}{3} \\ 0 \\ \frac{2}{3} \end{pmatrix}, \vec{\pi}^{(5)} \approx \begin{pmatrix} 0.3952 \\ 0.2624 \\ 0.3424 \end{pmatrix}, \vec{\pi}^{(20)} \approx \begin{pmatrix} 0.416667 \\ 0.25 \\ 0.333333 \end{pmatrix}.$$

Eigenvalues of Transition Matrices

Note that the sum of each row of P^t is equal to one, which implies $(1, \cdots, 1)^t$ is an eigenvector of P^t corresponding to the eigenvalue one.

Theorem

Let λ be an eigenvalue of the transition matrix P, then $|\lambda| \leq 1$.

Since P and P^t have the same eigenvalues, it is sufficient to study the eigenvalues of P^t. Let $\vec{v} = (v_1, \cdots, v_n)^t$ be an eigenvector of P^t such that $P^t(\vec{v}) = \lambda\vec{v}$. Suppose v_i has the largest absolute value among of entries of $\vec{v}$. Then

$$|\lambda v_i| = |(P^t\vec{v})_i| = \left|\sum_{j=1}^{n} P_{ji}v_j\right| \leq \sum_{j=1}^{n} P_{ji}\,|v_j| \leq \sum_{j=1}^{n} P_{ji}|v_i| = |v_i|$$

We conclude that $|\lambda| \leq 1$.

Suppose $|\lambda| = 1$. From the previous inequality, we obtain

$$|\lambda v_i| = |(P^t \vec{v})_i| = \left| \sum_{j=1}^{n} P_{ji} v_j \right| = \sum_{j=1}^{n} P_{ji} \, |v_j| = \sum_{j=1}^{n} P_{ji} |v_i| = |v_i|.$$

The equality $\sum_{j=1}^{n} P_{ji} \, |v_j| = \sum_{j=1}^{n} P_{ji} |v_i|$ implies that $|v_i| = |v_j|$ if $P_{ji} > 0$; together with $\left| \sum_{j=1}^{n} P_{ji} v_j \right| = \sum_{j=1}^{n} P_{ji} \, |v_j|$, we conclude that $v_i = v_j$ if $P_{ji} > 0$;

When P is positive, we must have $v = v_i(1, \cdots, 1) = v_i \vec{1}$ and $\lambda = 1$. Therefore, one is the only eigenvalue with the absolute value one and its geometric multiplicity equals one.

Algebraic Multiplicity of the Eigenvalue One

Let us show that the algebraic multiplicity of the eigenvalue one of P^t is also equal to one. Suppose not, then P^t has a unique one-Jordan block of size greater than one and there exists a nonzero vector $\vec{u}$ such that

$$P^t(\vec{u}) = \vec{1} + \vec{u} \quad \text{and} \quad (P^t)^k(\vec{u}) = k\vec{1} + \vec{u}, \forall k.$$

On the other hand, for all $k > 0$,

$$\|(P^t)^k(\vec{u})\|^2 = \sum_i \left| \sum_j ((P^t)^k)_{ij} u_j \right|^2 \leq \sum_i (\sum_j |u_j|)^2$$

but

$$\|k\vec{1} + \vec{u}\| \geq \|k\vec{1}\| - |\vec{u}\| \to \infty \quad \text{as } k \to \infty.$$

Thus, we obtain a contradiction.

Results

Since P and P^t have the same characteristic polynomial, we have the following result.

Theorem

For a transition matrix with positive entries, the following hold.

- *1 is an eigenvalue with geometric/algebraic multiplicity one.*
- *All the other eigenvalues have absolute values less then 1.*

Now, we are ready to study the asymptotic behavior $\vec{\pi}^{(k)}$. Recall that we can decompose $\mathbb{C}^n$ as a direct sum of L_P-invariant subspace $W_1 \oplus \cdots \oplus W_r$, such that the restriction of L_P on W_i can be represented as a single Jordan block.

Write $\vec{\pi}^{(0)} = \vec{v}_1 + \cdots + \vec{v}_r$ with $\vec{v}_i \in W_i$, then

$$\vec{\pi}^{(k)} = P^k \vec{\pi}^{(0)} = \sum_{i=1}^{r} P^k(\vec{v}_i) = \sum_{i=1}^{r} \left(L_P\big|_{W_i} \right)^k \vec{(v_i)}.$$

Asymptotic Behaviour of $\vec{\pi}^{(k)}$

Now suppose all entries of P are positive. By the previous theorem, we may assume that W_1 is the one-eigenspace of P which is of one dimension. Then

$$\left(L_P\big|_{W_1}\right)^k \vec{v}_1 = (1)^k \vec{v}_1 = \vec{v}_1.$$

Asymptotic Behaviour of $\vec{\pi}^{(k)}$

Next we consider the action of L_P on W_i for $i > 1$. Let m be the dimension of W_i and α be the Jordan basis of L_P on W_i such that

$$J = \mathrm{Rep}_\alpha(L_P\big|_{W_i}) = \lambda I_m + N,$$

Here λ is an eigenvalue of P and N is a nilpotent matrix with $N^m = 0$. By the previous theorem again, we have $|\lambda| < 1$. Note that when $k > m$,

$$J^k = (\lambda I_m + N)^k = \lambda^k I_m + \sum_{j=1}^{m-1} \binom{k}{j} \lambda^{k-j} N^j.$$

Since $|\lambda| < 1$, we have $\lim_{k\to\infty} J^k = 0$.

Corollary

Let B be a complex matrix which all eigenvalues have absolute values less than one. Then $\lim_{k\to\infty} B^k = 0$.

Therefore,

$$\lim_{k\to\infty} \mathrm{Rep}_\alpha(P^k \vec{v}_i) = \lim_{k\to\infty} J^k \mathrm{Rep}_\alpha(\vec{v}_i) = \vec{0}.$$

This implies that

$$\lim_{k\to\infty} P^k \vec{v}_i = \vec{0}.$$

Combining the above results, we conclude that

$$\lim_{k\to\infty} \vec{\pi}^{(k)} = \vec{v}_1 + \vec{0} + \cdots + \vec{0} = \vec{v}_1.$$

Since $\vec{\pi}^{(k)}$ is a probability vectors for all k, its limit $\vec{v}_1$ is also a probability vector, denoted by $\vec{\pi}^{(\infty)}$. Next we will show that $\vec{\pi}^{(\infty)}$ is indeed independent of the initial probability vector $\vec{\pi}^{(0)}$.

Uniqueness of Stationary Probability Vector

A probability vector is called stationary if its is P-invariant. For instance, $\vec{\pi}^{(\infty)}$ is stationary. Let $\vec{\tau}^{(0)}$ be another initial probability vector. Note that $\vec{\pi}^{(\infty)}$ and $\vec{\tau}^{(\infty)}$ are both contained in the one-eigensapce of P, which is of one-dimension (provided that P is positive). On the other hand, in a one dimensional subspace, there exists at most one probability vector. Therefore, we must have $\vec{\pi}^{(\infty)} = \vec{\tau}^{(\infty)}$.

Conclusion

Theorem

For a positive transition matrix P, the following hold.

- *One is an eigenvalue of P with algebraic and geometric multiplicity one.*
- *All the other eigenvalues have absolute values less than one.*
- *There exists a unique stationary probability vector $\vec{\pi}_\infty$ such that for any initial probability vector $\vec{\pi}$, we have*

$$\lim_{k \to \infty} P^k \vec{\pi} = \vec{\pi}_\infty.$$

Example

For the transition matrix $P = \begin{pmatrix} \frac{1}{5} & \frac{1}{3} & \frac{3}{4} \\ \frac{2}{5} & 0 & \frac{1}{4} \\ \frac{2}{5} & \frac{2}{3} & 0 \end{pmatrix}$, we have

$$P^2 = \begin{pmatrix} \frac{71}{150} & \frac{17}{30} & \frac{7}{30} \\ \frac{9}{50} & \frac{3}{10} & \frac{3}{10} \\ \frac{26}{75} & \frac{2}{15} & \frac{7}{15} \end{pmatrix}.$$

If $\vec{\pi}^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, we have $\vec{\pi}^{(20)} \approx \begin{pmatrix} 0.416667 \\ 0.25 \\ 0.333333 \end{pmatrix}$.

If $\vec{\pi}^{(0)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, we have $\vec{\pi}^{(20)} \approx \begin{pmatrix} 0.416667 \\ 0.25 \\ 0.33333 \end{pmatrix}$. To find the stationary probability vector of P, we compute

$$\ker(P - I) = \mathrm{span}\{\begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}\}.$$

Thus the stationary probability vector is $\frac{1}{12}\begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}$.