THE FUNDAMENTAL THEOREM OF MARKOV CHAINS

AARON PLAVNICK

ABSTRACT. This paper provides some background for and proves the Fundamental Theorem of Markov Chains. It provides some basic definitions and notation for recursion, periodicity, and stationary distributions.

CONTENTS

1. DEFINITIONS AND BACKGROUND

So what is a Markov Chain, let's define it.

Definition 1.1. Let $\{X_0, X_1, \ldots\}$ be a sequence of random variables and $Z = 0, \pm 1, \pm 2, \ldots$ be the union of the sets of their realizations.Then $\{X_0, X_1, \ldots\}$ is called a *discrete-time Markov Chain* with state space Z if:

$$P(X_{n+1} = i_{n+1} | X_n = i_n, \ldots, X_1 = i_1) = P(X_{n+1} = i_{n+1} | X_n = i_n)$$

Now lets set up some notation for the *one-step transition probabilities* of the Markov Chain Let:

$$p_{ij}(n) = P(X_{n+1} = j | X_n = i); \;\; n = 0, 1, \ldots$$

We will limit ourselves to *homogeneous* Markov Chains. Or Markov Chains that do not evolve in time.

Definition 1.2. We say that a Markov Chain is *homogeneous* if its one-step transition probabilities do not depend on n ie.

$$\forall n, m \in N \;\; and \;\; i, j \in Z \quad p_{ij}(n) = p_{ij}(m)$$

We then define the *n-step transition probabilities* of a homogeneous Markov Chain by

$$p_{ij}^{(m)} = P(X_{n+m} = j | X_n = i)$$

Where by convention we define

$$p_{ij}^{(0)} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Date: August 22, 2008.

With these definitions we are now ready for a quick theorem: the *Chapman-Kolmogorov equation*. This serves to relate the transition probabilities of a Markov Chain.

Theorem 1.3.
$$p_{ij}^{(m)} = \sum_{k\in\mathbf{Z}} p_{ik}^{(r)} p_{kj}^{(m-r)}; \quad \forall\ r \in N \cup \{0\}$$

Proof. Making use of the total probability rule as well as the Markovian property in this proof yields:
By total probability rule:

$$p_{ij}^{(m)} = P(X_m = j|X_0 = i) = \sum_{k\in\mathbf{Z}} P(X_m = j, X_r = k|X_0 = i)$$
$$= \sum_{k\in\mathbf{Z}} P(X_m = j|X_r = k, X_0 = i)P(X_r = k|X_0 = i)$$

and now by the Markovian property:

$$= \sum_{k\in\mathbf{Z}} P(X_m = j|X_r = k)P(X_r = k|X_0 = i) = \sum_{k\in\mathbf{Z}} p_{ik}^{(r)} p_{kj}^{(m-r)}$$

□

The Chapaman-Kolmogrov equation allows us to associate our n-step transition probabilities to a matrix of n-step transistion probabilities. Lets define the matrix:

$$P^{(m)} = ((p_{ij}^{(m)}))$$

Corollary 1.4. Observe now that:
$$P^{(m)} = P^m$$
Where P^m here denotes the typical matrix multiplication.

Proof. We can rewrite the Chapaman-Kolmogrov equation in matrix form as:
$$P^{(m)} = P^{(r)}P^{(m-r)}; \quad \forall\ r\ \in\ N \cup \{0\}$$
Now we can induct on m. It's apparent that
$$P^{(1)} = P$$

Giving us:

$$P^{(m+1)} = P^{(r)}P^{(m+1-r)}$$

□

A Markov Chain's transition probabilities along with an initial distribution completely determine the chain.

Definition 1.5. An initial distribution is a probability distribution $\{\pi_i = P(X_0 = i)|i \in \mathbb{Z}\}$ $\sum_{i\in\mathbb{Z}} \pi_i = 1$
Such a distribution is said to be stationary if it satisfies
$$\pi_j = \sum_{i\in\mathbb{Z}} \pi_i p_{ij}$$

Definition 1.6. We say that a subset space state $\mathbf{C} \subset \mathbf{Z}$ is *closed* if

$$\sum_{j \in \mathbf{C}} p_{ij} \ = \ 1 \ \ \forall \, i \in \mathbf{C}$$

If $\mathbf{Z}$ itself has no proper closed subsets then the Markov Chain is said to be *irreducible.*

Definition 1.7. We define the *period* of a state i by

$$d_i = gcd(m \in \mathbb{Z} | p_{ii}^{(m)} > 0)$$

State i is *aperiodic* if $d_i = 1$

Definition 1.8. A state i is said to be *accessible* from from a state j if there is an $m \geq 1$ s.t

$$p_{ij}^{(m)} > 0$$

If i is accessible from j and j is accessible from i then i and j are said to *communicate.*

Definition 1.9. We define the *first-passage time probabilities* by

$$f_{ij}^{(m)} = P(X_m = j; \ X_k \neq j, 0 < k < m-1 | X_0 = i) \ ; \quad i, j \in \mathbb{Z}$$

And we will denote the expected value, the expected return time, of this distribution by

$$\mu_{ij} = \sum_{m=1}^{\infty} m f_{ij}^{(m)}$$

Definition 1.10. We say that a state i is *recurrent* if

$$\sum_{m=1}^{\infty} f_{ij}^{(m)} = 1$$

and *transient* if

$$\sum_{m=1}^{\infty} f_{ij}^{(m} < 1$$

A recurrent state i is *positive-recurrent* if $\mu_{ii} < \infty$ and *null-recurrent* if $\mu_{ii} = \infty$

2. Fundamental Theorem of Markov Chains

Theorem 2.1. For any irreducible, aperiodic, positive-recurrent Markov Chain there exist a unique stationary distribution $\{\pi_j, j \in \mathbb{Z}\}$ s.t $\forall i \in \mathbb{Z}$

Proof. Because our chain is irreducible, aperiodic, and positive-recurrent we know that for all $i \in \mathbb{Z} \quad \pi_j = \lim_{n \to \infty} p_{ij}^{(n)} > 0$. Likewise as π_j is a probability distribution $\sum_{j \in \mathbb{Z}} \pi_j$. We know that for any m

$$\sum_{i=0}^{m} p_{ij}^{(m)} \leq \sum_{i=0}^{\infty} p_{ij}^{(m)} \leq 1$$

taking the limit

$$\lim_{m\to\infty}\sum_{i=0}^{m} p_{ij}^{(m)} = \sum_{i=0}^{\infty}\pi_j \leq 1$$

which implies that for any M

$$\sum_{i=0}^{M}\pi_j \leq 1$$

Now we can use the *Chapman-Kolmogorov* equation analogously

$$p_{ij}^{(m+1)} = \sum_{i=0}^{\infty} p_{ik}^{(m)} p_{kj} \geq \sum_{i=0}^{M} p_{ik}^{(m)} p_{kj}$$

as before we take the limit as $m, M \to \infty$ yielding

$$\pi_j \geq \sum_{i=0}^{\infty}\pi_k p_{kj}$$

Now we search for contradiction and assume strict inequality holds for at least one state j. Summing over these inequalities yields

$$\sum_{j=0}^{\infty}\pi_j > \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\pi_k p_{kj} = \sum_{k=0}^{\infty}\pi_k \sum_{j=0}^{\infty} p_{kj} = \sum_{k=0}^{\infty}\pi_k$$

But this is a contradiction, hence equality must hold

$$\pi_j = \sum_{k=0}^{\infty}\pi_k p_{kj}$$

Thus a unique stationary distribution exists. □

Futher it can be shown that this unique distribution is related to the expected return value of the Markov chain by

$$\pi_j = \frac{1}{\mu_{jj}}$$

References

[1] F. E. Beichelt. L. P. Fatti. Stochastic Processes And Their Applications. Taylor and Francis. 2002.