

11.4: Fundamental Limit Theorem for Regular Chains**

The fundamental limit theorem for regular Markov chains states that if \mathbf{P} is a regular transition matrix then

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \mathbf{W}, \quad (11.4.1)$$

where \mathbf{W} is a matrix with each row equal to the unique fixed probability row vector \mathbf{w} for \mathbf{P} . In this section we shall give two very different proofs of this theorem.

Our first proof is carried out by showing that, for any column vector \mathbf{y} , $\mathbf{P}^n \mathbf{y}$ tends to a constant vector. As indicated in Section 1.3, this will show that \mathbf{P}^n converges to a matrix with constant columns or, equivalently, to a matrix with all rows the same.

The following lemma says that if an r -by- r transition matrix has no zero entries, and \mathbf{y} is any column vector with r entries, then the vector $\mathbf{P}\mathbf{y}$ has entries which are "closer together" than the entries in \mathbf{y} .

Lemma 11.4.1

Let \mathbf{P} be an r -by- r transition matrix with no zero entries. Let d be the smallest entry of the matrix. Let \mathbf{y} be a column vector with r components, the largest of which is M_0 and the smallest m_0 . Let M_1 and m_1 be the largest and smallest component, respectively, of the vector $\mathbf{P}\mathbf{y}$. Then

$$M_1 - m_1 \leq (1 - 2d)(M_0 - m_0). \quad (11.4.2)$$

Proof: In the discussion following Theorem 11.3.1, it was noted that each entry in the vector $\mathbf{P}\mathbf{y}$ is a weighted average of the entries in \mathbf{y} . The largest weighted average that could be obtained in the present case would occur if all but one of the entries of \mathbf{y} have value M_0 and one entry has value m_0 , and this one small entry is weighted by the smallest possible weight, namely d . In this case, the weighted average would equal

$$dm_0 + (1 - d)M_0. \quad (11.4.3)$$

Similarly, the smallest possible weighted average equals

$$dM_0 + (1 - d)m_0. \quad (11.4.4)$$

Thus,

$$\begin{aligned} M_1 - m_1 &\leq (dm_0 + (1 - d)M_0) - (dM_0 + (1 - d)m_0) \\ &= (1 - 2d)(M_0 - m_0). \end{aligned}$$

This completes the proof of the lemma.

We turn now to the proof of the fundamental limit theorem for regular Markov chains.

Theorem Fundamental Limit Theorem for Regular Chains 11.4.1

If \mathbf{P} is the transition matrix for a regular Markov chain, then

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \mathbf{W}, \quad (11.4.5)$$

where \mathbf{W} is matrix with all rows equal. Furthermore, all entries in \mathbf{W} are strictly positive.

Proof. We prove this theorem for the special case that \mathbf{P} has no 0 entries. The extension to the general case is indicated in Exercise [exer 11.4.6]. Let be any r -component column vector, where r is the number of states of the chain. We assume that $r > 1$, since otherwise the theorem is trivial. Let M_n and m_n be, respectively, the maximum and minimum components of the vector $\mathbf{P}^n \mathbf{y}$. The vector $\mathbf{P}^n \mathbf{y}$ is obtained from the vector $\mathbf{P}^{n-1} \mathbf{y}$ by multiplying on the left by the matrix \mathbf{P} . Hence each component of $\mathbf{P}^n \mathbf{y}$ is an average of the components of $\mathbf{P}^{n-1} \mathbf{y}$. Thus

$$M_0 \geq M_1 \geq M_2 \geq \cdots \quad (11.4.6)$$

and

$$m_0 \leq m_1 \leq m_2 \leq \cdots . \quad (11.4.7)$$

Each sequence is monotone and bounded:

$$m_0 \leq m_n \leq M_n \leq M_0 . \quad (11.4.8)$$

Hence, each of these sequences will have a limit as n tends to infinity.

Let M be the limit of M_n and m the limit of m_n . We know that $m \leq M$. We shall prove that $M - m = 0$. This will be the case if $M_n - m_n$ tends to 0. Let d be the smallest element of \mathbf{P} . Since all entries of \mathbf{P} are strictly positive, we have $d > 0$. By our lemma

$$M_n - m_n \leq (1 - 2d)(M_{n-1} - m_{n-1}) . \quad (11.4.9)$$

From this we see that

$$M_n - m_n \leq (1 - 2d)^n (M_0 - m_0) . \quad (11.4.10)$$

Since $r \geq 2$, we must have $d \leq 1/2$, so $0 \leq 1 - 2d < 1$, so the difference $M_n - m_n$ tends to 0 as n tends to infinity. Since every component of $\mathbf{P}^n \mathbf{y}$ lies between m_n and M_n , each component must approach the same number $u = M = m$. This shows that

$$\lim_{n \rightarrow \infty} \mathbf{P}^n \mathbf{y} = \mathbf{u} , \quad (11.4.11)$$

where \mathbf{u} is a column vector all of whose components equal u .

Now let \mathbf{y} be the vector with j th component equal to 1 and all other components equal to 0. Then $\mathbf{P}^n \mathbf{y}$ is the j th column of \mathbf{P}^n . Doing this for each j proves that the columns of \mathbf{P}^n approach constant column vectors. That is, the rows of \mathbf{P}^n approach a common row vector \mathbf{w} , or,

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \mathbf{W} . \quad (11.4.12)$$

It remains to show that all entries in \mathbf{W} are strictly positive. As before, let \mathbf{y} be the vector with j th component equal to 1 and all other components equal to 0. Then $\mathbf{P} \mathbf{y}$ is the j th column of \mathbf{P} , and this column has all entries strictly positive. The minimum component of the vector $\mathbf{P} \mathbf{y}$ was defined to be m_1 , hence $m_1 > 0$. Since $m_1 \leq m$, we have $m > 0$. Note finally that this value of m is just the j th component of \mathbf{w} , so all components of \mathbf{w} are strictly positive.

Doebelin's Proof

We give now a very different proof of the main part of the fundamental limit theorem for regular Markov chains. This proof was first given by Doebelin,¹⁷ a brilliant young mathematician who was killed in his twenties in the Second World War.

Theorem 11.4.2

Let \mathbf{P} be the transition matrix for a regular Markov chain with fixed vector \mathbf{w} . Then for any initial probability vector \mathbf{u} , $\mathbf{u} \mathbf{P}^n \rightarrow \mathbf{w}$ as $n \rightarrow \infty$.

Proof. Let X_0, X_1, \dots be a Markov chain with transition matrix \mathbf{P} started in state s_i . Let Y_0, Y_1, \dots be a Markov chain with transition probability \mathbf{P} started with initial probabilities given by \mathbf{w} . The X and Y processes are run independently of each other.

We consider also a third Markov chain \mathbf{P}^* which consists of watching both the X and Y processes. The states for \mathbf{P}^* are pairs (s_i, s_j) . The transition probabilities are given by

$$\mathbf{P}^*[(i, j), (k, l)] = \mathbf{P}(i, k) \cdot \mathbf{P}(j, l) . \quad (11.4.13)$$

Since \mathbf{P} is regular there is an N such that $\mathbf{P}^N(i, j) > 0$ for all i and j . Thus for the \mathbf{P}^* chain it is also possible to go from any state (s_i, s_j) to any other state (s_k, s_l) in at most N steps. That is \mathbf{P}^* is also a regular Markov chain.

We know that a regular Markov chain will reach any state in a finite time. Let T be the first time the the chain \mathbf{P}^* is in a state of the form (s_k, s_k) . In other words, T is the first time that the X and the Y processes are in the same state. Then we have shown that

$$P[T > n] \rightarrow 0 \text{ as } n \rightarrow \infty . \quad (11.4.14)$$

If we watch the X and Y processes after the first time they are in the same state we would not predict any difference in their long range behavior. Since this will happen no matter how we started these two processes, it seems clear that the long range behaviour should not depend upon the starting state. We now show that this is true.

We first note that if $n \geq T$, then since X and Y are both in the same state at time T ,

$$P(X_n = j \mid n \geq T) = P(Y_n = j \mid n \geq T). \quad (11.4.15)$$

If we multiply both sides of this equation by $P(n \geq T)$, we obtain

$$P(X_n = j, n \geq T) = P(Y_n = j, n \geq T). \quad (11.4.16)$$

We know that for all n ,

$$P(Y_n = j) = w_j. \quad (11.4.17)$$

But

$$P(Y_n = j) = P(Y_n = j, n \geq T) + P(Y_n = j, n < T), \quad (11.4.18)$$

and the second summand on the right-hand side of this equation goes to 0 as n goes to ∞ , since $P(n < T)$ goes to 0 as n goes to ∞ . So,

$$P(Y_n = j, n \geq T) \rightarrow w_j, \quad (11.4.19)$$

as n goes to ∞ . From Equation 11.4.1, we see that

$$P(X_n = j, n \geq T) \rightarrow w_j, \quad (11.4.20)$$

as n goes to ∞ . But by similar reasoning to that used above, the difference between this last expression and $P(X_n = j)$ goes to 0 as n goes to ∞ . Therefore,

$$P(X_n = j) \rightarrow w_j, \quad (11.4.21)$$

as n goes to ∞ .

This completes the proof.

In the above proof, we have said nothing about the rate at which the distributions of the X_n 's approach the fixed distribution \mathbf{w} . In fact, it can be shown that¹⁸

$$\sum_{j=1}^r |P(X_n = j) - w_j| \leq 2P(T > n). \quad (11.4.22)$$

The left-hand side of this inequality can be viewed as the distance between the distribution of the Markov chain after n steps, starting in state s_i , and the limiting distribution \mathbf{w} .

Exercises

Exercise 11.4.1

Define \mathbf{P} and \mathbf{y} by

$$\mathbf{P} = \begin{pmatrix} .5 & .5 \\ .25 & .75 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (11.4.23)$$

Compute $\mathbf{P}\mathbf{y}$, $\mathbf{P}^2\mathbf{y}$, and $\mathbf{P}^4\mathbf{y}$ and show that the results are approaching a constant vector. What is this vector?

Exercise 11.4.2

Let \mathbf{P} be a regular $r \times r$ transition matrix and \mathbf{y} any r -component column vector. Show that the value of the limiting constant vector for $\mathbf{P}^n\mathbf{y}$ is $\mathbf{w}\mathbf{y}$.

Exercise 11.4.3

Let

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ .25 & 0 & .75 \\ 0 & 0 & 1 \end{pmatrix} \quad (11.4.24)$$

be a transition matrix of a Markov chain. Find two fixed vectors of \mathbf{P} that are linearly independent. Does this show that the Markov chain is not regular?

Exercise 11.4.4

Describe the set of all fixed column vectors for the chain given in Exercise 11.4.3

Exercise 11.4.5

The theorem that $\mathbf{P}^n \rightarrow \mathbf{W}$ was proved only for the case that \mathbf{P} has no zero entries. Fill in the details of the following extension to the case that \mathbf{P} is regular. Since \mathbf{P} is regular, for some N , \mathbf{P}^N has no zeros. Thus, the proof given shows that $M_{nN} - m_{nN}$ approaches 0 as n tends to infinity. However, the difference $M_n - m_n$ can never increase. (Why?) Hence, if we know that the differences obtained by looking at every N th time tend to 0, then the entire sequence must also tend to 0.

Exercise 11.4.6

Let \mathbf{P} be a regular transition matrix and let \mathbf{w} be the unique non-zero fixed vector of \mathbf{P} . Show that no entry of \mathbf{w} is 0.

Exercise 11.4.7

Here is a trick to try on your friends. Shuffle a deck of cards and deal them out one at a time. Count the face cards each as ten. Ask your friend to look at one of the first ten cards; if this card is a six, she is to look at the card that turns up six cards later; if this card is a three, she is to look at the card that turns up three cards later, and so forth. Eventually she will reach a point where she is to look at a card that turns up x cards later but there are not x cards left. You then tell her the last card that she looked at even though you did not know her starting point. You tell her you do this by watching her, and she cannot disguise the times that she looks at the cards. In fact you just do the same procedure and, even though you do not start at the same point as she does, you will most likely end at the same point. Why?

Exercise 11.4.8

Write a program to play the game in Exercise 11.4.7.

Exercise 11.4.9

(Suggested by Peter Doyle) In the proof of Theorem 11.4.1, we assumed the existence of a fixed vector \mathbf{w} . To avoid this assumption, beef up the coupling argument to show (without assuming the existence of a stationary distribution \mathbf{w}) that for appropriate constants C and $r < 1$, the distance between αP^n and βP^n is at most $C r^n$ for any starting distributions α and β . Apply this in the case where $\beta = \alpha P$ to conclude that the sequence αP^n is a Cauchy sequence, and that its limit is a matrix W whose rows are all equal to a probability vector w with $wP = w$. Note that the distance between αP^n and w is at most $C r^n$, so in freeing ourselves from the assumption about having a fixed vector we've proved that the convergence to equilibrium takes place exponentially fast.

This page titled [11.4: Fundamental Limit Theorem for Regular Chains**](#) is shared under a [GNU Free Documentation License 1.3](#) license and was authored, remixed, and/or curated by [Charles M. Grinstead & J. Laurie Snell](#) ([American Mathematical Society](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.