Discrete Time Markov Chains
Part 1

Chapter 3 of notes (section 3.1)

STAT2003 / STAT3102

Contents

- Definition of discrete time Markov chains
- Transition matrices
- Chapman-Kolmogorov equations
- Invariant distributions

We will consider only discrete time, discrete state space processes for now.

Discrete-time Markov chains

Definition

A discrete-time Markov chain is a sequence of random variables $X_0, X_1, X_2, \ldots$ taking values in a finite or countable state space S such that, for all $n, i, j, i_0, i_1, \ldots, i_{n-1}$

$$\begin{aligned} &\mathrm{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\ =&\mathrm{P}(X_{n+1} = j | X_n = i), \end{aligned}$$

that is, a discrete-time stochastic process satisfying the Markov property.

Time homogeneity

Definition

A Markov chain is *time-homogeneous* if

$$P(X_{n+1} = j \mid X_n = i)$$

does not depend on n for all $i, j \in S$.

Note that when we assume time homogeneity, even when $n \neq m$,

$$P(X_{n+1} = j \mid X_n = i) = P(X_{m+1} = j \mid X_m = i), \qquad \text{for } i, j, \in S.$$

Transition probabilities

Definition

When we assume time homogeneity, the probabilities

$$p_{ij} = \mathrm{P}(X_{n+1} = j \mid X_n = i), \qquad \text{for } i, j, \in S,$$

are called the (1-step) *transition probabilities* of the Markov chain.

Note: p_{ij} does not depend on n since we assume time homogeneity, so

$$p_{ij} = \mathrm{P}(X_{n+1} = j \mid X_n = i) = \mathrm{P}(X_{m+1} = j \mid X_m = i)$$

even for $m \neq n$.

Transition matrices

- The p_{ij}s are often collected into a matrix P called the *transition matrix*.
- The transition probability p_{ij} is the element of P in the ith row and the jth column, that is, the (i, j)th element of P.
- We will only deal with time-homogeneous Markov chains.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ p_{31} & p_{32} & \cdots & p_{3k} \\ \cdots & \cdots & \cdots & \cdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}$$

For example, p_{32} is the probability that the process moves **from** state 3 **to** state 2 in one step.

Transition matrices

- The p_{ij}s are often collected into a matrix P called the *transition matrix*.
- The transition probability p_{ij} is the element of P in the ith row and the jth column, that is, the (i, j)th element of P.
- We will only deal with time-homogeneous Markov chains.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ p_{31} & p_{32} & \cdots & p_{3k} \\ \cdots & \cdots & \cdots & \cdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}$$

For example, p_{32} is the probability that the process moves **from** state 3 **to** state 2 in one step.

Transition matrices

- The p_{ij}s are often collected into a matrix P called the *transition matrix*.
- The transition probability p_{ij} is the element of P in the ith row and the jth column, that is, the (i, j)th element of P.
- We will only deal with time-homogeneous Markov chains.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ p_{31} & p_{32} & \cdots & p_{3k} \\ \cdots & \cdots & \cdots & \cdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}$$

For example, p_{32} is the probability that the process moves **from** state 3 **to** state 2 in one step.

Transition matrices

- The p_{ij}s are often collected into a matrix P called the *transition matrix*.
- The transition probability p_{ij} is the element of P in the ith row and the jth column, that is, the (i, j)th element of P.
- We will only deal with time-homogeneous Markov chains.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ p_{31} & p_{32} & \cdots & p_{3k} \\ \cdots & \cdots & \cdots & \cdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}$$

For example, p_{32} is the probability that the process moves **from** state 3 **to** state 2 in one step.

Transition matrices

- The p_{ij}s are often collected into a matrix P called the *transition matrix*.
- The transition probability p_{ij} is the element of P in the ith row and the jth column, that is, the (i, j)th element of P.
- We will only deal with time-homogeneous Markov chains.

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1k} \\ p_{21} & p_{22} & \dots & p_{2k} \\ p_{31} & p_{32} & \dots & p_{3k} \\ \dots & \dots & \dots & \dots \\ p_{k1} & p_{k2} & \dots & p_{kk} \end{bmatrix}$$

For example, p_{32} is the probability that the process moves **from** state 3 **to** state 2 in one step.

Properties of a transition matrix

(i) All entries are non-negative.
   - Also $\leqslant 1$ because they are probabilities.

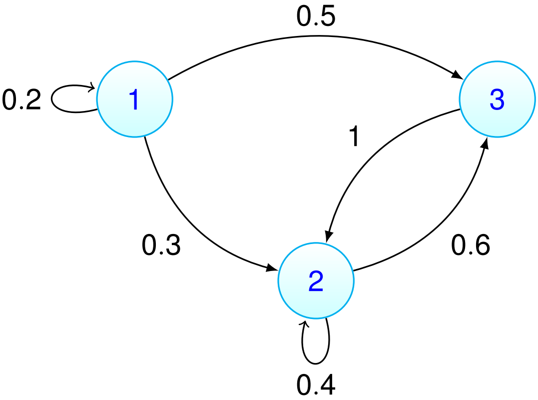
(ii) Each row sums to 1

$$\sum_j p_{ij} = 1, \text{ for all } j \in S.$$

Question: Why does each row sum to 1?

State space diagrams

Helpful to ‘visualise’ our Markov process using a state space diagram (where possible).



n-step transition probabilities

Definition

The n-**step transition probabilities** are

$$p_{ij}^{(n)} = \mathrm{P}(X_n = j | X_0 = i) = \mathrm{P}(X_{m+n} = j | X_m = i),$$

for all m if $\{X_n\}$ is time-homogeneous.

So $p_{ij}^{(n)}$ is the probability that a process in state i will be in state j after n steps.

n-step transition matrices

The *n-step transition matrix* $P^{(n)}$ has $p_{ij}^{(n)}$ as its (i, j)th element.

$$P^{(n)} = \begin{bmatrix} p_{11}^{(n)} & p_{12}^{(n)} & \dots & p_{1k}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} & \dots & p_{2k}^{(n)} \\ p_{31}^{(n)} & p_{32}^{(n)} & \dots & p_{3k}^{(n)} \\ \dots & \dots & \dots & \dots \\ p_{k1}^{(n)} & p_{k2}^{(n)} & \dots & p_{kk}^{(n)} \end{bmatrix}$$

For example, $p_{32}^{(n)}$ is the probability that the process moves **from** state 3 **to** state 2 in n steps.

n-step transition

Big question:

How do we calculate the n-step transition probabilities and/ or matrices?

Answer:

Use the Chapman-Kolmogorov equations.

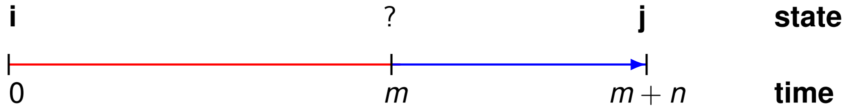
The Chapman-Kolmogorov equations

You try to calculate

$$p_{ij}^{(n+m)} = \mathrm{P}(X_{m+n} = j | X_0 = i).$$

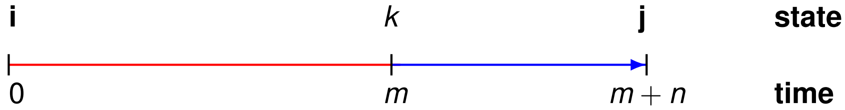
To calculate this probability, we split the timeline into two independent chunks:

- What happens up to time m;
- What happens from time m to time $m + n$.



The Chapman-Kolmogorov equations

Look at **all possible states** k that the Markov chain could have been in at time m...



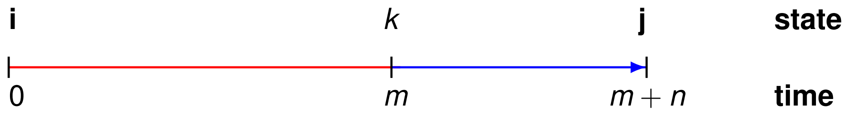
... and sum over all possibilities for k...

... so that

$$p_{ij}^{(n+m)} = \mathrm{P}(X_{m+n} = j | X_0 = i) = \sum_{k \in S} \mathrm{P}(X_{m+n} = j, X_m = k | X_0 = i).$$

The Chapman-Kolmogorov equations

Look at **all possible states** k that the Markov chain could have been in at time m...



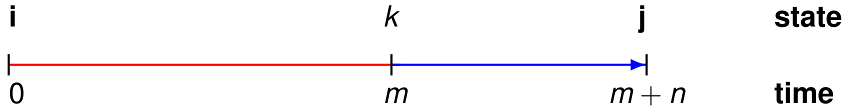
... and sum over all possibilities for k...

... so that

$$p_{ij}^{(n+m)} = \mathrm{P}(X_{m+n} = j | X_0 = i) = \sum_{k \in S} \mathrm{P}(X_{m+n} = j, X_m = k | X_0 = i).$$

The Chapman-Kolmogorov equations

Look at **all possible states** k that the Markov chain could have been in at time m...



... and sum over all possibilities for k...

... so that

$$p_{ij}^{(n+m)} = \mathrm{P}(X_{m+n} = j | X_0 = i) = \sum_{k \in S} \mathrm{P}(X_{m+n} = j, X_m = k | X_0 = i).$$

The Chapman-Kolmogorov equations

By conditioning on the state k at time m we derive $p_{ij}^{(m+n)}$:

$$\begin{aligned} p_{ij}^{(n+m)} &= \mathrm{P}(X_{m+n} = j \mid X_0 = i) \\ &= \sum_{k \in S} \mathrm{P}(X_{m+n} = j, X_m = k \mid X_0 = i) \\ &= \sum_{k \in S} \mathrm{P}(X_{m+n} = j \mid X_m = k, X_0 = i)\, \mathrm{P}(X_m = k \mid X_0 = i) \\ &= \sum_{k \in S} p_{ik}^{(m)}\, p_{kj}^{(n)}, \qquad \text{for all } n, m \geqslant 0, \text{all } i, j \in S. \end{aligned}$$

These equations (one for each (i, j) pair) are called the **Chapman-Kolmogorov equations**.

Interpreting the C-K equations

$$\boxed{p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(m)} \, p_{kj}^{(n)}, \qquad \text{for all } n, m \geqslant 0, \text{all } i, j \in S}$$

Note that $p_{ik}^{(m)} \, p_{kj}^{(n)}$ is the probability that a process:

- is in state i at time 0, AND
- will go to state j in $m + n$ steps, AND
- passes through state k at time m.

Summing over all possible values for k gives us the probability we require.

Interpreting the C-K equations

The derivation of the Chapman-Kolmogorov equations uses

- the Markov property;
- time homogeneity of transition probabilities.

Easy exercise: find where each of these assumptions is used in the derivation of the C-K equations.

Matrix version of the C-K equations

Since

$$\{p_{ik}^{(m)}, \, k \in S\} \text{ is the } i\text{th row of } P^{(m)}$$

and

$$\{p_{kj}^{(n)}, \, k \in S\} \text{ is the } j\text{th column of } P^{(n)},$$

the Chapman-Kolmogorov equations can be written in matrix form as

$$\boxed{P^{(m+n)} = P^{(m)} \cdot P^{(n)}}$$

where $\cdot$ denotes **matrix multiplication.**

How do the C-K equations help us to calculate $P^{(n)}$?

RECALL:

Big question:

How do we calculate the n-step transition probabilities and/ or matrices?

Answer:

Use the Chapman-Kolmogorov equations.

But how do the C-K equations help?

How do the C-K equations help us to calculate $P^{(n)}$?

NOTE: the C-K equations tell us that

$$P^{(m+n)} = P^{(m)} \cdot P^{(n)}.$$

In addition, we have:

1. $P^{(1)}$ is the transition matrix P.
2. $P^{(0)}$ is the identity matrix I.

So, starting with $n = 2$,

$$P^{(2)} = P^{(1+1)} = \boxed{P^{(1)} \cdot P^{(1)} = P \cdot P} = P^2.$$

And for $n = 3$,

$$P^{(3)} = P^{(1+2)} = \boxed{P^{(1)} \cdot P^{(2)} = P \cdot P^2} = P^3.$$

How do the C-K equations help us to calculate $P^{(n)}$?

So with repeated applications of the Chapman-Kolmogorov equations, we find that:

$$\boxed{P^{(n)} = P^n.}$$

That is, the n-step transition matrix $P^{(n)}$ is equal to the nth matrix power of P.

Example - modelling HIV progression

See paper on Moodle

- Six-monthly progression of HIV infected subjects, with high CD4-cell depletion.
- At each 6 month interval, patients are categorised as having a CD4-cell count which is:
  - 0-49 cells/mm^3 ('Low')
  - 50-74 cells/mm^3 ('Med')
  - 75+ cells/mm^3 ('High')
- X_n= CD4-cell count category at time point n.

Markov chain assumption:

> A patient's CD4-cell count today is dependent only on their last CD4-cell count (measured 6 months ago).

That is, $X_{n+1} \perp\!\!\!\perp \{X_{n-1}, \ldots, X_0\} \mid X_n$.

Example - modelling HIV progression

See paper on Moodle

- Six-monthly progression of HIV infected subjects, with high CD4-cell depletion.
- At each 6 month interval, patients are categorised as having a CD4-cell count which is:
  - 0-49 cells/mm^3 (‘Low’)
  - 50-74 cells/mm^3 (‘Med’)
  - 75+ cells/mm^3 (‘High’)
- X_n= CD4-cell count category at time point n.

Markov chain assumption:

> A patient’s CD4-cell count today is dependent only on their last CD4-cell count (measured 6 months ago).

That is, $X_{n+1} \perp\!\!\!\perp \{X_{n-1}, \ldots, X_0\} \mid X_n$.

Example - modelling HIV progression

See paper on Moodle

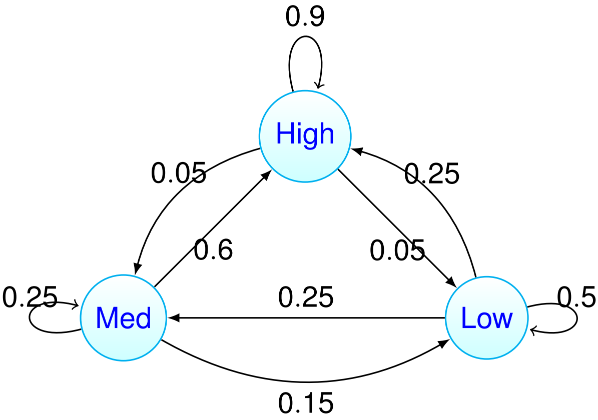
- Six-monthly progression of HIV infected subjects, with high CD4-cell depletion.
- At each 6 month interval, patients are categorised as having a CD4-cell count which is:
  - 0-49 cells/mm^3 ('Low')
  - 50-74 cells/mm^3 ('Med')
  - 75+ cells/mm^3 ('High')
- X_n= CD4-cell count category at time point n.

Markov chain assumption:

> A patient's CD4-cell count today is dependent only on their last CD4-cell count (measured 6 months ago).

That is, $X_{n+1} \perp\!\!\!\perp \{X_{n-1}, \ldots, X_0\} \mid X_n$.

Example - modelling HIV progression



(a) Find the 2-step transition matrix.

(b) Find $\mathrm{P}(X_4 = \text{High}, X_2 = \text{High} \mid X_0 = \text{High})$.

Example - modelling HIV progression

Transition matrix is

$$P = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.6 & 0.25 & 0.15 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

So the two-step transition matrix is (via C-K equations)

$$P^{(2)} = P^2 = \begin{bmatrix} 0.8525 & 0.07 & 0.0775 \\ 0.7275 & 0.13 & 0.1425 \\ 0.5 & 0.2 & 0.3 \end{bmatrix}$$

$$\begin{aligned} &\mathrm{P}(X_4 = \mathrm{High}, X_2 = \mathrm{High} \mid X_0 = \mathrm{High}) \\ =&\mathrm{P}(X_4 = \mathrm{High} | X_2 = \mathrm{High}, X_0 = \mathrm{High})\mathrm{P}(X_2 = \mathrm{High} \mid X_0 = \mathrm{High}) \\ =&\mathrm{P}(X_4 = \mathrm{High} | X_2 = \mathrm{High})\mathrm{P}(X_2 = \mathrm{High} \mid X_0 = \mathrm{High}) \\ =&0.8525 \times 0.8525 \simeq 0.73 \end{aligned}$$

Example - modelling HIV progression

Transition matrix is

$$P = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.6 & 0.25 & 0.15 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

So the two-step transition matrix is (via C-K equations)

$$P^{(2)} = P^2 = \begin{bmatrix} 0.8525 & 0.07 & 0.0775 \\ 0.7275 & 0.13 & 0.1425 \\ 0.5 & 0.2 & 0.3 \end{bmatrix}$$

$$\begin{aligned} & \mathrm{P}(X_4 = \text{High}, X_2 = \text{High} \mid X_0 = \text{High}) \\ =& \mathrm{P}(X_4 = \text{High} | X_2 = \text{High}, X_0 = \text{High}) \mathrm{P}(X_2 = \text{High} \mid X_0 = \text{High}) \\ =& \mathrm{P}(X_4 = \text{High} | X_2 = \text{High}) \mathrm{P}(X_2 = \text{High} \mid X_0 = \text{High}) \\ =& 0.8525 \times 0.8525 \simeq 0.73 \end{aligned}$$

Example - modelling HIV progression

Transition matrix is

$$P = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.6 & 0.25 & 0.15 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

So the two-step transition matrix is (via C-K equations)

$$P^{(2)} = P^2 = \begin{bmatrix} 0.8525 & 0.07 & 0.0775 \\ 0.7275 & 0.13 & 0.1425 \\ 0.5 & 0.2 & 0.3 \end{bmatrix}$$

$$\begin{aligned} & \mathrm{P}(X_4 = \text{High}, X_2 = \text{High} \mid X_0 = \text{High}) \\ =& \mathrm{P}(X_4 = \text{High} | X_2 = \text{High}, X_0 = \text{High}) \mathrm{P}(X_2 = \text{High} \mid X_0 = \text{High}) \\ =& \mathrm{P}(X_4 = \text{High} | X_2 = \text{High}) \mathrm{P}(X_2 = \text{High} \mid X_0 = \text{High}) \\ =& 0.8525 \times 0.8525 \simeq 0.73 \end{aligned}$$

Conditional vs marginal probabilities

So far, we have only considered *conditional* probabilities.

However, we may also be interested in unconditional or marginal probabilities.

What is the probability that the chain is in state j at time n?

That is, what is $\mathrm{P}(X_n = j)$?

Conditional vs marginal probabilities

So far, we have only considered *conditional* probabilities.

However, we may also be interested in unconditional or marginal probabilities.

What is the probability that the chain is in state j at time n?

That is, what is $\mathrm{P}(X_n = j)$?

Marginal probabilities and the initial distribution

Let $p_j^{(n)} = \mathrm{P}(X_n = j)$. If, for example, $S = \mathbb{N}_0 = \{0, 1, 2, \ldots\}$, then

$$\underline{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \ldots)$$

is a *probability row vector* (i.e. a row vector with non-negative entries summing to 1) specifying the distribution of X_n.

When $n = 0$:

- $\underline{p}^{(0)}$ is the **distribution** of X_0;
- $\underline{p}^{(0)}$ is called the **initial distribution.**

Calculating marginal probabilities

Remark

Do not confuse the notation $p_{ij}^{(n)}$ and $p_i^{(n)}$. The former is the **conditional probability** of the chain moving from state i to state j in n steps, while the latter is the **marginal probability** of the chain being in state i at time n.

How do we calculate the marginal probabilities, $p_j^{(n)}$?

$$
\begin{aligned}
p_j^{(n)} &= \mathrm{P}(X_n = j) = \sum_i \mathrm{P}(X_n = j \mid X_0 = i)\,\mathrm{P}(X_0 = i) \\
&= \sum_i p_i^{(0)} p_{ij}^{(n)}.
\end{aligned}
$$

In matrix notation, $\underline{p}^{(n)} = \underline{p}^{(0)} \cdot P^{(n)}$.

Calculating marginal probabilities

Remark

Do not confuse the notation $p_{ij}^{(n)}$ and $p_i^{(n)}$. The former is the **conditional probability** of the chain moving from state i to state j in n steps, while the latter is the **marginal probability** of the chain being in state i at time n.

How do we calculate the marginal probabilities, $p_j^{(n)}$?

$$
\begin{aligned}
p_j^{(n)} &= \mathrm{P}(X_n = j) = \sum_i \mathrm{P}(X_n = j \mid X_0 = i)\,\mathrm{P}(X_0 = i) \\
&= \sum_i p_i^{(0)} p_{ij}^{(n)}.
\end{aligned}
$$

In matrix notation, $\underline{p}^{(n)} = \underline{p}^{(0)} \cdot P^{(n)}$.

Calculating marginal probabilities

Remark

Do not confuse the notation $p_{ij}^{(n)}$ and $p_i^{(n)}$. The former is the **conditional probability** of the chain moving from state i to state j in n steps, while the latter is the **marginal probability** of the chain being in state i at time n.

How do we calculate the marginal probabilities, $p_j^{(n)}$?

$$\begin{aligned} p_j^{(n)} &= \mathrm{P}(X_n = j) = \sum_i \mathrm{P}(X_n = j \mid X_0 = i)\,\mathrm{P}(X_0 = i) \\ &= \sum_i p_i^{(0)} p_{ij}^{(n)}. \end{aligned}$$

In matrix notation, $\underline{p}^{(n)} = \underline{p}^{(0)} \cdot P^{(n)}$.

Putting it all together

Thus, the initial distribution $\underline{p}^{(0)}$ and the transition matrix P together contain all the probabilistic information about the chain.

- Conditional one-step probabilities are contained in P;
- Conditional n-step probabilities are contained in $P^{(n)} = P^n$;
- The initial distribution of the chain is contained in $\underline{p}^{(0)}$;
- The marginal probabilities, $p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}$;
- The distribution of X_n is $\underline{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \dots)$.

For example

$$\mathrm{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = p_{i_0}^{(0)} \, p_{i_0 i_1} \, p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Putting it all together

Thus, the initial distribution $\underline{p}^{(0)}$ and the transition matrix P together contain all the probabilistic information about the chain.

- Conditional one-step probabilities are contained in P;
- Conditional n-step probabilities are contained in $P^{(n)} = P^n$;
- The initial distribution of the chain is contained in $\underline{p}^{(0)}$;
- The marginal probabilities, $p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}$;
- The distribution of X_n is $\underline{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \dots)$.

For example

$$\mathrm{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = p_{i_0}^{(0)} \, p_{i_0 i_1} \, p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Putting it all together

Thus, the initial distribution $\underline{p}^{(0)}$ and the transition matrix P together contain all the probabilistic information about the chain.

- Conditional one-step probabilities are contained in P;
- Conditional n-step probabilities are contained in $P^{(n)} = P^n$;
- The initial distribution of the chain is contained in $\underline{p}^{(0)}$;
- The marginal probabilities, $p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}$;
- The distribution of X_n is $\underline{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \ldots)$.

For example

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = p_{i_0}^{(0)} \, p_{i_0 i_1} \, p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Putting it all together

Thus, the initial distribution $\underline{p}^{(0)}$ and the transition matrix P together contain all the probabilistic information about the chain.

- Conditional one-step probabilities are contained in P;
- Conditional n-step probabilities are contained in $P^{(n)} = P^n$;
- The initial distribution of the chain is contained in $\underline{p}^{(0)}$;
- The marginal probabilities, $p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}$;
- The distribution of X_n is $\underline{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \ldots)$.

For example

$$\mathrm{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = p_{i_0}^{(0)} \, p_{i_0 i_1} \, p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Putting it all together

Thus, the initial distribution $\underline{p}^{(0)}$ and the transition matrix P together contain all the probabilistic information about the chain.

- Conditional one-step probabilities are contained in P;
- Conditional n-step probabilities are contained in $P^{(n)} = P^n$;
- The initial distribution of the chain is contained in $\underline{p}^{(0)}$;
- The marginal probabilities, $p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}$;
- The distribution of X_n is $\underline{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \ldots)$.

For example

$$\mathrm{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = p_{i_0}^{(0)} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Putting it all together

Thus, the initial distribution $\underline{p}^{(0)}$ and the transition matrix P together contain all the probabilistic information about the chain.

- Conditional one-step probabilities are contained in P;
- Conditional n-step probabilities are contained in $P^{(n)} = P^n$;
- The initial distribution of the chain is contained in $\underline{p}^{(0)}$;
- The marginal probabilities, $p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}$;
- The distribution of X_n is $\underline{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \ldots)$.

For example

$$\mathrm{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = p_{i_0}^{(0)} \, p_{i_0 i_1} \, p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Putting it all together

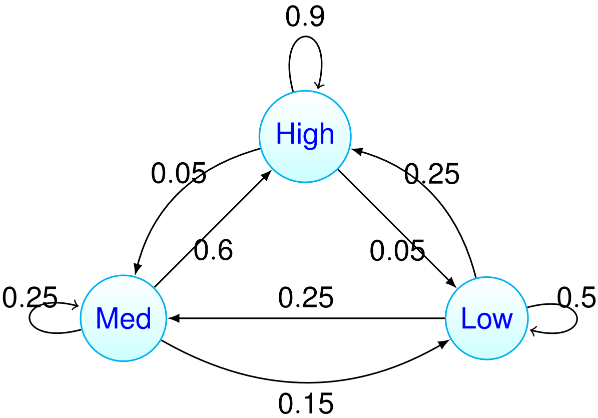
Thus, the initial distribution $\underline{p}^{(0)}$ and the transition matrix P together contain all the probabilistic information about the chain.

- Conditional one-step probabilities are contained in P;
- Conditional n-step probabilities are contained in $P^{(n)} = P^n$;
- The initial distribution of the chain is contained in $\underline{p}^{(0)}$;
- The marginal probabilities, $p_j^{(n)} = \sum_i p_i^{(0)} p_{ij}^{(n)}$;
- The distribution of X_n is $\underline{p}^{(n)} = (p_0^{(n)}, p_1^{(n)}, \ldots)$.

For example

$$\mathrm{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = p_{i_0}^{(0)} \, p_{i_0 i_1} \, p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Example - HIV progression again



Suppose $p^{(0)}_{\text{High}} = 0.7$, $p^{(0)}_{\text{Med}} = 0.2$ and $p^{(0)}_{\text{Low}} = 0.1$.

Find $\underline{p}^{(0)}, \underline{p}^{(1)}$ and $\underline{p}^{(2)}$.

Example - HIV progression again

$$P = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.6 & 0.25 & 0.15 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

By C-K equations:

$$P^{(2)} = P^2 = \begin{bmatrix} 0.8525 & 0.07 & 0.0775 \\ 0.7275 & 0.13 & 0.1425 \\ 0.5 & 0.2 & 0.3 \end{bmatrix}$$

Suppose $p^{(0)}_{\text{High}} = 0.7$, $p^{(0)}_{\text{Med}} = 0.2$ and $p^{(0)}_{\text{Low}} = 0.1$.

$$\underline{p}^{(0)} = (0.7, 0.2, 0.1)$$
$$\underline{p}^{(1)} = \underline{p}^{(0)} P = (0.775, 0.11, 0.115)$$
$$\underline{p}^{(2)} = \underline{p}^{(0)} P^2 = (0.79225, 0.095, 0.11275)$$

Example - HIV progression again

$$P = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.6 & 0.25 & 0.15 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

By C-K equations:

$$P^{(2)} = P^2 = \begin{bmatrix} 0.8525 & 0.07 & 0.0775 \\ 0.7275 & 0.13 & 0.1425 \\ 0.5 & 0.2 & 0.3 \end{bmatrix}$$

Suppose $p^{(0)}_{\text{High}} = 0.7$, $p^{(0)}_{\text{Med}} = 0.2$ and $p^{(0)}_{\text{Low}} = 0.1$.

$$\underline{p}^{(0)} = (0.7, 0.2, 0.1)$$

$$\underline{p}^{(1)} = \underline{p}^{(0)} P = (0.775, 0.11, 0.115)$$

$$\underline{p}^{(2)} = \underline{p}^{(0)} P^2 = (0.79225, 0.095, 0.11275)$$

Example - HIV progression again

$$P = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.6 & 0.25 & 0.15 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$$

By C-K equations:

$$P^{(2)} = P^2 = \begin{bmatrix} 0.8525 & 0.07 & 0.0775 \\ 0.7275 & 0.13 & 0.1425 \\ 0.5 & 0.2 & 0.3 \end{bmatrix}$$

Suppose $p_{\text{High}}^{(0)} = 0.7$, $p_{\text{Med}}^{(0)} = 0.2$ and $p_{\text{Low}}^{(0)} = 0.1$.

$$\begin{aligned} \underline{p}^{(0)} &= (0.7, 0.2, 0.1) \\ \underline{p}^{(1)} &= \underline{p}^{(0)} P = (0.775, 0.11, 0.115) \\ \underline{p}^{(2)} &= \underline{p}^{(0)} P^2 = (0.79225, 0.095, 0.11275) \end{aligned}$$

How long will the chain stay in a particular state?

For any Markov chain the duration of stay in any state i satisfies

$$\begin{aligned} &\mathrm{P}(\text{stay in state } i \text{ lasts exactly } n \text{ time steps} | \text{chain is in state } i) \\ =&\mathrm{P}(\text{stay in } i \text{ for a further } n-1 \text{ steps, then leave } i | \text{in } i) \\ =&p_{ii}^{n-1}(1-p_{ii}). \end{aligned}$$

Remark

This is a geometric pmf. Therefore, the duration of stay in state i has a **geometric distribution**,

$$\text{Duration of stay in state } i \sim \mathrm{Geom}(1-p_{ii}).$$

Therefore, the expected time the chain stays in state i is

$$\mathrm{E}[\text{time chain stays in state } i] = \frac{1}{1-p_{ii}}.$$

How long will the chain stay in a particular state?

For any Markov chain the duration of stay in any state i satisfies

$$\begin{aligned} &\mathrm{P}(\text{stay in state } i \text{ lasts exactly } n \text{ time steps} | \text{chain is in state } i) \\ =&\mathrm{P}(\text{stay in } i \text{ for a further } n-1 \text{ steps, then leave } i | \text{in } i) \\ =&p_{ii}^{n-1}(1-p_{ii}). \end{aligned}$$

Remark

This is a geometric pmf. Therefore, the duration of stay in state i has a **geometric distribution**,

$$\text{Duration of stay in state } i \sim \mathrm{Geom}(1-p_{ii}).$$

Therefore, the expected time the chain stays in state i is

$$\mathrm{E}[\text{time chain stays in state } i] = \frac{1}{1-p_{ii}}.$$

Where will the chain go when it leaves a state?

Where will the Markov chain go when it leaves i? For $j \neq i$,

$$
\begin{aligned}
&\mathrm{P}(\text{MC moves to } j | \text{MC leaves state } i) \\
=&\mathrm{P}(X_{n+1} = j | X_n = i, X_{n+1} \neq i) \\
=&\frac{\mathrm{P}(X_{n+1} = j, X_{n+1} \neq i | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)} \\
=&\frac{\mathrm{P}(X_{n+1} = j | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)} \\
=&\frac{p_{ij}}{1 - p_{ii}}.
\end{aligned}
$$

This answer makes sense: the probability is proportional to p_{ij} and is scaled so that it sums to 1 over $j \neq i$.

Where will the chain go when it leaves a state?

Where will the Markov chain go when it leaves i? For $j \neq i$,

$$
\begin{aligned}
&\mathrm{P}(\text{MC moves to } j | \text{MC leaves state } i) \\
=&\mathrm{P}(X_{n+1} = j | X_n = i, X_{n+1} \neq i) \\
=&\frac{\mathrm{P}(X_{n+1} = j, X_{n+1} \neq i | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)} \\
=&\frac{\mathrm{P}(X_{n+1} = j | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)} \\
=&\frac{p_{ij}}{1 - p_{ii}}.
\end{aligned}
$$

This answer makes sense: the probability is proportional to p_{ij} and is scaled so that it sums to 1 over $j \neq i$.

Where will the chain go when it leaves a state?

Where will the Markov chain go when it leaves i? For $j \neq i$,

$$
\begin{aligned}
&\mathrm{P}(\text{MC moves to } j | \text{MC leaves state } i) \\
=&\mathrm{P}(X_{n+1} = j | X_n = i, X_{n+1} \neq i) \\
=&\frac{\mathrm{P}(X_{n+1} = j, X_{n+1} \neq i | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)} \\
=&\frac{\mathrm{P}(X_{n+1} = j | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)} \\
=&\frac{p_{ij}}{1 - p_{ii}}.
\end{aligned}
$$

This answer makes sense: the probability is proportional to p_{ij} and is scaled so that it sums to 1 over $j \neq i$.

Where will the chain go when it leaves a state?

Where will the Markov chain go when it leaves i? For $j \neq i$,

$$
\begin{aligned}
&\mathrm{P}(\text{MC moves to } j \mid \text{MC leaves state } i) \\
=&\mathrm{P}(X_{n+1} = j \mid X_n = i, X_{n+1} \neq i) \\
=&\frac{\mathrm{P}(X_{n+1} = j, X_{n+1} \neq i \mid X_n = i)}{\mathrm{P}(X_{n+1} \neq i \mid X_n = i)} \\
=&\frac{\mathrm{P}(X_{n+1} = j \mid X_n = i)}{\mathrm{P}(X_{n+1} \neq i \mid X_n = i)} \\
=&\frac{p_{ij}}{1 - p_{ii}}.
\end{aligned}
$$

This answer makes sense: the probability is proportional to p_{ij} and is scaled so that it sums to 1 over $j \neq i$.

Where will the chain go when it leaves a state?

Where will the Markov chain go when it leaves i? For $j \neq i$,

$$
\begin{aligned}
&\mathrm{P}(\text{MC moves to } j | \text{MC leaves state } i)\\
=&\mathrm{P}(X_{n+1} = j | X_n = i, X_{n+1} \neq i)\\
=&\frac{\mathrm{P}(X_{n+1} = j, X_{n+1} \neq i | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)}\\
=&\frac{\mathrm{P}(X_{n+1} = j | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)}\\
=&\frac{p_{ij}}{1 - p_{ii}}.
\end{aligned}
$$

This answer makes sense: the probability is proportional to p_{ij} and is scaled so that it sums to 1 over $j \neq i$.

Where will the chain go when it leaves a state?

Where will the Markov chain go when it leaves i? For $j \neq i$,

$$
\begin{aligned}
& \mathrm{P}(\text{MC moves to } j | \text{MC leaves state } i) \\
= & \mathrm{P}(X_{n+1} = j | X_n = i, X_{n+1} \neq i) \\
= & \frac{\mathrm{P}(X_{n+1} = j, X_{n+1} \neq i | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)} \\
= & \frac{\mathrm{P}(X_{n+1} = j | X_n = i)}{\mathrm{P}(X_{n+1} \neq i | X_n = i)} \\
= & \frac{p_{ij}}{1 - p_{ii}}.
\end{aligned}
$$

This answer makes sense: the probability is proportional to p_{ij} and is scaled so that it sums to 1 over $j \neq i$.