Markov chains and applications

Dr. András Pongrácz

Week 9, University of Debrecen

Absorbing Chains

Revision

A state is absorbing if we stay in it with probability one in the next move.

Revision

A state is absorbing if we stay in it with probability one in the next move.

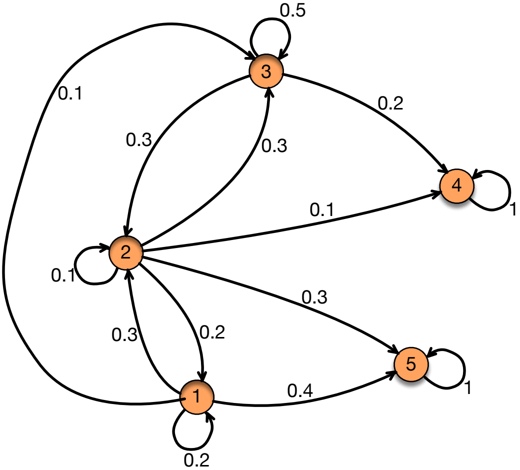
A finite Markov chain is absorbing if, from every initial state, we can reach an absorbing state with positive probability in finitely many steps.

Revision

A state is absorbing if we stay in it with probability one in the next move.

A finite Markov chain is absorbing if, from every initial state, we can reach an absorbing state with positive probability in finitely many steps.

The non-absorbing states in such chains are called transient; cf. the last class for a more precise introduction to transient states.


0.5
3
0.1
0.2
0.3
0.3
4
0.1
1
2
0.1
0.3
0.3
0.2
5
1
0.4
1
0.2

Markov chains

Transition matrix:

$$
\begin{array}{c}
\begin{array}{ccccc} Q & & & & R \end{array} \\
\begin{pmatrix}
0.2 & 0.3 & 0.1 & 0 & 0.4 \\
0.2 & 0.1 & 0.3 & 0.1 & 0.3 \\
0 & 0.3 & 0.5 & 0.2 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}
$$

$N = (I - Q)^{-1}$

fundamental matrix

Transition matrix:

$$\begin{array}{c} \begin{array}{ccccc} Q & & & & R \end{array} \\ \left(\begin{array}{ccc|cc} 0.2 & 0.3 & 0.1 & 0 & 0.4 \\ 0.2 & 0.1 & 0.3 & 0.1 & 0.3 \\ 0 & 0.3 & 0.5 & 0.2 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \end{array}$$

$N = (I - Q)^{-1}$
fundamental matrix

(Given an initial state) what is the expected time to absorption?

Markov chains

Transition matrix:

$$
\begin{array}{c}
\begin{array}{ccccc} Q & & & & R \end{array} \\
\begin{pmatrix}
0.2 & 0.3 & 0.1 & 0 & 0.4 \\
0.2 & 0.1 & 0.3 & 0.1 & 0.3 \\
0 & 0.3 & 0.5 & 0.2 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}
$$

$N = (I - Q)^{-1}$
fundamental matrix

(Given an initial state) what is the expected time to absorption?

(Given an initial state) what is the probability of absorption at each absorbing state?

Markov chains

Transition matrix:

$$\begin{array}{c}
\begin{array}{ccccc} \color{red}{Q} & & & & \color{blue}{R} \end{array} \\
\begin{pmatrix}
0.2 & 0.3 & 0.1 & 0 & 0.4 \\
0.2 & 0.1 & 0.3 & 0.1 & 0.3 \\
0 & 0.3 & 0.5 & 0.2 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}$$

$N = (I - Q)^{-1}$
fundamental matrix

(Given an initial state) what is the expected time to absorption?
Answer: $N\underline{\mathbf{1}}$

(Given an initial state) what is the probability of absorption at each absorbing state? Answer: NR

Sketch of arguments

Claim

$$\lim_{k\to\infty} Q^k = 0$$

Sketch of arguments

Claim

$$\lim_{k \to \infty} Q^k = 0$$

with an exponential convergence speed (at each position, on the long-run, the moduli of terms are below a given geometric series with quotient $0 < q < 1$).

Sketch of arguments

Claim

$$\lim_{k \to \infty} Q^k = 0$$

with an exponential convergence speed (at each position, on the long-run, the moduli of terms are below a given geometric series with quotient $0 < q < 1$).

Proof: By definition of an absorbing chain, for a large enough K we have that the sum of terms in each row of Q^K is at most r for some $0 < r < 1$.

Sketch of arguments

Claim

$$\lim_{k\to\infty} Q^k = 0$$

with an exponential convergence speed (at each position, on the long-run, the moduli of terms are below a given geometric series with quotient $0 < q < 1$).

Proof: By definition of an absorbing chain, for a large enough K we have that the sum of terms in each row of Q^K is at most r for some $0 < r < 1$. By Gershgorin's theorem, the spectrum of Q^K is in $[-r, r]$.

Sketch of arguments

Claim

$$\lim_{k \to \infty} Q^k = 0$$

with an exponential convergence speed (at each position, on the long-run, the moduli of terms are below a given geometric series with quotient $0 < q < 1$).

Proof: By definition of an absorbing chain, for a large enough K we have that the sum of terms in each row of Q^K is at most r for some $0 < r < 1$. By Gershgorin's theorem, the spectrum of Q^K is in $[-r, r]$. If $k \geq mK$, the spectrum is in $[-r^m, r^m]$; making any $\sqrt[K]{r} < q < 1$ a good choice.

Corollary I

The fundamental matrix $N = (I - Q)^{-1}$ exists, and equals to $\sum_{k=0}^{\infty} Q^k$.

Corollary I

The fundamental matrix $N = (I - Q)^{-1}$ exists, and equals to $\sum_{k=0}^{\infty} Q^k$.

Corollary II

The probability of absorption in a finite number of steps from any initial state is 1.

Expected runtime

The entry $Q^k[i,j]$ is the probability that, starting from the i-th state, we are at the j-th state after k steps.

The entry $Q^k[i,j]$ is the probability that, starting from the i-th state, we are at the j-th state after k steps.

Thus $N\underline{1} = (I - Q)^{-1}\underline{1} = (\sum_{k=0}^{\infty} Q^k)\underline{1}$ is the expected number of times the walk is in a non-absorbing state.

Expected runtime

The entry $Q^k[i,j]$ is the probability that, starting from the i-th state, we are at the j-th state after k steps.

Thus $N\underline{1} = (I-Q)^{-1}\underline{1} = (\sum_{k=0}^{\infty} Q^k)\underline{1}$ is the expected number of times the walk is in a non-absorbing state.

Alternative proof: we can find a system of linear equations corresponding to the law of total expectation.

Expected runtime

The entry $Q^k[i,j]$ is the probability that, starting from the i-th state, we are at the j-th state after k steps.

Thus $N\underline{1} = (I-Q)^{-1}\underline{1} = (\sum_{k=0}^{\infty} Q^k)\underline{1}$ is the expected number of times the walk is in a non-absorbing state.

Alternative proof: we can find a system of linear equations corresponding to the law of total expectation. It is

$$\underline{x} = Q\underline{x} + \underline{1}$$

Expected runtime

The entry $Q^k[i,j]$ is the probability that, starting from the i-th state, we are at the j-th state after k steps.

Thus $N\underline{1} = (I - Q)^{-1}\underline{1} = (\sum_{k=0}^{\infty} Q^k)\underline{1}$ is the expected number of times the walk is in a non-absorbing state.

Alternative proof : we can find a system of linear equations corresponding to the law of total expectation. It is

$$\underline{x} = Q\underline{x} + \underline{1}$$

yielding $\underline{x} = N\underline{1}$.

Generalization

The same method works in a more general situation.

Generalization

The same method works in a more general situation.

If $\underline{v}$ is a vector of numbers assigned to all transient states,

Generalization

The same method works in a more general situation.

If $\underline{v}$ is a vector of numbers assigned to all transient states, then $N\underline{v}$ is the expectation of the sum of all values of entries during a random walk until absorption.

Generalization

The same method works in a more general situation.

If $\underline{v}$ is a vector of numbers assigned to all transient states, then $N\underline{v}$ is the expectation of the sum of all values of entries during a random walk until absorption.

The proof is based on the law of total probability. (Cf. the exrecizes.)

Absorption probabilities

Starting from state i, what is the probability of absorption is each state?

Absorption probabilities

Starting from state i, what is the probability of absorption is each state? The total for all absorbing states is 1.

Absorption probabilities

Starting from state i, what is the probability of absorption is each state? The total for all absorbing states is 1.

Law of total probability: if $\underline{x}$ is the vector of results for all transient states, then

$$\underline{x} = Q\underline{x} + R$$

Absorption probabilities

Starting from state i, what is the probability of absorption is each state? The total for all absorbing states is 1.

Law of total probability: if $\underline{x}$ is the vector of results for all transient states, then

$$\underline{x} = Q\underline{x} + R$$

yielding $\underline{x} = NR$.

Higher moments

Moment generating funciton

Given a random variable X, its moment generating function is

$$t \mapsto E(e^{tX})$$

(wherever it is defined).

Theorem

If the moment generating function is defined in a neighborhood of 0, then

$$\left(\frac{d^n}{dt^n} E(e^{tX}) \right)_{t=0} = E(X^n)$$

is the n-th moment of X.

In order to use the law of total expectation (similarly as before), it is better to create a vector of random variables.

In order to use the law of total expectation (similarly as before), it is better to create a vector of random variables.

$X_1, X_2, \ldots, X_k$ are the number of steps in the walks starting from each transient state.

In order to use the law of total expectation (similarly as before), it is better to create a vector of random variables.

$X_1, X_2, \ldots, X_k$ are the number of steps in the walks starting from each transient state.

$$\underline{v}(t) = (E(e^{tX_1}), E(e^{tX_2}), \ldots, E(e^{tX_k}))^*$$

In order to use the law of total expectation (similarly as before), it is better to create a vector of random variables.

$X_1, X_2, \ldots, X_k$ are the number of steps in the walks starting from each transient state.

$$\underline{v}(t) = (E(e^{tX_1}), E(e^{tX_2}), \ldots, E(e^{tX_k}))^*$$

We obtain the system of linear equations

$$\underline{v}(t) = e^t(Q\underline{v}(t) + (I - Q)\underline{1})$$

Rearranging:

$$(I - e^t Q)\underline{v}(t) = e^t (I - Q)\underline{1}$$

Rearranging:

$$(I - e^t Q)\underline{v}(t) = e^t (I - Q)\underline{1}$$

The matrix $(I - e^t Q)$ is singular for only a finite number of values of t.

Rearranging:

$$(I - e^t Q)\underline{v}(t) = e^t (I - Q)\underline{1}$$

The matrix $(I - e^t Q)$ is singular for only a finite number of values of t.

$$\underline{v}(t) = (I - e^t Q)^{-1} e^t (I - Q)\underline{1}$$

Rearranging:

$$(I - e^t Q)\underline{v}(t) = e^t (I - Q)\underline{1}$$

The matrix $(I - e^t Q)$ is singular for only a finite number of values of t.

$$\underline{v}(t) = (I - e^t Q)^{-1} e^t (I - Q)\underline{1}$$

yielding

$$\underline{v}(t) = e^t (I + e^t Q + e^{2t} Q^2 + e^{3t} Q^3 + \cdots)(I - Q)\underline{1}$$

Rearranging:

$$(I - e^t Q)\underline{v}(t) = e^t (I - Q)\underline{1}$$

The matrix $(I - e^t Q)$ is singular for only a finite number of values of t.

$$\underline{v}(t) = (I - e^t Q)^{-1} e^t (I - Q)\underline{1}$$

yielding

$$\underline{v}(t) = e^t (I + e^t Q + e^{2t} Q^2 + e^{3t} Q^3 + \cdots)(I - Q)\underline{1}$$

The first derivative at zero is
$\underline{v}'(0) = (I + 2Q + 3Q^2 + 4Q^3 + \cdots)(I - Q)\underline{1} = (I - Q)^{-2}(I - Q)\underline{1} = N\underline{1}$.

In general

Similar calculation, yielding more complicated power series.

In general

Similar calculation, yielding more complicated power series. They are always linear combinations of powers of $(I - Q)^{-m}$, multiplied by $(I - Q)\underline{1}$.

In general

Similar calculation, yielding more complicated power series. They are always linear combinations of powers of $(I - Q)^{-m}$, multiplied by $(I - Q)\underline{1}$.

Lemma 1: $((1 - x)^{-1})^{(m)} = n! \cdot (1 - x)^{-(m+1)}$

In general

Similar calculation, yielding more complicated power series. They are always linear combinations of powers of $(I - Q)^{-m}$, multiplied by $(I - Q)\underline{1}$.

Lemma 1: $((1-x)^{-1})^{(m)} = n! \cdot (1-x)^{-(m+1)}$

Lemma 2: $((1-x)^{-1})^{(m)} = \sum_{j=0}^{\infty} (j+1)\cdots(j+m)x^j$

In general

Similar calculation, yielding more complicated power series. They are always linear combinations of powers of $(I - Q)^{-m}$, multiplied by $(I - Q)\underline{1}$.

Lemma 1: $((1-x)^{-1})^{(m)} = n! \cdot (1-x)^{-(m+1)}$

Lemma 2: $((1-x)^{-1})^{(m)} = \sum_{j=0}^{\infty} (j+1)\cdots(j+m)x^j$

Theorem

$\underline{v}^{(m)}(0) = f_m(N)\underline{1}$, where $f_m(N)$ is a degree m polynomial of N, whose constant term is 0 and the main coefficient is $m!$.

In general

Similar calculation, yielding more complicated power series. They are always linear combinations of powers of $(I-Q)^{-m}$, multiplied by $(I-Q)\underline{1}$.

Lemma 1: $((1-x)^{-1})^{(m)} = n! \cdot (1-x)^{-(m+1)}$

Lemma 2: $((1-x)^{-1})^{(m)} = \sum_{j=0}^{\infty} (j+1)\cdots(j+m)x^j$

Theorem

$\underline{v}^{(m)}(0) = f_m(N)\underline{1}$, where $f_m(N)$ is a degree m polynomial of N, whose constant term is 0 and the main coefficient is $m!$.

$f_1(N) = N$, $f_2(N) = 2N^2 - N$, $f_3(N) = 6N^3 - 6N^2 + N$, $f_4(N) = 24N^4 - 36N^2 + 14N - N$ can be easily computed from the above.

Exercises

Exercises

1. Show that the fundamental matrix N is non-negative.

Exercises

1. Show that the fundamental matrix N is non-negative.
2. Compute the fundamental matrix, and answer the basic problems for the absorbing chain provided in the talk as an example.

Exercises

1. Show that the fundamental matrix N is non-negative.
2. Compute the fundamental matrix, and answer the basic problems for the absorbing chain provided in the talk as an example.
3. Compute the first four moments of the same chain with the general method presented. Compare the first moment with the result of Problem 1.

1. Show that the fundamental matrix N is non-negative.
2. Compute the fundamental matrix, and answer the basic problems for the absorbing chain provided in the talk as an example.
3. Compute the first four moments of the same chain with the general method presented. Compare the first moment with the result of Problem 1.
4. Prove the general formula about $N\underline{v}$ being the expected value of the sum of entries on $\underline{v}$ during a walk until absorption.

Exercises

1. Show that the fundamental matrix N is non-negative.
2. Compute the fundamental matrix, and answer the basic problems for the absorbing chain provided in the talk as an example.
3. Compute the first four moments of the same chain with the general method presented. Compare the first moment with the result of Problem 1.
4. Prove the general formula about $N\underline{v}$ being the expected value of the sum of entries on $\underline{v}$ during a walk until absorption.
5. Observe that the hypergeometrical distribution is a special absorbing Markov chain with 2 states. Deduce the formula $1/p$ for the expected value as a special case of the expected runtime of an absorbing Markov chain.