MS-C2111 Stochastic Processes

Lecture 5

General Markov chains and random walks

Jukka Kohonen
Aalto University

Contents

Contents

Countable state spaces

Let S be a countable (finite or countably infinite) state space

Initial distribution is a map $\mu_0 : S \to [0,1]$ such that

$$\sum_{x \in S} \mu_0(x) \; = \; 1$$

Transition matrix is a map $P : S \times S \to [0,1]$ such that

$$\sum_{y \in S} P(x,y) \; = \; 1 \quad \text{for all } x \in S$$

Product $\mu_0 P$ (row vector $\times$ square matrix) is a map defined by

$$(\mu_0 P)(y) \; = \; \sum_{x \in S} \mu_0(x) P(x,y) \quad \text{for all } y \in S$$

Multiplication of infinite matrices

The product of nonnegative matrices P and Q is the nonnegative matrix $R = PQ$ defined by

$$R(x,y) \;=\; \sum_{z \in S} P(x,z) Q(z,y) \;\in\; [0,\infty].$$

When P and Q are transition matrices, then so is R, because $R(x,y) \geq 0$ and

$$\sum_{y \in S} R(x,y) \;=\; \sum_{z \in S} P(x,z) \sum_{y \in S} Q(z,y) \;=\; \sum_{z \in S} P(x,z) \;=\; 1.$$

Matrix powers are defined by $P^0 = I$ and $P^{t+1} = P^t P$.

Time-dependent distributions

Consider a Markov chain on a countably infinite state space with transition matrix P and initial distribution μ_0.

Theorem

The distribution of a Markov chain at time $t = 0, 1, 2, \ldots$ *can be computed by* $\mu_t = \mu_0 P^t$. *Moreover,*

$$\mathbb{P}(X_t = y \mid X_0 = x) \ = \ P^t(x, y).$$

Proof.

The proof for the finite state space works here equally well. $\square$

What is different?

Theorem

Any irreducible Markov chain on a finite state space admits a unique invariant distribution π.

For an irreducible Markov chain on an infinite state space, two things may happen:

- There is a unique invariant distribution π
- There is no invariant distribution at all

We will see examples of both cases soon.

Convergence theorem

Theorem

If *an irreducible and aperiodic Markov chain on a countable state space has an invariant distribution* π*, then the invariant distribution is unique and, regardless of the initial state,*

$$\sum_y |\mu_t(y) - \pi(y)| \to 0 \quad \textit{as } t \to \infty.$$

Especially:

- $\mu_t(y) \to \pi(y)$ for every y
- $d_{\mathrm{tv}}(\mu_t, \pi) \to 0$ where d_{tv} is the total variation metric
- $\mathbb{P}(X_t = y \mid X_0 = x) \to \pi(y)$ for all x, y

The proof (Lecture notes, Sec 5.4) is based on a stochastic coupling method + Markov chain covering theorem.

Markov chain covering theorem

Theorem

If an irreducible Markov chain on a countable state space S has an invariant distribution π, then the chain visits every state of the state space infinitely often with probability one.

Proof.

Lecture notes, Sec 5.3. □

The positive passage time into state y is defined by
$T_y^+ \;=\; \min\{t \geq 1 : X_t = y\}$.
The probability that a chain started at x later visits y is
$\rho(x, y) = \mathbb{P}_x(T_y^+ < \infty)$.
When $x = y$, the above number is the return probability of x.
Always

$$\rho(x, y) \geq P(x, y).$$

State x is recurrent if $\rho(x, x) = 1$ and transient otherwise.

Contents

Reversibility

Transition matrix P is reversible with respect to distribution π (or π-reversible) if the following detailed balance conditions are valid:

$$\pi(x)P(x,y) = \pi(y)P(y,x) \quad \text{for all } x, y.$$

Theorem

If P is π-reversible, then π is an invariant distribution of P.

Proof.

$$(\pi P)(y) = \sum_x \pi(x)P(x,y) = \sum_x \pi(y)P(y,x) = \pi(y)\sum_x P(y,x) = \pi(y)$$

□

Reversibility in time

For a Markov chain with a π-reversible transition matrix P, such that X_0 (and hence also every X_t) is π-distributed,

$$
\begin{aligned}
&\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \\
&\quad = \pi(x_0) P(x_0, x_1) P(x_1, x_2) P(x_2, x_3) \cdots P(x_{t-1}, x_t) \\
&\quad = P(x_1, x_0) \pi(x_1) P(x_1, x_2) P(x_2, x_3) \cdots P(x_{t-1}, x_t) \\
&\quad = P(x_1, x_0) P(x_2, x_1) \pi(x_2) P(x_2, x_3) \cdots P(x_{t-1}, x_t) \\
&\quad = \cdots \\
&\quad = P(x_1, x_0) P(x_2, x_1) P(x_3, x_2) \cdots P(x_t, x_{t-1}) \pi(x_t) \\
&\quad = \pi(x_t) P(x_t, x_{t-1}) \cdots P(x_1, x_0) \\
&\quad = \mathbb{P}(X_t = x_0, X_{t-1} = x_1, \ldots, X_0 = x_t)
\end{aligned}
$$

Statistically the chain looks the same when observed backwards in time.

Birth–death chains

A birth–death chain is a Markov chain on $S \subset \mathbb{Z}_+$ with a transition matrix such that $P(i,j) = 0$ for $|j - i| > 1$.

Note

- Birth–death chain can only move to its nearby states (or stay in the current state)
- Birth–death chains with constant transition probabilities (outside boundaries) are called random walks
- The state space can be finite (e.g. gambler's ruin) or countably infinite (e.g. random walk on $\mathbb{Z}_+$)

Invariant distributions of birth–death chains

Theorem

If a birth–death chain has an invariant distribution π, then the chain is π-reversible.

Proof.

(i) If $|j - i| > 1$, then evidently $\pi_i P_{i,j} = 0 = \pi_j P_{j,i}$.
(ii) If $j = i + 1$, then a chain with initial distribution π satisfies (draw a picture)

$$
\begin{aligned}
&\mathbb{P}(X_{t+1} \le i) \\
&= \mathbb{P}(X_t \le i - 1) + \mathbb{P}(X_t = i)(1 - P_{i,i+1}) + \mathbb{P}(X_t = i + 1)P_{i+1,i} \\
&= \mathbb{P}(X_t \le i) - \mathbb{P}(X_t = i)P_{i,i+1} + \mathbb{P}(X_t = i + 1)P_{i+1,i} \\
&= \mathbb{P}(X_t \le i) - \pi_i P_{i,i+1} + \pi_{i+1}P_{i+1,i}
\end{aligned}
$$

Because both X_t and X_{t+1} are π-distributed,

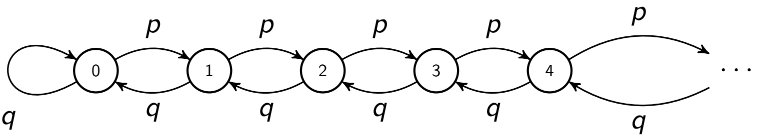
$$\pi_i P_{i,i+1} = \pi_{i+1} P_{i+1,i}.$$

□

Contents

Random walk on $\mathbb{Z}_+$

A random walk on $\mathbb{Z}_+ = \{0, 1, \dots\}$ moves from $x > 0$ to the right with probability $0 < p < 1$ and to the left with probability $q = 1 - p$.



When the boundary condition is $P(0,0) = q$ we obtain a transition matrix

$$P = \begin{bmatrix} q & p & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & \\ 0 & q & 0 & p & 0 & \\ 0 & 0 & q & 0 & p & \\ \vdots & & & \ddots & \ddots & \ddots \end{bmatrix}$$

Irreducible? YES. Aperiodic? YES.
Invariant distribution exists? (Nontrivial question!)

. . . Random walk on $\mathbb{Z}_+$

The random walk is a birth–death chain, so if an invariant distribution exists, it must satisfy the detailed balance conditions

$$\pi(x)P(x,x-1) \;=\; \pi(x-1)P(x-1,x)$$

$$\text{that is} \quad \pi(x)\,q \;=\; \pi(x-1)\,p, \qquad x \geq 1.$$

$\implies \pi(x) = \alpha^x \pi(0)$, $x \geq 0$, where $\alpha = \frac{p}{q}$.
If π is a probability distribution, then

$$1 = \sum_{x=0}^{\infty} \pi(x) = \pi(0) \sum_{x=0}^{\infty} \alpha^x = \begin{cases} \pi(0)\left(\frac{1}{1-\alpha}\right), & 0 < \alpha < 1, \\ \pi(0) \cdot \infty, & \alpha \geq 1. \end{cases}$$

- If $p < 1/2$, then the chain does have an invariant distribution $\pi(x) = (1-\alpha)\alpha^x$.
- If $p \geq 1/2$, the chain does not have an invariant distribution.

Random walk on $\mathbb{Z}_+$, $p < \frac{1}{2}$

Now $\alpha = p/q < 1$, so the invariant distribution is a geometric distribution

$$\pi(x) \;=\; (1-\alpha)\alpha^x, \quad x = 0, 1, \ldots$$

Note

- By the convergence theorem, the distribution of X_t converges to π, as $t \to \infty$.
- Every state of the chain is recurrent, so the chain visits all states of $\mathbb{Z}_+$ infinitely often.
- Hence $\limsup_{t\to\infty} X_t = \infty$ with probability one, so the path of the chain does not converge anywhere.
- Nevertheless the chain reaches its statistical equilibrium

$$\mathbb{P}_i(X_t = j) \;\to\; (1-\alpha)\alpha^j \quad \text{for all } i, j \geq 0.$$

What about $p > \frac{1}{2}$?

Random walk on $\mathbb{Z}_+$, $p > \frac{1}{2}$

When $p > \frac{1}{2}$, the chain does not have an invariant distribution.

What happens in the long run?

$$\text{Define} \quad Y_t = \begin{cases} +1, & \text{if the } t\text{-th step is to the right,} \\ -1, & \text{else} \end{cases}$$

Note

- $\mathbb{P}(Y_t = 1) = 1 - \mathbb{P}(Y_t = -1) = p$ for all $t \geq 1$.
- $\mathbb{E}(Y_t) = 2p - 1 > 0$.
- $Y_1, Y_2, \ldots$ are independent
- $(Y_t)_{t \geq 1}$ is a Markov chain on $\{-1, 1\}$, irreducible, aperiodic, invariant distribution π such that $\pi(-1) = 1 - p$, $\pi(1) = p$
- By ergodic theorem, $\frac{1}{t} \sum_{s=1}^{t} Y_s \to \sum_x x\pi(x) = 2p - 1$ w.pr. 1
- $X_t - X_{t-1} \geq Y_t$ for all $t \geq 1$.

Random walk on $\mathbb{Z}_+$, $p > \frac{1}{2}$

The law of large numbers implies that with probability one,

$$\lim_{t\to\infty} \frac{\sum_{s=1}^{t} Y_s}{t} = 2p - 1 > 0 \quad \Longrightarrow \quad \lim_{t\to\infty} \sum_{s=1}^{t} Y_s = \infty.$$

On the other hand, $X_t - X_0 = \sum_{s=1}^{t}(X_s - X_{s-1}) \geq \sum_{s=1}^{t} Y_s$, so that

$$X_t \to \infty \quad \text{with probability one.}$$

Hence if $p > 1/2$, then all states are transient.

What about $p = \frac{1}{2}$?

Random walk on $\mathbb{Z}_+$, $p = \frac{1}{2}$

$p = \frac{1}{2} \implies$ no invariant distribution

What happens in the long run?

T_j = passage time into state j

- Starting from state 1, the probability to reach N before 0 is $\mathbb{P}_1(T_N < T_0)$
- Gambler's ruin on $\{0, \dots, N\} \implies \mathbb{P}_1(T_N < T_0) = \frac{1}{N}$
- The probability of never visiting 0 equals

$$\mathbb{P}_1(T_0 = \infty) = \lim_{N\to\infty} \mathbb{P}_1(T_N < T_0) = \lim_{N\to\infty} \frac{1}{N} = 0$$

- Hence $\mathbb{P}_0(T_0^+ < \infty) = \mathbb{P}_1(T_0 < \infty) = 1$, so state 0 is recurrent.
- Because the chain is irreducible, all states are recurrent.

Random walk on $\mathbb{Z}_+$, $p = \frac{1}{2}$

What about the expected passage time $\mathbb{E}_i(T_0)$ from state i to state 0?

- Clearly $\mathbb{E}_i(T_0) \geq \mathbb{E}_i(\min\{T_0, T_N\})$.
- When $i \leq N$, the latter expectation is the same as for a gambler's ruin, so that $\mathbb{E}_i(\min\{T_0, T_N\}) = i(N - i)$.
- When $N \to \infty$, it hence follows that

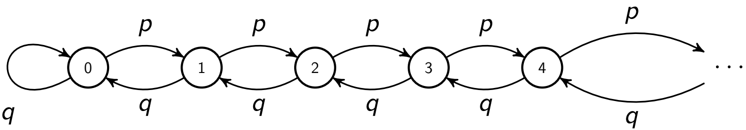
$$\mathbb{E}_i(T_0) \geq \mathbb{E}_i(\min\{T_0, T_N\}) = i(N - i) \to \infty.$$

  for all $i \geq 1$.
- The expected return time to state zero is $1 + \mathbb{E}_1(T_0) = \infty$.

Although the return time to state 0 is surely finite, the expected return time is infinite.

Random walk on $\mathbb{Z}_+$ — Summary

Irreducible and aperiodic Markov chain with infinite state space



$$P = \begin{bmatrix} q & p & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & \\ 0 & q & 0 & p & 0 & \\ 0 & 0 & q & 0 & p & \\ \vdots & & & \ddots & \ddots & \ddots \end{bmatrix}$$

- If $p < \frac{1}{2}$, then chain has the geometric distribution $\pi(x) = (1 - p/q)(p/q)^x$ as the unique invariant distribution.
- If $p = \frac{1}{2}$, then there is no invariant distribution, and all states are recurrent (the chain surely visits all states infinitely many times)
- If $p > \frac{1}{2}$, then there is no invariant distribution, and all states are transient (after visiting a state, the chain might never visit it again)

Random walks in multiple dimensions

The symmetric random walk on the d-dimensional integer lattice is recurrent for $d = 1, 2$ but transient for $d \geq 3$. (Georg Pólya 1920)

A drunk man will find his way home, but a drunk bird might get lost forever. (Shizuo Kakutani)

References

Richard Durrett.
Essentials of Stochastic Processes.
Springer, second edition, 2012.

Sources

Photos

1. Image courtesy of think4photop at FreeDigitalPhotos.net