First steps in random walks

(a brief introduction to Markov chains)

Paul-André Melliès

CNRS, Université Paris Denis Diderot

ANR Panda

Ecole Polytechnique   4 mai 2010
Step 1

Random variables

Before the walk
**Measurable spaces**

A **measurable space** is a set $\Omega$ equipped with a family of sets $A \subseteq \Omega$ called the **events** of the space, such that

(i) the set $\Omega$ is an event

(ii) if $A_1, A_2, \ldots$ are events, then $\bigcup_{i=1}^{\infty} A_i$ is an event

(iii) if $A$ is an event, then its complement $\Omega \setminus A$ is an event
Every topological space $\Omega$ induces a measurable space whose events $A \subseteq \Omega$ are defined by induction:

- the events of level $0$ are the **open sets** and the **closed sets**, 
- the events of level $k + 1$ are the **countable** unions and intersections

$$\bigcup_{i=1}^{\infty} A_i$$

$$\bigcap_{i=1}^{\infty} A_i$$

of events $A_i$ of level $k$. 
Typically...

The measurable space $\mathbb{R}$ equipped with its borelian events
Probability spaces

A measurable set $\Omega$ equipped with a **probability measure**

$$A \mapsto P(A) \in [0, 1]$$

which assigns a value to every event, in such a way that

(i) the event $\Omega$ has probability $P(\Omega) = 1$

(ii) the event $\bigcup_{i=1}^{\infty} A_i$ has probability

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

when the events $A_i$ are pairwise disjoint.
Random variable

A random variable on a measurable space $\mathcal{Y}$

$$X : \Omega \rightarrow \mathcal{Y}$$

is a measurable function from a probability space $(\Omega, P)$

called the universe of the random variable.

Notation for an event $A$ of the space $\mathcal{Y}$:

$$\{ X \in A \} := \{ \omega \in \Omega | X(\omega) \in A \} = X^{-1}(A)$$
Conditional probabilities

Given two random variables

$$X, Y : \Omega \rightarrow \mathcal{Y}$$

and two events $A, B$ such that

$$P \{ Y \in B \} \neq 0$$

the probability of $\{ X \in A \}$ conditioned by $\{ Y \in B \}$ is defined as

$$P \{ X \in A \mid Y \in B \} := \frac{P \{ X \in A \cap Y \in B \}}{P \{ Y \in B \}}$$

where

$$\{ X \in A \cap Y \in B \} = X^{-1}(A) \cap Y^{-1}(B).$$
Expected value

The expected value of a random variable

$$X : \Omega \rightarrow \mathbb{R}$$

is defined as

$$E (X) = \int_{\Omega} X \ dP$$

when the integral converges absolutely.

In the case of a random variable $X$ with finite image:

$$E (X) = \sum_{x \in \mathbb{R}} x \ P \{ X = x \}$$
Step 2

Markov chains

Stochastic processes
Finite Markov chains

A Markov chain is a sequence of random variables

\[ X_0, X_1, X_2, \ldots : \Omega \rightarrow \mathcal{Y} \]

on a measurable space \( \mathcal{Y} \) such that

\[ \mathbb{P} \{ X_{n+1} = y \mid X_1 = x_1, \ldots, X_n = x_n \} = \mathbb{P} \{ X_{n+1} = y \mid X_n = x_n \} \]

Every Markov chain is described by its transition matrix

\[ P(x, y) := \mathbb{P} \{ X_{n+1} = y \mid X_n = x \} \]
Stationary distribution

A stationary distribution of the Markov chain

\[ P \]

is a probability measure \( \pi \) on the state space \( \Upsilon \) such that

\[ \pi = \pi P \]

A stationary distribution \( \pi \) is a fixpoint of the transition matrix \( P \)
Reversible Markov chains

A probability distribution \( \pi \) on the state space \( \Upsilon \) satisfies the **detailed balance equations**

when

\[
\pi(x) \ P(x, y) = \pi(y) \ P(y, x)
\]

for all elements \( x, y \) of the state space \( \Upsilon \).

**Property.** Every such probability distribution \( \pi \) is stationary.
Proof of the statement

Suppose that

\[ \pi(x) P(x, y) = \pi(y) P(y, x) \]

for all elements \( x, y \) of the state space \( \Upsilon \). In that case,

\[
\pi P(x) = \sum_{y \in \Upsilon} \pi(y) P(y, x) \quad \text{by definition}
\]
\[
= \sum_{y \in \Upsilon} \pi(x) P(x, y) \quad \text{detailed balance equation}
\]
\[
= \pi(x) \quad \text{property of the matrix } P
\]
Irreducible Markov chains

A Markov chain is \textit{irreducible} when for any two states

\[ x, y \in \Omega \]

there exists an integer

\[ t \in \mathbb{N} \]

such that

\[ P^t(x, y) > 0 \]

where \( P^t \) is the transition matrix \( P \) composed \( t \) times with itself.
Step 3

Random walk

A concrete account of reversible Markov chains
Networks

A finite undirected connected graph

\[ G = (V, E) \]

where every edge \( e = \{x, y\} \) has a **conductance**

\[ c(e) \in \{ x \in \mathbb{R} \mid x > 0 \}. \]

The inverse of the conductance

\[ r(e) = \frac{1}{c(e)} \]

is called the **resistance** of the edge.
**Weighted random walk**

Every network defines a Markov chain

\[ P(x, y) = \frac{c(x, y)}{c(x)} \]

where

\[ c(x) = \sum_{x \sim y} c(x, y) \]

Here, \( x \sim y \) means that \( \{x, y\} \) is an edge of the graph \( G \).
A stationary probability

Define the probability distribution

\[ \pi(x) = \frac{c(x)}{c_G} \]

where

\[ c_G = \sum_{x \in V} \sum_{x \sim y} c(x, y) \]

The Markov chain \( P \) is reversible with respect to the distribution \( \pi \).

**Consequence.**

the distribution \( \pi \) is **stationary** for the Markov chain \( P \).
Conversely...

Every Markov chain $P$ on a finite set $\Upsilon$ reversible with respect to the probability $\pi$ may be recovered from the random walk on the graph

$$G = (V, E)$$

with set of vertices

$$V = \Upsilon$$

and edges

$$\{x, y\} \in E \iff P(x, y) > 0$$

weighted by the conductance

$$c(x, y) = \pi(x) P(x, y).$$
Step 4

Harmonic functions

Expected value of hitting time is harmonic
Harmonic functions

A function

\[ h : \Omega \rightarrow \mathbb{R} \]

is \textbf{harmonic} at a vertex \( x \) when

\[ h(x) = \sum_{y \in \Omega} P(x, y) \, h(y) \]

Here, \( P \) denotes a given transition matrix.

Harmonic functions at a vertex \( x \) define a vector space
Expected value

The **expected value** of a random variable on \( \mathbb{R} \) is defined as

\[
E(X) = \int_{\Omega} X \, dP
\]

In the finite case:

\[
E(X) = \sum_{x \in \mathbb{R}} x \, P \{ X = x \}
\]
Hitting time

The hitting time $\tau_B$ associated to a set of vertices $B$ is defined as

$$\tau_B = \min \{ t \geq 0 \mid X_t \in B \}$$

This defines a random variable

$$X_{\tau_B} : \mathcal{Y} \rightarrow B$$

which maps every $v \in \mathcal{Y}$ to the first element $b$ it reaches in the set $B$. 
Proof of the statement

\[ X_{\tau_B}^{-1}(b) = \bigcup_{n=0}^{\infty} \text{Hit}_n(b) \]

where

\[ \text{Hit}_0(b) = X_0^{-1}(b) \]
\[ \text{Hit}_1(b) = X_1^{-1}(b) \setminus X_0^{-1}(B) \]
\[ \text{Hit}_{n+1}(b) = X_{n+1}^{-1}(b) \setminus \bigcup_{b \in B} \text{Hit}_n(b) \]

This establishes that each \( X_{\tau_B}^{-1}(b) \) is an event of the universe \( \Omega \), and thus that \( X_{\tau_B} \) is a random variable.
Expected value

Given a function

\[ h_B : B \rightarrow \mathbb{R} \]

define the random variable:

\[ h_B \circ X_{\tau_B} : \Upsilon \rightarrow \Omega \rightarrow \mathbb{R} \]

whose expected value at the vertex \( x \) is denoted

\[ E_x [ h_B \circ X_{\tau_B} ] \]
Existence of an harmonic function

Observation: the function

\[ h : x \mapsto E_x \left[ h_B \circ X_{\tau_B} \right] \]

(i) coincides with \( h_B \) on the vertices of \( B \)

(ii) is harmonic on every vertex \( x \) in the complement \( \Omega \setminus B \).
Proof of the statement

\[ E_b \left[ h_B \circ X_{\tau_B} \right] = h_B (b) \]

\[ E_x \left[ h_B \circ X_{\tau_B} \right] = \sum_{y \in \Omega} P(x, y) \ E_x \left[ h_B \circ X_{\tau_B} \mid X_1 = y \right] \]

\[ = \sum_{y \in \Omega} P(x, y) \ E_y \left[ h_B \circ X_{\tau_B} \right] \]

\[ = \sum_{y \sim x} E_y \left[ h_B \circ X_{\tau_B} \right] \]
Uniqueness of the harmonic function

There exists a unique function

\[ h : \Omega \rightarrow \mathbb{R} \]

such that

\( (i) \) coincides with \( h_B \) on the vertices of \( B \)

\( (ii) \) is harmonic on every vertex \( x \) in the complement \( \Omega \setminus B \).
Proof of the statement

First, reduce the statement to the particular case

\[ h_B = 0 \]

Then, consider a vertex \( x \in \Omega \setminus B \) such that

\[ h(x) = \max \{ h(z) \mid z \in \Omega \} \]

Then, for every vertex \( y \) connected to \( x \), one has

\[ h(y) = \max \{ h(z) \mid z \in \Omega \} \]

because the function \( h \) is harmonic.
Step 5

Electric networks

Expected values as conductance
Idea

Now that we know that

\[ h : x \mapsto E_x \left[ h_B \circ X_{\tau_B} \right] \]

defines the unique harmonic function on the vertices of \( \Omega \setminus B \)...

let us find another way to define this harmonic function!
Voltage

We consider a source \( a \) and a sink \( z \) and thus define

\[
B = \{ a, z \}
\]

and define a voltage as any function

\[
W : V \rightarrow \mathbb{R}
\]

harmonic on the vertices of \( V \setminus \{a, z\} \).

A voltage \( W \) is determined by its boundary values \( W(a) \) and \( W(z) \).
Flows

A flow $\theta$ is a function on oriented edges of the graph, such that

$$\theta(x\hat{y}) = -\theta(y\hat{x})$$

The divergence

$$\text{div } \theta : x \mapsto \sum_{y \sim x} \theta(x\hat{y})$$

Observe that

$$\sum_{x \in V} \text{div } \theta(x) = 0$$
A flow from $a$ to $z$ is a flow such that

(i) Kirchnoff’s node law: \( \text{div} \ \theta (x) = 0 \)

(ii) the vertex $a$ is a source: \( \text{div} \ \theta (a) \geq 0 \)

Observe that

\[
\text{div} \ \theta (z) = - \text{div} \ \theta (a)
\]
Current flow

The current flow $I$ induced by a voltage $W$ is defined as

$$I(\vec{x}, \vec{y}) = \frac{W(x) - W(y)}{r(x, y)} = c(x, y) \left[ W(x) - W(y) \right]$$

From this follows Ohm’s law:

$$r(\vec{x}, \vec{y}) I(\vec{x}, \vec{y}) = W(y) - W(x)$$
Main theorem

\[ P_a (\tau_z < \tau_a^+ ) = \frac{1}{c(a) \mathcal{R}(a \leftrightarrow z)} = \frac{C(a \leftrightarrow z)}{c(a)} \]

where

\[ \mathcal{R}(a \leftrightarrow z) = \frac{W(a) - W(z)}{\|I\|} = \frac{W(a) - W(z)}{\text{div } \theta(a)} \]
**Edge-cutset**

An edge-cutset separating \( a \) from \( z \) is a set of vertices \( \Pi \) such that every path from \( a \) to \( z \) crosses \( \Pi \).

If \( \Pi_k \) is a set of disjoint edge-cutset separating sets, then

\[
\mathcal{R} ( a \leftrightarrow z ) \geq \sum_{k} \left( \sum_{e \in \Pi_k} c(e) \right)^{-1}
\]
Energy of a flow

The energy of a flow is defined as:

$$\mathcal{E}(\theta) = \sum_e [\theta(e)]^2 r(e)$$

**Theorem.** (Thompson’s Principle) For any finite connected graph,

$$R(a \leftrightarrow z) = \inf \{ \mathcal{E}(\theta) \mid \theta \text{ is a unit flow from } a \text{ to } z \}$$

where a unit flow $\theta$ is a flow from $a$ to $z$ such that

$$\text{div } \theta(a) = 1.$$