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 Ω equipped with a family of sets

$$A \quad \subseteq \quad \Omega$$

called the **events** of the space, such that

(i) the set Ω 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

Illustration

Every topological space Ω induces a measurable space whose events

$$A \quad \subseteq \quad \Omega$$

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

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

of events A_i of level k.

Typically...

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

Probability spaces

A measurable set Ω equipped with a **probability measure**

$$A \quad \mapsto \quad \mathbf{P}(A) \quad \in \quad [0,1]$$

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

(i) the event Ω has probability $\mathbf{P}(\Omega) = 1$

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

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

when the events A_i are pairwise disjoint.

Random variable

A **random variable** on a measurable space Υ

$$X \quad : \quad \Omega \quad \longrightarrow \quad \Upsilon$$

is a **measurable** function from a probability space

$$(\, \Omega \, , \, \mathbf{P} \,)$$

called the **universe** of the random variable.

Notation for an event A of the space Υ :

$$\boxed{\{ \quad X \in A \quad \} \quad := \quad \{ \quad \omega \in \Omega \quad | \quad X(\omega) \in A \quad \} \quad = \quad X^{-1} \, (A)}$$

Conditional probabilities

Given two random variables

$$X, Y \quad : \quad \Omega \quad \longrightarrow \quad \Upsilon$$

and two events A, B such that

$$\mathbf{P}\{\, Y \in B \,\} \quad \neq \quad 0$$

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

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

where

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

Expected value

The **expected value** of a random variable

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

is defined as

$$\mathbf{E}(X) = \int_\Omega X \, d\mathbf{P}$$

when the integral converges absolutely.

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

$$\mathbf{E}(X) = \sum_{x \in \mathbb{R}} x \, \mathbf{P}\{X = x\}$$

Step 2

Markov chains

Finite Markov chains

A **Markov chain** is a sequence of random variables

$$X_0\ ,\ X_1\ ,\ X_2\ ,\ \dots \quad : \quad \Omega \quad \longrightarrow \quad \Upsilon$$

on a measurable space Υ such that

$$\mathbf{P}\ \{\ X_{n+1} = y \mid X_1 = x_1, \dots, X_n = x_n\ \} \quad = \quad \mathbf{P}\ \{\ X_{n+1} = y \mid X_n = x_n\ \}$$

Every Markov chain is described by its **transition matrix**

$$P(x,y) \quad := \quad \mathbf{P}\ \{\ X_{n+1} = y \mid X_n = x\ \}$$

Stationary distribution

A **stationary distribution** of the Markov chain

$$P$$

is a probability measure π on the state space Υ such that

$$\pi = \pi P$$

A stationary distribution π is a fixpoint of the transition matrix P

Reversible Markov chains

A probability distribution π on the state space Υ satisfies the

detailed balance equations

when

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

for all elements x, y of the state space Υ.

Property. Every such probability distribution π is stationary.

Proof of the statement

Suppose that

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

for all elements x, y of the state space Υ. In that case,

$\pi\,P(x)$	$=$	$\sum_{y\in\Upsilon} \pi(y)\,P(y,x)$	by definition
	$=$	$\sum_{y\in\Upsilon} \pi(x)\,P(x,y)$	detailed balance equation
	$=$	$\pi(x)$	property of the matrix P

Irreducible Markov chains

A Markov chain is **irreducible** when for any two states

$$x, y \quad \in \quad \Omega$$

there exists an integer

$$t \quad \in \quad \mathbb{N}$$

such that

$$P^t(x, y) \quad > \quad 0$$

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**

$$c(e) \quad \in \quad \{ \quad x \in \mathbb{R} \quad | \quad x > 0 \quad \}.$$

The inverse of the conductance

$$r(e) \quad = \quad \frac{1}{c(e)}$$

is called the **resistance** of the edge.

Weighted random walk

Every network defines a Markov chain

$$P(x, y) \quad = \quad \frac{c(x, y)}{c(x)}$$

where

$$c(x) \quad = \quad \sum_{x \sim y} \quad c(x, y)$$

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 π.

Consequence.

the distribution π is **stationary** for the Markov chain P.

Conversely...

Every Markov chain P on a finite set Υ reversible with respect to the probability π 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

$$h \quad : \quad \Omega \quad \longrightarrow \quad \mathbb{R}$$

is **harmonic** at a vertex x when

$$h(x) \quad = \quad \sum_{y \in \Omega} \quad P(x,y)\ h(y)$$

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

$$\mathbf{E}(X) = \int_{\Omega} X \, d\mathbf{P}$$

In the finite case:

$$\mathbf{E}(X) = \sum_{x \in \mathbb{R}} x \; \mathbf{P}\{X = x\}$$

Hitting time

The hitting time τ_B associated to a set of vertices B is defined as

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

This defines a random variable

$$X_{\tau_B} \quad : \quad \Upsilon \quad \longrightarrow \quad B$$

which maps every $v \in \Upsilon$ to the first element b it reaches in the set B.

Proof of the statement

$$X_{\tau_B}^{-1}(b) = \bigcup_{n=0}^{\infty} \mathbf{Hit}_n(b)$$

where

$$\mathbf{Hit}_0(b) = X_0^{-1}(b)$$

$$\mathbf{Hit}_1(b) = X_1^{-1}(b) \setminus X_0^{-1}(B)$$

$$\mathbf{Hit}_{n+1}(b) = X_{n+1}^{-1}(b) \setminus \bigcup_{b \in B} \mathbf{Hit}_n(b)$$

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

Expected value

Given a function

$$h_B \quad : \quad B \quad \longrightarrow \quad \mathbb{R}$$

define the random variable:

$$h_B \circ X_{\tau_B} \quad : \quad \Upsilon \quad \longrightarrow \quad \Omega \quad \longrightarrow \quad \mathbb{R}$$

whose expected value at the vertex x is denoted

$$\mathbf{E}_x \left[\, h_B \circ X_{\tau_B} \,\right]$$

Existence of an harmonic function

Observation: the function

$$h \quad : \quad x \quad \mapsto \quad \mathbf{E}_x \, [\, h_B \circ X_{\tau_B} \,]$$

(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

$$\mathbf{E}_b \left[h_B \circ X_{\tau_B} \right] = h_B(b)$$

$$\begin{aligned}
\mathbf{E}_x \left[h_B \circ X_{\tau_B} \right] &= \sum_{y \in \Omega} P(x,y)\; \mathbf{E}_x \left[h_B \circ X_{\tau_B} \mid X_1 = y \right] \\
&= \sum_{y \in \Omega} P(x,y)\; \mathbf{E}_y \left[h_B \circ X_{\tau_B} \right] \\
&= \sum_{y \sim x} \mathbf{E}_y \left[h_B \circ X_{\tau_B} \right]
\end{aligned}$$

Uniqueness of the harmonic function

There exists a unique function

$$h \quad : \quad \Omega \quad \longrightarrow \quad \mathbb{R}$$

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

$$h_B \quad = \quad 0$$

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

$$h(x) \quad = \quad max \, \{ \, h(z) \mid z \in \Omega \, \}$$

Then, for every vertex y connected to x, one has

$$h(y) \quad = \quad max \, \{ \, h(z) \mid z \in \Omega \, \}$$

because the function h is harmonic.

Step 5

Electric networks

Expected values as conductance

Idea

Now that we know that

$$h \quad : \quad x \quad \mapsto \quad \mathbf{E}_x \, [\, h_B \circ X_{\tau_B} \,]$$

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

$$W : V \longrightarrow \mathbb{R}$$

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 θ is a function on **oriented edges** of the graph, such that

$$\theta\,(\vec{xy}) \quad = \quad -\ \theta\,(\vec{yx})$$

The **divergence**

$$\mathbf{div}\ \theta \quad : \quad x \quad \mapsto \quad \sum_{y \sim x} \quad \theta\ (\vec{xy})$$

Observe that

$$\boxed{\quad \textstyle\sum_{x \in V} \quad \mathbf{div}\ \theta\ (x) \quad = \quad 0 \quad}$$

Flows from source to sink

A flow from a to z is a flow such that

(i) Kirchhoff's node law: $\mathbf{div}\ \theta\,(x) = 0$

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

Observe that

$$\boxed{\mathbf{div}\ \theta\,(z) = -\,\mathbf{div}\ \theta\,(a)}$$

Current flow

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

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

From this follows Ohm's law:

$$\boxed{\; r(\,\vec{xy}\,)\;\; I(\,\vec{xy}\,) \quad = \quad W(y) - W(x) \;}$$

Main theorem

$$\mathbf{P}_a\,(\,\tau_z \;<\; \tau_a^+\,) \;=\; \frac{1}{c\,(a)\;\mathcal{R}\,(\,a \leftrightarrow z\,)} \;=\; \frac{\mathcal{C}\,(\,a \leftrightarrow z\,)}{c\,(a)}$$

where

$$\mathcal{R}\,(\,a \leftrightarrow z\,) \;=\; \frac{W(a) \;-\; W(z)}{\|\,I\,\|} \;=\; \frac{W(a) \;-\; W(z)}{\mathbf{div}\;\theta\,(a)}$$

Edge-cutset

An edge-cutset separating a from z is a set of vertices

$$\Pi$$

such that every path from a to z crosses Π.

If Π_k is a set of disjoint edge-cutset separating sets, then

$$\mathcal{R}\,(\,a \leftrightarrow z) \quad \geq \quad \sum_k \quad \left(\quad \sum_{e \in \Pi_k} c(e) \quad \right)^{-1}$$

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,

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

where a unit flow θ is a flow from a to z such that

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