Communication Networks, Algorithms & Probability Theory

Philippe Robert

École Polytechnique & INRIA, France

2011

Markov Chains

Modeling

Observations of a system: M_0, M_1, $\ldots M_n$

- M_n random variable;
- M_n may depend, a priori, of M_0, $\ldots M_{n-1}$ (and even of M_{n+1}, $\ldots$)
- Statistical data:
  Probability of a given sample path

  $$\mathbb{P}(M_0 = x_0, M_1 = x_1, \ldots, M_n = x_n),\ n \geq 1,$$

  (x_i) sequence of the state space.

Modeling (II)

Problem

Mathematical Representation of (M_n)

- Minimal
  not all probabilities of all sample paths.
- Efficient
  mathematical analysis of asymptotic behavior of (M_n).

Possible (Solution)
$\Rightarrow$Markovian modeling

Framework

State space: $\mathcal{S}$ finite or countable.

- $\mathcal{S} = \{1, \ldots, N\}$;
- $\mathcal{S} = \mathbb{N}, \mathbb{Z}^d$;
- $\mathcal{S} = \{1, \ldots, C\}^{(\mathbb{N})}$: finite sequences in $\{1, \ldots, C\}$.

Transition Matrices (Booklet page 98)

A transition matrix

$$P = (p(x,y),\ x, y \in \mathcal{S})$$

is a matrix with non-negative coefficients such that

$$\sum_{y \in \mathcal{S}} p(x,y) = 1, \quad \forall x \in \mathcal{S}.$$

Linear Algebra: $P \cdot 1 = 1$.
1: right eigenvector for eigenvalue 1.

Markov Property

A sequence (M_n) of random variables has the Markov property if

$$\mathbb{P}(M_n{=}x_n | M_{n-1}{=}x_{n-1}, M_{n-2}{=}x_{n-2}, \ldots, M_0{=}x_0) \\ {=}\mathbb{P}(M_n{=}x_n \mid M_{n-1}{=}x_{n-1}).$$

for $n \geq 1$.
Translation:
Given all past states visited, the last location is sufficient to determine the future evolution.

Markov property (II)

The Markov chain (M_n) is homogeneous if,

$$\mathbb{P}(M_n = y | M_{n-1} = x) \\ = \mathbb{P}(M_1 = y | M_0 = x) = p(x,y).$$

for $n \geq 1$,

$(p(x,y), x, y \in \mathcal{S})$ is the transition matrix of the Markov chain.

Homogeneity:
Dynamics of (M_n) is invariant in time.

Example 1: independent variables

Successive coin tossing

$$M_0, M_1, \ldots, M_n, \ldots \in \{0,1\}$$

- Bias $q \in [0,1]$, $\mathbb{P}(M_k = 0) = q$, $\forall k \geq 0$.
- Independence $\Rightarrow$
  $\mathbb{P}(M_n{=}0|M_{n-1}, \ldots, M_0){=}q$
- Transition matrix

$$\begin{aligned} p(0,0) &= P(M_1 = 0|M_0 = 0) = q \\ &= P(M_1 = 0|M_0 = 1) = p(1,0), \\ p(0,1) &= p(1,1) = 1 - q. \end{aligned}$$

A trivial Markovian dependence

Example 2: Flip-flop chain

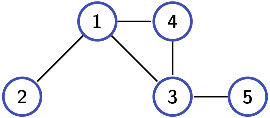
Observations in $\mathcal{S} = \{0,1\}$.

$0 \to 1$ with proba a_0
$1 \to 0$ with proba a_1

Transition Matrix

$$P = \begin{pmatrix} p(0,0) & p(0,1) \\ p(1,0) & p(1,1) \end{pmatrix} = \begin{pmatrix} 1 - a_0 & a_0 \\ a_1 & 1 - a_1 \end{pmatrix}$$

Example 3: Random walk on a graph



$\mathcal{G} = \{1, 2, \ldots N\}$. If x, $y \in \mathcal{G}$ and $x \sim y$,

$$p(x,y) = P(M_1 = y \mid M_0 = x) = \frac{1}{d_x}$$

d_x degree of x.

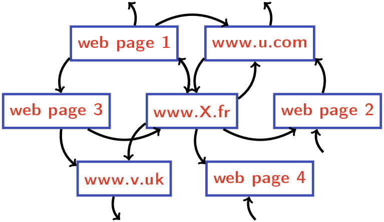
Random walk in $\mathbb{Z}^d$

Transition matrix:

$$p(x, x + e_i) = p(x, x - e_i) = \frac{1}{2d}$$

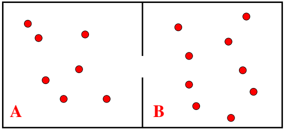
$e_i = (0, 0, \ldots, 0, 1, \ldots, 0)$, ith unit vector
$1 \leq i \leq d$.

Asymmetrical random walk on the web



$p(x,y) = 1/d_x^{\to}$, $d_x^{\to}$: out-degree of x.
Model used by Google.

Example 4:
Ehrenfest Urn with N particles



Dynamics:
A particle taken at random change location

Example 4:
Ehrenfest Urn with N particles

M_n: nb of part. in A at $t{=}n$, $0 \leq M_n \leq N$.
(M_n) Markov transition matrix:
$p(x, x-1) = x/N$ and
$p(x, x+1) = 1 - x/N$

Example 5: Coupon collector problem

(X_n) ind. uniform var. on $\{1, \ldots, N\}$

$$M_n = \mathrm{card}\{\mathrm{X}_1, \ldots, \mathrm{X}_\mathrm{n}\}$$

(M_n) Markov with transition matrix,

$$p(x, x{+}1) = 1 - \frac{x}{N} \quad \text{and} \quad p(x,x) = \frac{x}{N}$$

Remark: $p(N,N) = 1$,
N is an absorbing state.

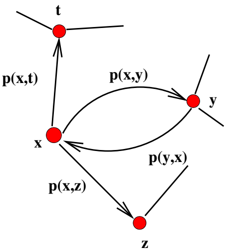
Example 6: Galton-Watson Process

- An individual gives birth to G children; G random variable.
- Z_n : nb. individuals of the nth generation, $Z_0 = 1$.
- (Z_n) Markov chain.

Transitions: $|u| < 1$,

$$\mathbb{E}\left(u^{Z_{n+1}} \middle| Z_n = x\right) = \left(\mathbb{E}\left(u^G\right)\right)^x$$

Representation in terms of graph of a Markov chain



Law of a Markov chain

(Booklet page 99)

If (M_n) is a Markov chain with transition matrix $P = (p(x,y))$
ν initial distribution: $\mathbb{P}(M_0 = x_0) = \nu(x_0)$,

Proba of a sample path

$$\mathbb{P}(M_0 = x_0, M_1 = x_1, \ldots, M_n = x_n)$$
$$= \nu(x_0)p(x_0,x_1)p(x_1,x_2)\cdots p(x_{n-1},x_n).$$

for $x_1, \ldots, x_n \in \mathcal{S}$.

Law of a Markov chain (II)

$$\mathbb{P}(M_n = y)$$
$$= \sum_{x_0,\ldots,x_{n-1}\in\mathcal{S}} \nu(x_0)p(x_0,x_1)p(x_1,x_2)\cdots p(x_{n-1},y).$$

Matrix representation:

$$\mathbb{P}(M_n = y) = (\nu \cdot P^n)(y)$$

The powers of matrix P determine the distribution of (M_n).

Law of a Markov chain (III)

Particular case:

$$\mathbb{P}(M_n = y | M_0 = x) = (P^n)(x, y)$$

probability of going from x to y in n steps.

$$p^n(x, y) \overset{\text{def.}}{=} \mathbb{P}(M_n = y | M_0 = x)$$

Markovian notation

- (M_n) Markov chain on $\mathcal{S}$;
- ν Proba on $\mathcal{S}$.

$\mathbb{P}_\nu$: law of (M_n) when
$\mathbb{P}(M_0 \in A) = \nu(A)$, for $A \subset \mathcal{S}$

$$\begin{aligned} &\mathbb{P}_\nu(M_n \in A) \\ &= \sum_{x_0 \in \mathcal{S}, y \in A} \nu(x_0) \mathbb{P}(M_n = y | M_0 = x_0) \\ &= \sum_{x_0 \in \mathcal{S}, y \in A} \nu(x_0) P^n(x_0, y) \end{aligned}$$

Markovian Notation (II)

If $x \in \mathcal{S}$,

$$\mathbb{P}_x(M_n \in A) = \mathbb{P}(M_n \in A \mid M_0 = x).$$

In fact

$$\mathbb{P}_x = \mathbb{P}_{\delta_x}$$

Properties of Markov chains

Definition: the past before time t

$\mathcal{F}_t$: set of events
before time t for the Markov chain (M_n) :

$$\mathcal{F}_t = \bigvee_{A_0,\dots,A_t \subset \mathcal{S}} \{M_0 \in A_0, M_1 \in A_1, \dots, M_t \in A_t\}$$

- $(\mathcal{F}_t) \nearrow$;
- $\cup_{t\geq 0}\mathcal{F}_t$: All events.

$(\mathcal{F}_t)$ is the natural filtration
of the Markov chain.

The Markov property rewritten

$$\mathbb{P}(M_{n+1} \in A \mid \mathcal{F}_n) = \mathbb{P}(M_{n+1} \in A \mid M_n)$$

Irreducibility property

(Booklet page 105)

The Markov chain (M_n) is irreducible if,

$$\forall x, y \in \mathcal{S}, \exists n \geq 0 : \mathbb{P}(M_n = y \mid M_0 = x) > 0.$$

On the graph of the Markov chain:
there exists a path from x to y.

If non-irreducible:
Decomposition into irreducible components.

Equilibrium of Markov chains

Invariant Distribution

(Booklet page 114)

A Probability $(\pi(x) : x \in \mathcal{S})$ on $\mathcal{S}$ is invariant for the transition matrix $(p(x,y))$ if

$$\pi(x) = \sum_y \pi(y) p(y,x), \forall x \in \mathcal{S}.$$

Matrix Notation: π invariant if

$$\pi = \pi \cdot P$$

π eigenvector for the eigenvalue 1.

Stationarity Property

If π is an invariant probability and (M_n) such that

$$\mathbb{P}(M_0 = x) = \pi(x) \qquad \forall x \in \mathcal{S}$$

then

$$\mathbb{P}(M_n = x) = \pi(x) \qquad \forall n \geq 1, \forall x \in \mathcal{S}$$

The distribution of the Markov chain does not change with time.

The Markov chain is at equilibrium.

Example: Flip-flop chain

Transition Matrix

$$\begin{pmatrix} 1-a_0 & a_0 \\ a_1 & 1-a_1 \end{pmatrix} \qquad \pi(x) = \sum_y \pi(y) p(y,x)$$

Equation for invariant distribution:

$$\begin{aligned} \pi(0) &= (1-a_0)\pi(0) + a_1\pi(1) \\ \pi(1) &= (1-a_1)\pi(1) + a_0\pi(0) \end{aligned}$$

and $\pi(1) + \pi(0) = 1$.

$$\pi(0) = \frac{a_1}{a_0 + a_1} \qquad \pi(1) = \frac{a_0}{a_0 + a_1}$$

Example: Random walk on a graph

$$\mathcal{G} = \{1, 2, \ldots N\}$$

$$p(x,y) = \frac{1}{d_x}$$

if $x \sim y$, d_x degree of x.
Invariant distribution:

$$\pi(x) = C d_x, \quad C = 1 \Big/ \sum_{y=1}^{N} d_y \,.$$

Example: Random walk on $\mathbb{Z}$

Transition Matrix

$$p(n, n+1) = p(n, n-1) = 1/2.$$

Solution of equations for invariant distribution

$$\pi(n) = C, \quad \forall n \in \mathbb{Z}.$$

Non-summable: no invariant probability.
No Equilibrium.

Uniqueness

Proposition
An irreducible Markov chain has at most one invariant probability.

Proposition
An irreducible Markov chain on a finite state space has a unique invariant probability.

Convergence in distribution

Theorem (Booklet page 118)
If (M_n) is an irreducible and aperiodic Markov chain with an invariant probability $(\pi(z), z \in \mathcal{S})$, then for $x \in \mathcal{S}$,

$$\lim_{n \to +\infty} \mathbb{P}(M_n = x) = \pi(x).$$

Expression of π :
powers of transition matrix $(p(x, y))$.

$\Rightarrow$ in general: Difficult and inefficient.

Ergodic Theorem

Proposition (Booklet page 120)
If (M_n) is an irreducible Markov chain with invariant probability $(\pi(z), z \in \mathcal{S})$,
then $\mathbb{P}$-almost surely

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} f(M_i) = \mathbb{E}_\pi(f) \overset{\text{def.}}{=} \sum_{x \in \mathcal{S}} \pi(x) f(x)$$

for any function f such that
$\mathbb{E}_\pi(|f|) < + = \infty$.

Application

If $A \subset \mathcal{S}$, then

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{\{M_i\in A\}} = \pi(A) = \sum_{x\in A}\pi(x)$$

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{\{M_i=x\}} = \pi(x)$$

Basic Principle of simulation

A model of the web by Google

Importance of a web page
Search of "XYZT" on the Internet:

- Set E of pages with "XYZT";
  Easy: Robots.
- Find a fn I such that $E=\{P_1,P_2,\ldots\}$

$$I(P_1) > I(P_2) > \cdots$$

$I(P)$: Importance of page P.

A good model for $I : P \mapsto I(P)$?

A model of the web by Google

Brin and Page (1997)
The model of the random surfer S

- S navigates at random on the web:
  If $X(t) = P$, at $t+1$, S chooses at random an outgoing-link of this page and goes to the corresponding web page, etc...

A model of the web by Google

Duration of navigation T, I defined by

$$I(P) \overset{\text{def.}}{=} \frac{1}{T}\sum_{n=1}^{T} \mathbb{1}_{\{X(n)=P\}}$$

If T large, ergodic theorem: $I(P)\sim\pi(P)$
π: invariant proba of the random walk on the web.

$$\pi(x) = \sum_{y\sim x}\pi(y)/d_y^{\rightarrow}$$

Explicit Expression Unknown.

Strong Markov Property

Definition: Stopping Time

(Booklet page 93)

T random variable in $\mathbb{N} \cup \{+\infty\}$ is a **stopping time** if,

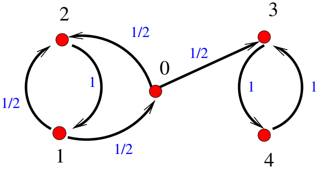
$$\{T = n\} \in \mathcal{F}_n, \qquad n \geq 1,$$

Translation:
An observer who looks at successive states of (M_n), knows when to stop at time T.

Stopping times: Examples

- $p \geq 0$, $T \equiv p$;
- Hitting time of a subset:
  $F \subset \mathcal{S}$,
  $$T_F = \inf\{n : M_n = F\}$$
  if $x \in \mathcal{S}$,
  $$T_x = \inf\{n : M_n = x\}$$
- If $\mathcal{S}$ finite: Cover time
  $$\tau = \inf\{n : \{M_0, M_1, \ldots, M_n\} = \mathcal{S}\}$$

Counter-example



$$M_0 = 0, \qquad L = \sup\{n : M_n = 0\}$$

is not a stopping time.

The past before a stopping time T

$$\mathcal{F}_T = \vee_{t=0}^{+\infty} \vee_{A_0,\dots,A_t \subset \mathcal{S}}$$

$$\mathcal{P}(\{M_0 \in A_0, M_1 \in A_1, \dots, M_t \in A_t, T \geq t\})$$

$\mathcal{F}_T$: all events before time T

Strong Markov

A Markov chain (M_n) (Booklet page 94)

satisfies the strong Markov property:
If for any stopping time T :

$$\mathbb{P}(M_{T+n} = y \mid T < +\infty, M_T = x_T, \dots, M_0 = x_0)$$
$$= \mathbb{P}(M_{T+n} = y \mid T < +\infty, M_T = x_T).$$

Strong Markov $\equiv$ Markov at non deterministic instants: stopping times.

Strong Markov (II)

Translation:

$$P(M_{T+n} \in \cdot \mid \mathcal{F}_T, T < +\infty)$$
$$= \mathbb{P}(M_{T+n} \in \cdot \mid M_T, T < +\infty)$$

Example

If $x \in \mathcal{S}$, $T_x = \inf\{n : M_n = x\}$
Assuming that $T_x < +\infty$ $\mathbb{P}$-a.s.
Strong Markov:

$$\mathbb{P}(M_{T_x+n} \in A \mid M_{T_x}, \dots, M_0)$$
$$= \mathbb{P}(M_{T_x+n} \in A \mid M_{T_x} = x)$$
$$= \mathbb{P}(M_n \in A \mid M_0 = x).$$

The sequence (M_{n+T_x}) is independent of the variables $M_{T_x-1}, \dots, M_0$.

Decomposition into independent cycles.