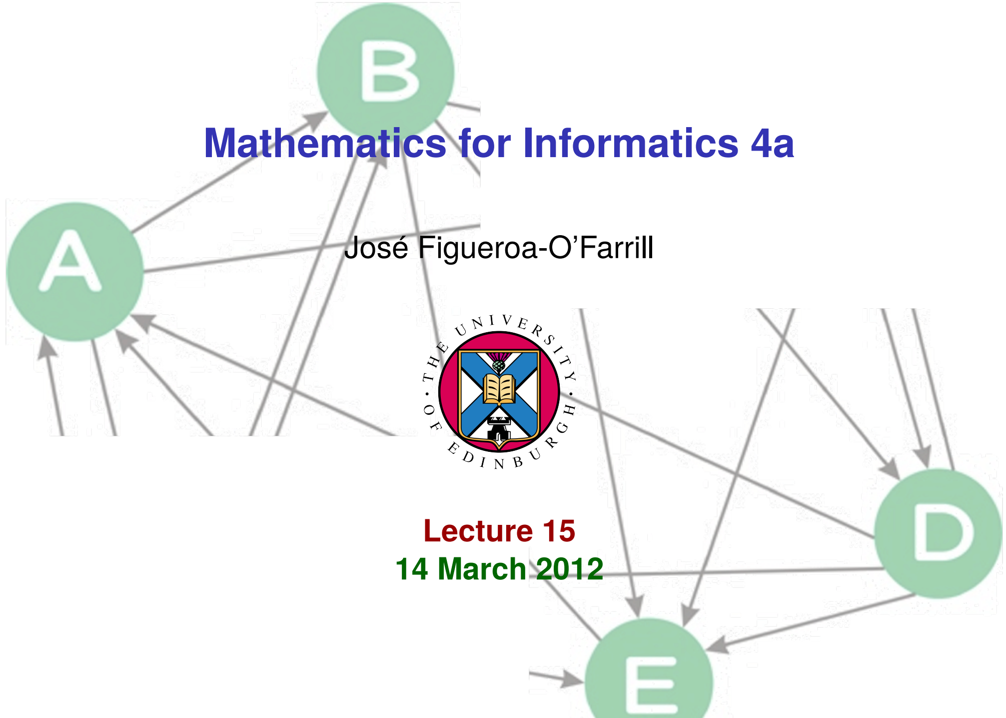


Mathematics for Informatics 4a

José Figueroa-O'Farrill

Lecture 15
14 March 2012

Determinism *vs* randomness

- There are two main kinds of processes in Nature, distinguished by their time evolution.
- In a **deterministic** process, the future state of the system is completely determined by the present state.
- Physical systems whose time evolution is described by differential equations are deterministic; e.g.,
  - classical mechanics (Newton's equation)
  - quantum mechanics (Schrödinger's equation)
  - the weather (chaotic but deterministic!)
- **Stochastic** (or **random**) processes are non-deterministic: the time evolution is subject to a probability distribution.
- Examples of stochastic processes are
  - Random walks
  - Markov chains
  - Birth-death processes
  - Queues
- These are the subject of the last part of this course.

Stochastic processes

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.
- Let $\mathcal{S}$ be a set called the **state space** of the system. The set $\mathcal{S}$ can be countable or uncountable.
- Let $\mathcal{T}$ be an index set, to be thought of as "time". It can be **continuous** or **discrete**.
- A **stochastic** (or **random**) process with state space $\mathcal{S}$ is a collection of random variables $X_t : \Omega \to \mathcal{S}$ indexed by $t \in \mathcal{T}$.
- The interpretation is that X_t is the state of the system at time t, which for a non-deterministic system is a random variable with some probability distribition.
- There are many kinds of stochastic processes, differing in how the probability of X_t being in a given state depends on the history of the system; that is, in which state the system was in times $t' < t$.

Markov chains

- We assume that $\mathcal{S}$ is **countable** so that the X_t are discrete random variables.
- We will also assume that we have a **discrete-time** process, so that $\mathcal{T} = \{0, 1, 2, \dots\}$.

Definition

A stochastic process $X = \{X_0, X_1, X_2, \dots\}$ is a **Markov chain** if it satisfies the **Markov property**:

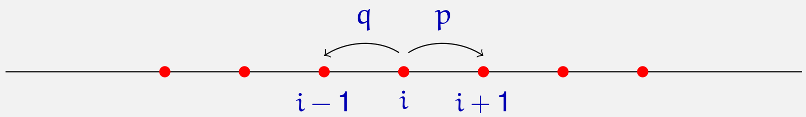
$$\mathbb{P}(X_{n+1} = s | X_0 = s_0, \dots, X_n = s_n) = \mathbb{P}(X_{n+1} = s | X_n = s_n)$$

for all $n \geqslant 0$ and $s_0, s_1, \dots, s_n, s \in \mathcal{S}$.

"given the present, the future does not depend on the past"

Random walks

Consider a particle moving on the integer lattice in $\mathbb{R}$:



Therefore $\mathcal{S} = \mathbb{Z}$ and J_i are independent random variables with

$$\mathbb{P}(J_i = 1) = p \qquad \mathbb{P}(J_i = -1) = q = 1 - p$$

Let X_n denote the position of the particle at time n, so that

$$X_n = X_0 + \sum_{i=1}^{n} J_i$$

Proposition

The sequence $\{X_0, X_1, X_2, \ldots\}$ exhibits **spatial homogeneity**:

$$\mathbb{P}(X_n = j \mid X_0 = a) = \mathbb{P}(X_n = j + b \mid X_0 = a + b)$$

Proof.

$$\mathbb{P}(X_n = j \mid X_0 = a) = \mathbb{P}\left(\sum_{i=1}^{n} J_i = j - a\right)$$

and also

$$\mathbb{P}(X_n = j + b \mid X_0 = a + b) = \mathbb{P}\left(\sum_{i=1}^{n} J_i = j + b - (a + b) = j - a\right)$$

$\square$

Proposition

The sequence $\{X_0, X_1, X_2, \ldots\}$ exhibits **temporal homogeneity**:

$$\mathbb{P}(X_n = j \mid X_0 = a) = \mathbb{P}(X_{n+m} = j \mid X_m = a)$$

Proof.

$$\mathbb{P}(X_n = j \mid X_0 = a) = \mathbb{P}\left(\sum_{i=1}^{n} J_i = j - a\right)$$

whereas

$$\mathbb{P}(X_{n+m} = j \mid X_m = a) = \mathbb{P}\left(\sum_{i=m+1}^{m+n} J_i = j - a\right)$$

but the J_i are i.i.d. $\square$

Proposition

The sequence $\{X_0, X_1, X_2, \ldots\}$ exhibits the **Markov property**:

$$\mathbb{P}(X_{m+n} = j \mid X_0 = i_0, \ldots, X_m = i_m) = \mathbb{P}(X_{m+n} = j \mid X_m = i_m)$$

Proof.

This follows because

$$X_{m+n} = X_m + \sum_{i=m+1}^{n} J_i$$

so X_{m+n} does not depend explicitly on the X_j for $j < m$. $\square$

Example (Gambler's ruin)

A gambler starts with £k and plays a game in which a fair coin is tossed repeatedly: winning £1 if heads and $-$£1 if tails. The game stops when the gambler's fortune is either £N ($N > k$) or £0. *What is the probability that the gambler is ultimately ruined?* This is an example of a random walk on a finite set $\{0, 1, 2, \dots, N\}$. Let R denote the event that the gambler is eventually ruined and let H and T denote the events that the first toss is heads and tails, respectively. Let $\mathbb{P}_k(R)$ denote the probability that gambler is eventually ruined starting with £k. Then

$$\mathbb{P}_k(R) = \mathbb{P}_k(R \mid H)\mathbb{P}(H) + \mathbb{P}_k(R \mid T)\mathbb{P}(T)$$

but clearly $\mathbb{P}_k(R \mid H) = \mathbb{P}_{k+1}(R)$ and $\mathbb{P}_k(R \mid T) = \mathbb{P}_{k-1}(R)$, whence

$$\mathbb{P}_k(R) = \tfrac{1}{2}\mathbb{P}_{k+1}(R) + \tfrac{1}{2}\mathbb{P}_{k-1}(R)$$

Example (Gambler's ruin – continued)

Letting $p_k = \mathbb{P}_k(R)$, we have the following difference equation:

$$p_k = \tfrac{1}{2}(p_{k+1} + p_{k-1}) \qquad p_0 = 1 \quad p_N = 0$$

Let $a_k = p_k - p_{k-1}$. Then

$$\begin{aligned} a_k - a_{k-1} &= p_k - p_{k-1} - (p_{k-1} - p_{k-2}) \\ &= p_k - 2p_{k-1} + p_{k-2} \\ &= p_k - (p_k + p_{k-2}) + p_{k-2} = 0 \end{aligned}$$

Therefore $a_k = a_1$ for all k and hence

$$p_k = a_1 + p_{k-1} = 2a_1 + p_{k-2} = \cdots = ka_1 + p_0$$

Since $p_0 = 1$ and $p_N = 0$, we find $a_1 = -\frac{1}{N}$, whence $p_k = 1 - \frac{k}{N}$.

Example (Gambler's ruin – continued)

What about if the coin is not fair?
Let $\mathbb{P}(H) = p$ and $\mathbb{P}(T) = q = 1 - p$, with $p \neq q$. Now

$$p_k = pp_{k+1} + qp_{k-1} \qquad 1 \leqslant k \leqslant N-1$$

with the same boundary conditions $p_0 = 1$ and $p_N = 0$. Try a solution $p_k = \theta^k$ for some θ. Then

$$\theta^k = p\theta^{k+1} + q\theta^{k-1} \implies p\theta^2 - \theta + q = 0$$

with roots $\theta_1 = 1$ and $\theta_2 = \frac{q}{p}$. The general solution is then

$$p_k = c_1\theta_1^k + c_2\theta_2^k$$

for some c_1, c_2 which are determined by $p_0 = 1$ and $p_N = 0$.

Example (Gambler's ruin – continued)

Imposing the boundary conditions

$$1 = p_0 = c_1 + c_2 \qquad 0 = p_N = c_1 + c_2\left(\frac{q}{p}\right)^N$$

whence

$$c_1 = -c_2\left(\frac{q}{p}\right)^N \qquad c_2 = \frac{1}{1 - \left(\frac{q}{p}\right)^N}$$

and hence

$$p_k = -\frac{\left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N} + \frac{\left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N} = \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}$$

Transition matrix

Let's go back to the case of a general Markov chain $\{X_0, X_1, X_2, \dots\}$.
Since $\mathcal{S}$ is countable we will assume it is a subset of $\mathbb{Z}$.
The evolution of a Markov chain is described by its **transition probabilities**

$$\mathbb{P}(X_{n+1} = j \mid X_n = i)$$

We will make the additional assumption of temporal homogeneity:

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_1 = j \mid X_0 = i)$$

Therefore the transition probabilities are encoded in a **transition matrix** $\mathbf{P} = (p_{ij})$, where

$$p_{ij} = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

Theorem

The transition matrix of a Markov chain is ***stochastic****; that is,*

1. $p_{ij} \geqslant 0$
2. $\sum_j p_{ij} = 1$ *for all* i *(i.e., rows sum to* 1*)*

Proof.

1. This is obvious since the p_{ij} are probabilities.
2. $$\sum_j p_{ij} = \sum_j \mathbb{P}(X_{n+1} = j \mid X_n = i) = 1$$
   since X_{n+1} must take *some* value.

□

Example

Let X_n denote the state of a computer at the start of the nth day. The computer can be in either of two states: $X_n = 0$ if it is broken or $X_n = 1$ if in working order.
Let $\pi_n(0) = \mathbb{P}(X_n = 0)$ and $\pi_n(1) = \mathbb{P}(X_n = 1) = 1 - \pi_n(0)$.
Let the transition probabilities be

$$\mathbb{P}(X_{n+1} = 1 \mid X_n = 0) = p \qquad \mathbb{P}(X_{n+1} = 0 \mid X_n = 0) = 1 - p$$
$$\mathbb{P}(X_{n+1} = 0 \mid X_n = 1) = q \qquad \mathbb{P}(X_{n+1} = 1 \mid X_n = 1) = 1 - q$$

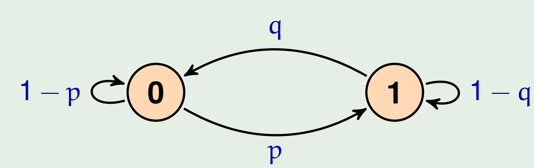
(Notice that $p + q$ need not equal 1!)
Therefore the transition matrix is

$$\mathbf{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

A typical question is: *What is* $\mathbb{P}(X_{n+1} = 0)$*?*
We will answer this naively at first.

Example (Continued)

We often represent Markov chains graphically; e.g.,



which allows us to read the transition probabilities at a glance and write down the transition matrix:

$$\mathbf{P} = \begin{pmatrix} 0 \to 0 & 0 \to 1 \\ 1 \to 0 & 1 \to 1 \end{pmatrix} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

Example (Continued)

$$
\begin{aligned}
\mathbb{P}(X_{n+1}=0) &= \mathbb{P}(X_{n+1}=0 \mid X_n=0)\mathbb{P}(X_n=0) \\
&\quad + \mathbb{P}(X_{n+1}=0 \mid X_n=1)\mathbb{P}(X_n=1) \\
&= (1-p)\pi_n(0) + q\pi_n(1) \\
&= (1-p)\pi_n(0) + q(1-\pi_n(0)) \\
\therefore \quad \pi_{n+1}(0) &= (1-p-q)\pi_n(0) + q \\
\text{e.g.} \quad \pi_1(0) &= (1-p-q)\pi_0(0) + q \\
\pi_2(0) &= (1-p-q)\left((1-p-q)\pi_0(0)+q\right) + q \\
&= (1-p-q)^2\pi_0(0) + q(1+(1-p-q)) \\
\implies \quad \pi_n(0) &= (1-p-q)^n\pi_0(0) + q\sum_{j=0}^{n-1}(1-p-q)^j
\end{aligned}
$$

Example (Continued)

Let us assume that $p+q>0$, otherwise $\pi_n(0)=\pi_0(0)$ for all n. Then

$$
\begin{aligned}
\pi_n(0) &= (1-p-q)^n\pi_0(0) + q\left(\frac{1-(1-p-q)^n}{p+q}\right) \\
&= (1-p-q)^n\left(\pi_0(0) - \frac{q}{p+q}\right) + \frac{q}{p+q}
\end{aligned}
$$

and similarly

$$
\pi_n(1) = (1-p-q)^n\left(\pi_0(1) - \frac{p}{p+q}\right) + \frac{p}{p+q}
$$

In other words, the probability of finding the machine in any given state on the nth day, depends only on the initial probabilities and the transition probabilities.

Example (Continued)

It turns out that we can arrive at the same result in a more automatic way using the transition matrix.

Let $\boldsymbol{\pi}_n = (\pi_n(0), \pi_n(1))$ be the row vector of probabilities. Then

$$
\begin{aligned}
\boldsymbol{\pi}_n \mathbf{P} &= (\pi_n(0), \pi_n(1)) \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \\
&= ((1-p)\pi_n(0) + q\pi_n(1), p\pi_n(0) + (1-q)\pi_n(1)) \\
&= ((1-p-q)\pi_n(0) + q, p + (1-p-q)\pi_n(1)) \\
&= \boldsymbol{\pi}_{n+1} (!)
\end{aligned}
$$

Therefore (see next lecture for a general proof)

$$
\boldsymbol{\pi}_n = \boldsymbol{\pi}_0 \underbrace{\mathbf{P}\dots\mathbf{P}}_{n} = \boldsymbol{\pi}_0 \mathbf{P}^n
$$

Summary

- Non-deterministic processes are subject to probabilistic analysis.
- A **stochastic process** is a collection of random variables indexed by "time" taking values in a **state space**, interpreted as the state of the system at a given time.
- **Markov chains** are discrete-time stochastic processes with countable states satisfying the **Markov property**: "given the present, the future does not depend on the past".
- (Temporally) homogeneous Markov chains are described by **transition matrices**, whose entries are the **transition probabilities**: non-negative and rows sum to 1.
- Random walks are examples of Markov chains.
- In a Markov chain, the probability of finding the system in a given state at a given time is determined by the transition probabilities and the initial probabilities.
- Finite-state Markov chains can be represented graphically.