# Dynamical systems and stochastic processes in ecology I

Alan McKane

Complex Systems Group, School of Physics and Astronomy, University of Manchester

Trieste, October 25, 2010

# OUTLINE

- Introduction
- Stochastic processes
- Markov chains
- The Wright-Fisher model
- The master equation

# 1.1 Introduction

Many quantitative models in science are *dynamical*, that is, the model consists of a set of rules so that the state of the system at time $t+1$ can be found from the state of the system at time $t$

Example: The logistic map, which in some situations tells us how the size of a population $x$ changes with time:

$$x(t+1) = r\,x(t)\left[1 - \frac{x(t)}{k}\right]$$

So if we know the state of the system $x$ at time $t$ this gives us a rule to find it at time $t+1$

• These are deterministic equations. If we know $x(0)$ then by iteration we can in principle find $x(t)$ for any $t$

• In a computer simulation we might take $\delta t = 1$ (time-step of unity) and so relate $x(t+1)$ to $x(t)$. In analytic work would take $\delta t \to 0$ to get differential equations

- The state space in the above is continuous: $x$ is a real variable, but it could also be discrete: $n = 0, 1, 2, \ldots$. For example, the simple exponential growth of organisms with time: $n(t+1) = rn(t)$

- In fact, a rule like $n(t+1) = rn(t)$ would not be true except on average. This is because births and deaths do not take place at regular fixed intervals. Instead there is a large element of randomness as to when they happen — they are examples of *stochastic events*

- So an example of a stochastic, rather than a deterministic, rule is the growth in the size of a population:

$$\begin{aligned} n(t+1) &= n(t) + 1 \text{ with probability } \beta \\ n(t+1) &= n(t) - 1 \text{ with probability } \delta \\ n(t+1) &= n(t) \text{ with probability } 1 - \beta - \delta \end{aligned}$$

Here $\beta$ is the probability of a birth in a single time interval and $\delta$ the probability of a death in a single time-interval

# 1.2. Stochastic Processes

A process where the state of the system at time $t+1$ (or $t+\delta t$) is a stochastic (random) variable which depends on the state of the system at earlier times, is called a stochastic process.

In this case we can only specify the probability of finding the system in a given state at time $t$. So to describe how the system evolves with time, we have to give the probability distribution function or pdf $P(n,t)$ or more accurately
$P(n,t|n_0,0)$ — the conditional probability distribution that the system is in state $n$ at time $t$ given that it was in state $n_0$ at time $t=0$

The dynamics will usually describe how pdf's change in time, not how the state variables themselves change in time

If the stochastic process is such that the state of the system at time $t+1$ only depends on the state of the system at time $t$, then the process is a Markov process

For a Markov process the history of the system, apart from the current state of the system, is forgotten. So if the rule to determine $n(t+1)$ depends on what $n(t-1)$ was, as well as $n(t)$, then the process is not Markov. So, for instance, for a Markov process $P(n_2, t_2|n_1, t_1; n_0, t_0) = P(n_2, t_2|n_1, t_1)$

There are only two independent pdfs for Markov processes: $P(n, t)$ and $P(n', t'|n, t)$

The state space in the above is discrete: $n = 0, 1, 2, \ldots$, but it may also be continuous: $P(x, t|x_0, 0)$, where $x$ is a real variable.

<u>Aside</u>: The conditional pdfs for Markov processes satisfy two equations:

$$\begin{aligned} P(n_3, t_3|n_1, t_1) &= \sum_{n_2} P(n_3, t_3|n_2, t_2)\, P(n_2, t_2|n_1, t_1) \\ P(n_2, t_2) &= \sum_{n_1} P(n_2, t_2|n_1, t_1)\, P(n_1, t_1) \end{aligned}$$

The first relation is called the Chapman-Kolmogorov equation

# 1.3. Markov chains

Suppose the system has just two states $n = 1$ and $n = 2$. We would like to be able to calculate $P(n, t+1)$ from $P(n, t)$ given the *transition probabilities* between the states

If $Q_{n\,n'}$ is the (time-independent) probability that there is a transition from state $n'$ to state $n$ in a single time interval, then

$$\begin{aligned} P(1, t+1) &= Q_{11} P(1, t) + Q_{12} P(2, t) \\ P(2, t+1) &= Q_{21} P(1, t) + Q_{22} P(2, t) \end{aligned}$$

In matrix form:

$$\begin{pmatrix} P(1, t+1) \\ P(2, t+1) \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \begin{pmatrix} P(1, t) \\ P(2, t) \end{pmatrix}.$$

This is called a *Markov chain* and can clearly be generalised to an arbitrary number of states, and not just 2 as in the above example

The ingredients for a Markov chain are firstly a *probability vector*

$$\underline{P}(t) = \left( \begin{array}{c} P(1,t) \\ P(2,t) \\ P(3,t) \\ \ldots \end{array} \right)$$

and secondly a *transition matrix*

$$Q = \left( \begin{array}{ccc} Q_{11} & Q_{12} & \ldots \\ Q_{21} & Q_{22} & \ldots \\ \ldots & \ldots & \ldots \end{array} \right)$$

The time evolution of the system is then given by the matrix equation

$$\underline{P}(t+1) = Q\underline{P}(t) \ \ \mathrm{or} \ \ P(n,t+1) = \sum_{n'} Q_{n\,n'} P(n',t)$$

Note that the sum of the components of the probability vector must add up to 1, and the sum of the columns of the transition matrix must add up to 1 (since $\sum_n P(n,t) = 1$ and $\sum_n Q_{n\,n'} = 1$)

In general we are given the transition matrix and we wish to calculate the probability that the system is in a given state at time $t$, that is, we wish to calculate the probability vector at time $t$

It is straightforward to formulate this mathematically:

$$\begin{aligned} \underline{P}(t) &= Q\,\underline{P}(t-1) = Q\,Q\,\underline{P}(t-2) = Q^2\,Q\,\underline{P}(t-3) = \ldots \\ &= Q^t\,\underline{P}(0) \end{aligned}$$

So if we are given the initial state of the system ($\underline{P}(0)$), then we can find the state of the system at time $t$ ($\underline{P}(t)$) by matrix multiplication by the $t$th power of the transition matrix

Unless $t$ is small, it is clearly not feasible to find $\underline{P}(t)$ by direct matrix multiplication. Instead the set of eigenvalues $\{\lambda_1, \lambda_2, \ldots\}$ of $Q$ together with the corresponding eigenvectors are found

From the fact that $Q^t$ has the same eigenvectors, but eigenvalues $\{\lambda_1^t, \lambda_2^t, \ldots\}$, the form of $\underline{P}(t)$ may be found for large $t$

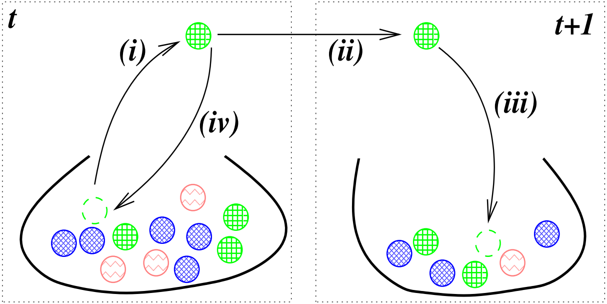
# 1.4. The Wright-Fisher model

An example of a Markov chain is the Wright-Fisher model of *random genetic drift*

Suppose we have a population of individuals at time $t$ who mate randomly to produce a new generation at time $t+1$. We focus on one particular gene which may be in one of two states (alleles) which we call $A$ and $B$. The Wright-Fisher model is based on the idea of the gene pool of the $t$ generation individuals being sampled to give the genetic structure of the $t+1$ generation.

In reality, individuals are diploid (have two sets of genes), but for convenience we assume that individuals are haploid, that is, there is 1-1 correspondence between individuals and genes

The process is analogous to sampling coloured balls — representing different alleles (types of gene) — from a bag or urn



The population in generation $t+1$ is constructed from generation $t$ by (i) selecting a gene from the current generation at random, (ii) copying this gene, (iii) placing a copy in the next generation, (iv) returning the original to the parent population. These steps are repeated until generation $t+1$ has the same size population as generation $t$

Suppose that at time $t$ there are $n'$ $A$ alleles and $(N - n')$ $B$ alleles, and let $P(n', t)$ be the probability that the system is in state $n'$ at time $t$ (or generation $t$).

The probability that in $N$ trials of sampling the pool to select a gene we will get $n$ $A$ alleles is

$$\binom{N}{n} p_1^n p_2^{N-n}$$

where $p_1$ is the probability of picking an $A$ $(= n'/N)$ and $p_2$ is the probability of picking a $B$ $(= (N - n')/N)$

$$\Rightarrow \quad Q_{n\,n'} = \binom{N}{n} \left(\frac{n'}{N}\right)^n \left(1 - \frac{n'}{N}\right)^{N-n}$$

Note that here $n$ and $n'$ run from 0 to $N$ and that the formula works at the boundaries: $Q_{n\,0} = 0$, unless $n = 0$ when it is 1, and similarly $Q_{n\,N} = 0$, unless $n = N$ when it is 1

# 1.5. The master equation

This is a Markov chain, but in continuous time. So instead of time intervals being $1, 2, \ldots$ they are $\delta t, 2\delta t, \ldots$. Therefore transition probabilities are replaced by transition *rates*

Suppose that the system consists of a set of states labelled $n = 1, 2, \ldots$. There may be a finite number of possible states ($N$) or an infinite number. Then,

Rate of change of $P(n, t) =$

[ Rate due to transitions into the state $n$ from all other states $n'$ ]

- [ Rate due to transitions out of the state $n$ into all other states $n'$ ]

Here we have suppressed the dependence on the initial state of the system: $P(n, t|n_0, 0)$ has become $P(n, t)$. Also, in what follows, we'll denote the transition rate from state $n$ to state $n'$ to be $T(n'|n)$

Then the probability of finding the system in state $n$ at time $t$ changes according to the master equation:

$$\frac{dP(n,t)}{dt} = \sum_{n' \neq n} T(n|n')\, P(n',t) - \sum_{n' \neq n} T(n'|n)\, P(n,t)\,.$$

The $T(n|n')$ are assumed given (this specifies the model) and we wish to determine the $P(n,t)$. The initial condition is that the system is in state $n_0$ at $t = 0$. Boundary conditions may also have to be given

For a *one-step processes*, transitions only take place between states whose label differs by one. That is, $T(n|n')$ and $T(n'|n)$ are zero unless $n' = n + 1$ and $n' = n - 1$

Then the master equation takes the simpler form

$$\begin{aligned} \frac{dP(n,t)}{dt} &= T(n|n+1)P(n+1,t) + T(n|n-1)P(n-1,t) \\ &- \{T(n-1|n) + T(n+1|n)\}\; P(n,t) \end{aligned}$$

## Examples

1) A simple example is the asymmetric random walk

$$\begin{aligned} T(n+1|n) &= \lambda \ \text{(move in the } + \text{ve direction)} \\ T(n-1|n) &= \mu \ \text{(move in the } - \text{ve direction)} \end{aligned}$$

2) The Moran model

This is a one-step process version of the Wright-Fisher model. Whereas in the Wright-Fisher model the gene pool of the $t$ generation individuals was sampled to give the genetic structure of the $t+1$ generation, in the Moran model the generations are overlapping

At each time step a gene is selected at random and copied. Another gene is selected at random and eliminated. The copied gene and the original are then placed back into the population

So the transition rates for the Moran model are ($n$ is the number of $A$s)

$$\begin{aligned} T(n+1|n) &= \left(\frac{n}{N}\right)\left(1-\frac{n}{N}\right) \ \text{(}A\text{ copied , }B\text{ eliminated)} \\ T(n-1|n) &= \left(1-\frac{n}{N}\right)\left(\frac{n}{N}\right) \ \text{(}B\text{ copied , }A\text{ eliminated)} \end{aligned}$$