Hidden Markov and related models as powerful and versatile devices for modelling ecological time series

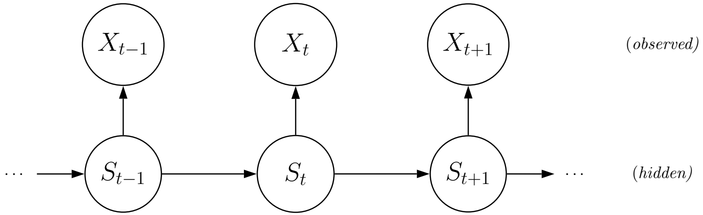
Roland Langrock

University of St Andrews

A quick overview

HMMs – a brief overview

- versatile & mathematically tractable **time series model**
- two (discrete-time) stochastic processes, one of them hidden
- **hidden state process** is an N-state **Markov chain**
- distribution of observations determined by underlying state



- a useful source: *Zucchini and MacDonald (2009, Chapman & Hall)* – watch out for the 2nd Edition, due to appear next year!

HMMs – model formulation

A basic HMM involves

1.) the initial state probabilities $\delta_i = \Pr(S_1 = i)$, $i = 1, \dots, N$

2.) the **state transition probabilities** $\gamma_{ij} = \Pr(S_{t+1} = j \mid S_t = i)$, $i, j = 1, \dots, N$, summarized in the t.p.m.

$$\boldsymbol{\Gamma} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1N} \\ \vdots & \ddots & \vdots \\ \gamma_{N1} & \cdots & \gamma_{NN} \end{pmatrix}$$

3.) the **state-dependent distributions** $f(x_t|s_t)$, e.g.

- Poisson, geometric, negative binomial (for count data)
- normal, gamma, Weibull, ... (for continuous observations)
- ...
- combinations of these, e.g. gamma for step lengths and von Mises for turning angles in an animal movement model

HMM machinery

HMMs – likelihood calculation using brute force

$$
\begin{aligned}
\mathcal{L}^{\text{HMM}} &= f(x_1, \ldots, x_T) \\
&= \sum_{s_1=1}^{N} \ldots \sum_{s_T=1}^{N} f(x_1, \ldots, x_T, s_1, \ldots, s_T) \\
&= \sum_{s_1=1}^{N} \ldots \sum_{s_T=1}^{N} f(x_1, \ldots, x_T | s_1, \ldots, s_T) f(s_1, \ldots, s_T) \\
&= \sum_{s_1=1}^{N} \ldots \sum_{s_T=1}^{N} \delta_{s_1} \prod_{t=1}^{T} f(x_t | s_t) \prod_{t=2}^{T} \gamma_{s_{t-1}, s_t}
\end{aligned}
$$

Simple form, but N^T summands, numerical maximization of *this* expression thus infeasible.

HMMs – likelihood calculation via forward algorithm

Consider instead the so-called **forward probabilities**,

$$\alpha_t(j) = f(x_1, \ldots, x_t, s_t = j).$$

These can be calculated using an **efficient recursive scheme**:

$$\begin{aligned} \boldsymbol{\alpha}_1 &= \boldsymbol{\delta}\mathbf{P}(x_1) \\ \boldsymbol{\alpha}_{t+1} &= \boldsymbol{\alpha}_t \boldsymbol{\Gamma}\mathbf{P}(x_{t+1}) \end{aligned}$$

with $\mathbf{P}(x_t) = \text{diag}\big(f(x_t|s_t = 1), \ldots, f(x_t|s_t = N)\big)$ and t.p.m. $\boldsymbol{\Gamma}$.

$$\Rightarrow \quad \boxed{\mathcal{L}^{\text{HMM}} = \boldsymbol{\delta}\mathbf{P}(x_1)\boldsymbol{\Gamma}\mathbf{P}(x_2) \cdot \ldots \cdot \boldsymbol{\Gamma}\mathbf{P}(x_T)\mathbf{1}}$$

Computational effort **linear** in T!

HMMs – example code

R code for computing the log-likelihood of a gamma HMM:

```
loglik<-function(x,delta,Gamma,pshape,pscale){
 llk<-0
 foo<-delta
 for (t in 1:length(x)){
  foo<-foo%*%Gamma*dgamma(x[t],pshape,1/pscale)
  llk<-llk+log(sum(foo)); foo<-foo/sum(foo)
 }
 return(llk)
}
```

A big advantage over alternative estimation techniques (EM or MCMC): modifications of the model usually require only minimal changes in the code.

Estimation times for a simple gamma HMM

Example times required to numerically maximize $\mathcal{L}^{\mathrm{HMM}}$:

	N=2	N=3	N=4
T=200	0.3s	3s	10s
T=2000	2s	13s	29s
T=20000	21s	107s	284s

Computational speed is the second big advantage over alternative estimation techniques.

Other inferential issues

- uncertainty quantification
  $\longrightarrow$ bootstrap or Hessian-based
- model selection
  $\longrightarrow$ information criteria
- model checking
  $\longrightarrow$ pseudo-residuals, simulation-based, ...
- state decoding
  $\longrightarrow$ Viterbi algorithm

Ecological applications

Animal movement modelling

- one of the standard movement models (an HMM!):
  - *N* behavioural states, switching governed by Markov chain
  - e.g. von Mises and gamma state-dependent distributions for turning angles and step lengths, respectively

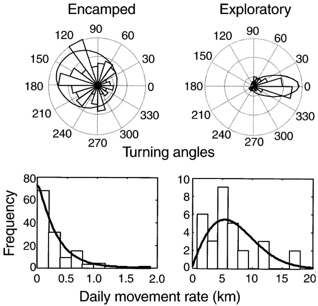
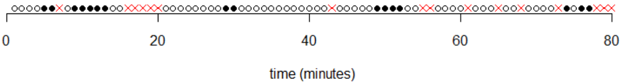


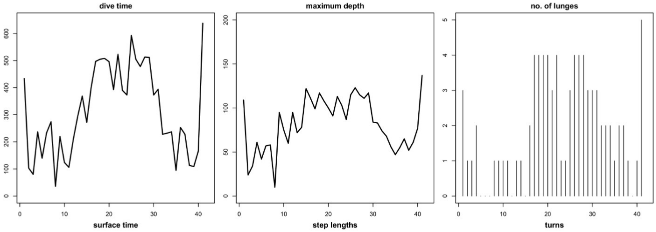
Figure: Turning angle and step length distributions for an elk in two behavioural states (taken from *Morales et al., 2004, Ecology*)

General animal behaviour

- The same type of model can be (and has been) applied to various other aspects of animal behaviour, e.g.
  - to model **feeding behaviour** (feeding vs. not feeding)



  - to model/classify **whale dive types** (shallow vs. deep, as indicated e.g. by dive duration or maximum depth)



Capture-recapture

- a capture-recapture **encounter history** such as

  $$1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0$$

  (0: not seen; 1: seen alive)

  can be regarded as the **outcome of an HMM**, with
  - the states corresponding to the animal's survival state, so that

    $$\boldsymbol{\Gamma} = \begin{pmatrix} \phi & 1-\phi \\ 0 & 1 \end{pmatrix},$$

  - Bernoulli state-dep. distribution for state 1
    (the probability of success being the recapture probability)

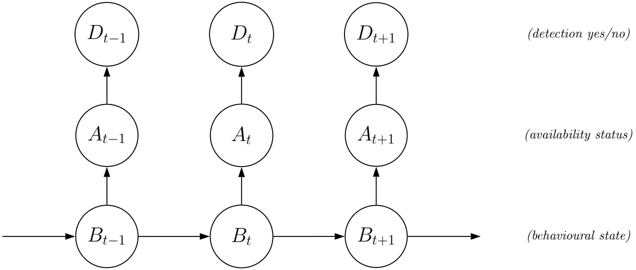
- more or less straightforward extensions:
  - capture-recapture-recovery data, multi-state (Arnason-Schwarz) models, multi-state models with state uncertainty, various types of covariates

Occupancy modelling

- see **Olivier's talk** later in this session!

Availability bias in distance sampling for marine mammals

- **Aim**: a model that accounts for non-detection due to both
  1) animals being unavailable for detection (submerged) and
  2) available animals not being detected



- N states (corresponding to "diving", "resting", ...)
- D_t depends on A_t and a covariate, namely the distance of the animal to observer at time t

Modelling of population dynamics (illustration)

- S_t: true (unknown) number of individuals at time t, and e.g.

$$S_t = S_t^* + N_t,$$

  where $S_t^* \sim \text{Binomial}(S_{t-1}, \phi)$ and $N_t \sim \text{Poisson}(\lambda \phi S_{t-1})$
- specifying some upper bound for $\{S_t\}$, this is a Markov chain (with many states, yet determined by only two parameters)
- observation process: animals seen, conditional on states,

$$X_t \mid S_t = j \sim \text{Binomial}(j, p_t)$$

- (various other formulations fit into the HMM framework)

Extensions & related models

Some extensions of the basic HMM setting

- covariates, seasonality $\longrightarrow$ straightforward
- random effects $\longrightarrow$ conceptually straightforward, but computationally challenging
- semi-Markov state processes $\longrightarrow$ simple using a trick
- feedback, and in fact some other modifications to the dependence structure $\longrightarrow$ straightforward
- continuous-valued state processes $\longrightarrow$ simple discretization renders HMM machinery applicable
- nonparametric approaches $\longrightarrow$ current research...