Introduction to Hidden Markov Models (HMMs)

Dr. Vianey Leos Barajas -- University of Toronto
Sofia Ruiz Suarez -- INIBIOMA-CONICET/Universidad Nacional de Rosario

RIIAA

MEETING ON ARTIFICIAL INTELLIGENCE AND ITS APPLICATIONS

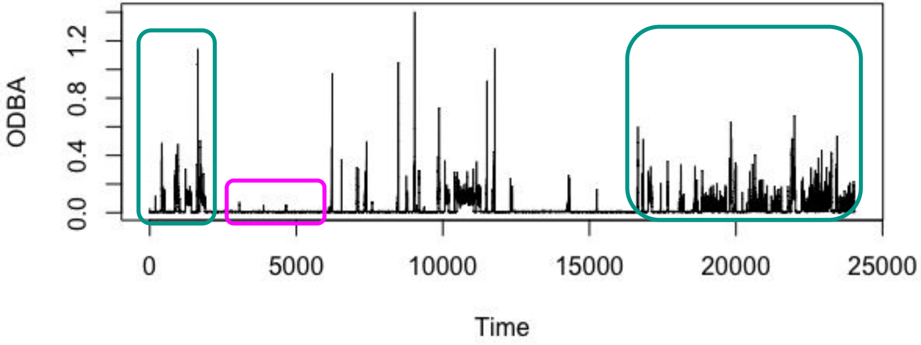
25-27 August 2021

Hidden Markov Models for Time Series

A **hidden Markov model (HMM)** is used to model data collected over time (or in sequence) where we can make two key assumptions:

- a set of *N* different patterns manifest in the data
- there is persistence in the pattern over time (or sequence)

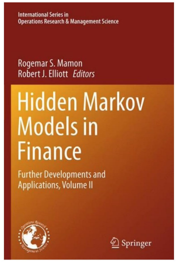
Example:



Where have HMMs been used?

Finance

market regimes



Medicine

disease progression

Ecology

animal movement

capture recapture

General Pattern Recognition

speech recognition

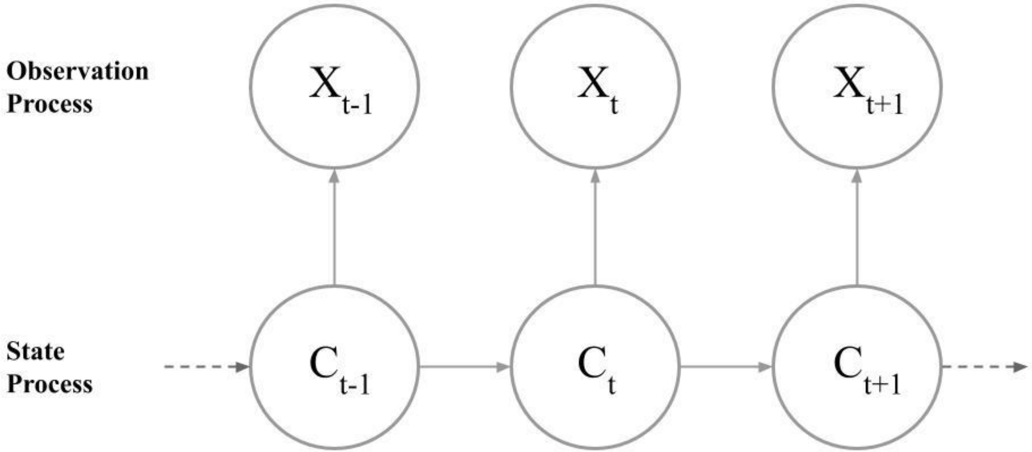
human movement

Hidden Markov Models (HMMs)

HMMs: A doubly stochastic process composed of an

observation process: $X_{t=1}^{T}$ and a **state process:** $C_{t=1}^{T}$

where the state process $C_{t=1}^{T}$ is taken to be a Markov chain



Data Types: *Observations + States*

$X_{t=1}^{T}$ can be:

- univariate
- multivariate
- continuous
- discrete
- measured with error

$C_{t=1}^{T}$ is taken to be:

- integer-valued (1, 2, 3,....)

Discrete-time HMM: $X_{t=1}^{T}$ is assumed to be observed regularly over time

Finite-state HMM: $C_{t=1}^{T}$ can take on a finite set of integer values

Components of an HMM

A basic (discrete-time, finite-state) HMM is fully specified by 4 components:

Mathematical Description:

Number of states:

$$N \in \mathbb{N}$$

State-dependent distributions:

$$\{f(X_t|C_t = n)\}_{n=1}^{N}$$

Transition probability matrix (t.p.m.):

$$\Gamma_{i,j} = \Pr\left(C_t = j | C_{t-1} = i\right) \quad \text{for } i,j \in \{1,\ldots,N\}$$

Initial state distribution:

$$\delta_i = \Pr\left(C_1 = i\right) \quad \text{for } i \in \{1,\ldots,N\}$$

Components of an HMM (cont.)

General Interpretation

Number of states

Translation: Number of patterns

Transition probability matrix

Translation: How the patterns evolve over time

State-dependent distributions:

Translation: Distribution of observations within each pattern

Initial state distribution

Translation: What pattern occurred at the first time point?

State dwell time distribution

As the state process is taken to be a Markov chain, that means that there is an expected amount of time a state will 'last' until it switches to another state process over time.

$$d_n \sim Geometric(1 - \Gamma_{nn})$$

The expected number of times that a state will be exhibited over time before there is a switch is $\mathrm{E}(d_n) = 1/(1 - \Gamma_{nn})$

Example

Number of states: $N = 3$

State-dependent distributions:

$$f(X_t|C_t = n) \sim N(\mu_n, \sigma_n) \quad \text{for } n \in \{1, 2, \ldots, N\}$$

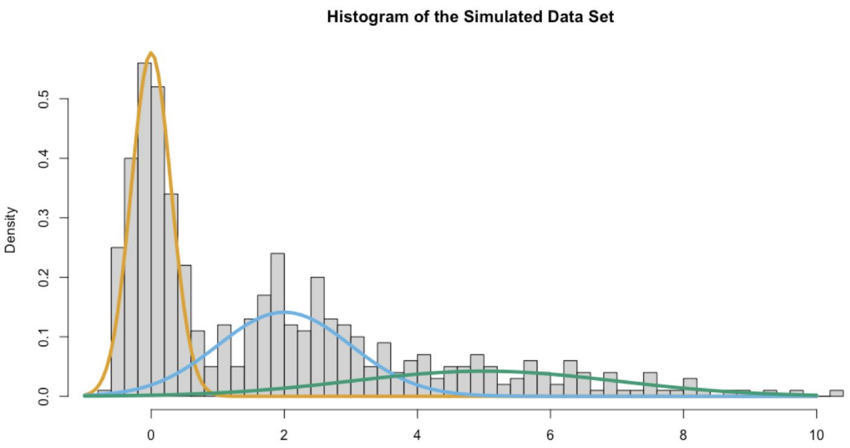
$$\{\mu_1 = 0, \sigma_1 = 0.3\}\{\mu_2 = 2, \sigma_2 = 1\}\{\mu_3 = 5, \sigma_3 = 2\}$$

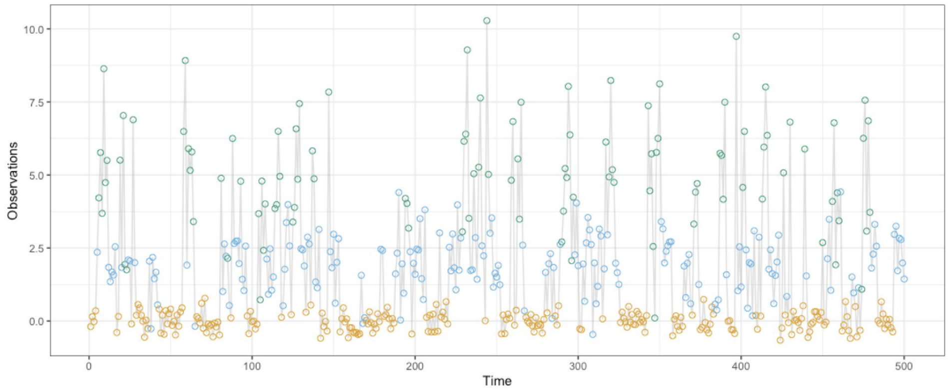
Transition probability matrix:

$$\mathbf{\Gamma} = \begin{bmatrix} .8 & 0.1 & 0.1 \\ 0.2 & 0.7 & 0.1 \\ 0.05 & 0.35 & 0.6 \end{bmatrix}$$

Initial state distribution:

$$\boldsymbol{\delta} = [1/3, 1/3, 1/3]$$





How to simulate data from an N-state HMM

1. Sample C_1 from the initial state distribution $\boldsymbol{\delta}$

2. Given $C_1 = i_1$, sample C_2 using $\boldsymbol{\Gamma}_{i,\cdot}$

3. Repeat step 2 for $C_3, \ldots, C_T$

4. Given $\{C_1 = i_1, C_2 = i_2, \ldots, C_T = i_T\}$, for $t = 1, \ldots, T$ simulate from

   $f(X_t | C_t = i_t)$

Simulation in R