Adam, Koslik April 26, 2024

Hidden Markov Models - Practical Session 3

Exercise 1 - Markov chains, Part 2

a) Write a function `simMC(T, N, Gamma, delta)` which generates a state vector of length T from the N-state Markov chain characterized by its t.p.m. $\mathbf{\Gamma}$ and its initial distribution $\boldsymbol{\delta}$.

b) Use your function to generate a 4-state Markov chain of length $T = 300$ choosing your own values for $\mathbf{\Gamma}$ and $\boldsymbol{\delta}$.

c) Write a function `getDelta(Gamma, N)` which calculates the stationary distribution given the t.p.m. of a Markov chain.

d) Use your function to find the stationary distribution for a Markov chain with:

$$\mathbf{\Gamma} = \begin{pmatrix} 0.9 & 0.05 & 0.05 \\ 0.05 & 0.9 & 0.05 \\ 0.05 & 0.05 & 0.9 \end{pmatrix}$$

Exercise 2 - Simulating from an HMM

To become more familiar with HMMs and the data they produce, this exercise focuses on generating samples from an HMM with 3 states. Remember that an N-state HMM consists of a (hidden) state process $S_1, \dots, S_T$ taking values in $\{1, \dots, N\}$, which is a (homogeneous) Markov chain, and an (observed) state-dependent process $X_1, \dots, X_T$. The distribution generating x_t only depends on the current state s_t.
The goal is to generate a vector of observations `x` of length $T = 300$ from a 3-state HMM. For an example (including R code) of a simulated HMM with 2 states see lecture slides 69-73.

a) Consider a Markov chain with the t.p.m. $\mathbf{\Gamma}$ defined in exercise 1d) starting in its stationary distribution. The 3 state-dependent distributions to generate the observations `x` are defined as:

$$X_t|S_t = 1 \sim \mathcal{N}(\mu_1 = -2, \sigma_1^2 = 0.7^2)$$
$$X_t|S_t = 2 \sim \mathcal{N}(\mu_2 = 0, \sigma_2^2 = 0.7^2)$$
$$X_t|S_t = 3 \sim \mathcal{N}(\mu_3 = 2, \sigma_3^2 = 0.7^2)$$

Generate data of length $T = 300$ from this 3-state hidden Markov model.
Hint: First, generate the Markov chain `s`. Then generate the state-dependent observation vector `x` using the function `rnorm(...)` and the current state `s[t]`.

b) Plot the observation sequence from a). Identify a couple of observations for which it would be difficult based on the plot alone to guess which state they were generated from. Then use colour code within the `plot()` function to indicate the states underlying the observations.

c) Generate and plot additional vectors of observations, varying the parameters of both the t.p.m. and of the state-dependent distributions. What are circumstances that facilitate the decoding of the underlying states by eye?

d)* Building on the previously considered function – `simMC(T,N,Gamma,delta)` – write a function `simNormHMM(T,N,Gamma,delta,mu,sigma)` in R which generates a sequence of T values from the N-state hidden Markov model characterized by its t.p.m. $\mathbf{\Gamma}$, its initial distribution $\boldsymbol{\delta}$ and the state-dependent normal means $\boldsymbol{\mu}$ and standard deviations $\boldsymbol{\sigma}$. Then simulate a 4-state HMM choosing your own parameter values.