Markov Process

October 9, 2023

1 Markov processes

A Markov process can be used to describe the dependency between consecutive data samples. This is done using the state-transition matrix

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$$

1.1 Markov model for the weather of the Land of Oz

Consider that the weather in the land of Oz can be one of n=3 states: x_{1}=R: rainy weather, x_{2}=N: nice weather, and x_{3}=S: snowy weather. The state diagram of a stationary Markov process that is characterized by its state-transition matrix

$$P = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/4 & 1/4 & 1/2 \end{bmatrix}$$

where the elements i, j specify the probability

$$P_{ij} = P[X(k) = x_j(k)|X(k-1) = x_i(k-1)]$$

that the next state is x_j when you are currently $(k-1)$ in state x_i.

Remark that this matrix contains probabilities on each row (each row corresponding to a different current state x_i.) Hence,

$$\sum_{j=1}^{n} P_{ij} = 1 \qquad \forall i.$$

The matrix is therefore called to be a row stochastic matrix.

Define the matrix P in numpy and show that it is a row stochastic matrix.

```
Stochastic matrix P=
[[0.5  0.25 0.25]
 [0.5  0.   0.5 ]
 [0.25 0.25 0.5 ]]
Checking if the sum over all rows equal to one?
[1. 1. 1.]
[0.5  0.25 0.25]
```

1.2 Generate a realization of the weather of the land of Oz

Generate and print a realization of length 20 of the weather of the land of Oz starting from the initial state: the nice weather.

```
Generated realization: [1, 0, 0, 0, 2, 2, 0, 0, 0, 1, 0, 2, 2, 1, 2, 2, 0, 0, 0,
1]
```

1.3 Probability to be in a state

To determine the entropy of a Markov source, it is necessary to determine the probability to be in a given state. Using the Markov process generator defined before, generate the a realization a sequence of 100000 state and determine the probability of have rainly, nice and snowy weather. This corresponds with a finite sample estimation of the probability vector π introduced in the theory.

Note that for different realizations, one obtains similar results for the probabilty of being in a certain state (try running this algorthm different times and look to the results)

```
p[0] = 0.39460
p[1] = 0.20140
p[2] = 0.40400
```

1.4 Properties stochastic matrix

It is known that for a row stochastic matrix such as P that its largest left eigenvalue $\lambda_1 = 1$. If there is a unique stationary distribution (only the largest eigenvalue equals to one, or mathematically $1 = \lambda_1 > \lambda_2$), then the largest eigenvalue is unique (=1) and hence, the corresponding eigenvector ($W_{[1,:]}$) are both unique

$$P = W^{-1}DW$$

with W the matrix of eigenvectors and D the diagonal matrix of *left* eigenvalues.

First, compute the eigenvalues and eigenvectors using the linalg.eig function. As the eig function computes the *right* eigenvalues/eigenvector problem, we have to work on the transposed probability matrix.

Verify that exactly one eigenvalue equals 1 and print the eigenvectors (but don't forget the transpose operation)

```
Eigenvalues:
[ 1.    0.25 -0.25]
Eigenvectors:
[[-6.66666667e-01 -3.33333333e-01 -6.66666667e-01]
 [-7.07106781e-01  1.66935271e-16  7.07106781e-01]
 [ 4.08248290e-01 -8.16496581e-01  4.08248290e-01]]
```

Note that one eigenvalue is equal to 1 while all others are $|\lambda_i| < 1$ for all $i = 2, ..., n$.

Check that the eigenvector coresponding to the largest eigenvalue is not representing a probability vectors with $0 \le p_i \le 1$ and

$$\sum_{i=1}^{n} p_i = 1$$

but with a quadratic norm scaling

$$\sum_{i=1}^{n} w_i^2 = 1.$$

```
Eigenvector:
[-0.66666667 -0.33333333 -0.66666667]
2-norm = 1.00000
```

It is possible to scale an eigenvector with an arbitrary constant. Determine the scaling of the eigenvector such that it becomes an probability vector, which is mathematically represented by π.

```
Probability vector:
[0.4 0.2 0.4]
```

1.5 Entropy of a Markov process

The entropy of a (stationary and ergodic) Markov process equals

$$H(X) = -\sum_{i=1}^{n}\sum_{j=1}^{n} \pi_i P_{ij} \log P_{ij}.$$

Determine the entropy of the considered Markov process.

```
Entropy of individual processes
[[-5.00000000e-01 -5.00000000e-01 -5.00000000e-01]
 [-5.00000000e-01 -5.31508495e-15 -5.00000000e-01]
 [-5.00000000e-01 -5.00000000e-01 -5.00000000e-01]]
Entropy of Markov process = 1.40000 bit
```

1.6 Covergence to the asymptotic probability

The vector π gives use the probabiltiy $p(x_i)$ that the Markov process is in state x_i. The assumption that the Markov process is stationary and ergodic implies that there is only the largest eigenvalue equal is to one. Due to this, the process will always converge to this distribution, irrespective of probability distribution at the initial time of $k = 0$.

To illustrate this, consider an arbitrary probability distribution, perform 5 successive steps in the Markov process, and compare the final probability vector with the asymptotic one.

```
Initial probability vector:
[0.59614929 0.16572488 0.23812583]
Iteration 0
[0.44046854 0.20856878 0.35096268]
Iteration 1
[0.41225933 0.1978578  0.38988286]
Iteration 2
[0.40252928 0.20053555 0.39693517]
Iteration 3
[0.40076621 0.19986611 0.39936768]
Iteration 4
[0.40015808 0.20003347 0.39980845]
```