Machine Learning for Graphs and Sequential Data Exercise Sheet 06

Autoregressive Models, Markov Chains

Problem 1: Consider the stationary AR(p) process $X_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma^2)$. We denote by μ the mean $E[X_t]$ and by γ_i the autocovariance $Cov(X_t, X_{t-i})$. Show:

1. $\mu = \frac{c}{1-\sum_{i=1}^{p} \phi_i}$, for all t
2. $\gamma_0 = \sum_{j=1}^{p} \phi_j \gamma_{-j} + \sigma^2$
3. $\gamma_i = \sum_{j=1}^{p} \phi_j \gamma_{i-j}$, for all $t, i \in [1, p]$

Problem 2:

a) Consider the following AR(1) process $X_t = c + \phi_1 X_{t-1} + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Show that this process is stationary iff the following condition is fulfilled

$$|\phi_1| < 1.$$

b) Consider the following AR processes:

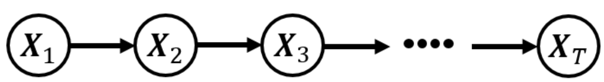
- $X_t = c + .8 \times X_{t-1} + .1 \times X_{t-2} + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma^2)$
- $X_t = -\sum_{k=1}^{p} \binom{p}{k} X_{t-k} + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma^2)$

Are these processes stationary ?

Problem 3: Let $\mathbf{X}_t$ be a 2-D random vector:

$$\mathbf{X}_t = \begin{bmatrix} u_t \\ v_t \end{bmatrix}, \qquad \text{where } u_t, v_t \in \{1, 2, ..., K\}. \tag{1}$$

Consider the following Markov chain.



Model parameters are as follows:

- initial distribution $\boldsymbol{\pi}_x \in \mathbb{R}^{K \times K}$ that parametrizes $Pr(\mathbf{X}_1)$:

$$Pr(\mathbf{X}_1 = \begin{bmatrix} i \\ j \end{bmatrix}) = \boldsymbol{\pi}_x(i, j). \tag{2}$$

- transition probability matrix $\mathbf{A}_x \in \mathbb{R}^{K\times K\times K\times K}$ that parametrizes $Pr(\mathbf{X}_{t+1}|\mathbf{X}_t)$:

$$Pr(\mathbf{X}_{t+1} = \begin{bmatrix} i_{t+1} \\ j_{t+1} \end{bmatrix} \mid \mathbf{X}_t = \begin{bmatrix} i_t \\ j_t \end{bmatrix}) = \mathbf{A}_x(i_t, j_t, i_{t+1}, j_{t+1}). \tag{3}$$

The joint probability can be factorized as:

$$Pr\big(\mathbf{X}_1, ..., \mathbf{X}_T\big) = Pr\big(\mathbf{X}_1\big) \prod_{t=1}^{T-1} Pr\big(\mathbf{X}_{t+1}|\mathbf{X}_t\big).$$

In this task, we refer to this model as *"2-D first-order Markov chain"*.

a) Does the sequence $[u_1, ..., u_T]$ (where $u_t \in \{1, 2, ..., K\}$ is defined in Eq. 1) have the first-order Markov property? Why or why not?

b) Let $[Y_1, ..., Y_T]$ be a 1-D first-order Markov chain with the following initial and transition probabilities ($Y_1, ..., Y_T$ are binary-valued).

$$\boldsymbol{\pi}_y = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}, \qquad \mathbf{A}_y = \begin{bmatrix} 0.2 & 0.8 \\ 0.5 & 0.5 \end{bmatrix}.$$

- Briefly explain why the sequence $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \begin{bmatrix} Y_2 \\ Y_3 \end{bmatrix}, ..., \begin{bmatrix} Y_{T-1} \\ Y_T \end{bmatrix}$ is a 2-D first-order Markov chain.

- Compute initial and transition probabilities, $\boldsymbol{\pi}_x$ and $\mathbf{A}_x$ (defined in Eqs. 2 and 3) for the sequence $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \begin{bmatrix} Y_2 \\ Y_3 \end{bmatrix}, ..., \begin{bmatrix} Y_{T-1} \\ Y_T \end{bmatrix}$.