Pattern Recognition Week 7 Assignment Report

Name: Chen Peng
StuID: 6930-30-2948
Lab: Yoshikawa & Ma lab

1. Discuss the similarity and difference between Kalman filter and hidden Markov Model (foward algorithm)

Difference

(1) The biggest difference is that an HMM has a finite-number of state while the states of a kalman filter are continuous variables.

(2) Different evolution method on the Markovian chain, an HMM transit its states by probability transition matrix whereas in a kalman filter, the state evolves based on state transition matrix as well as a Gaussian noise.

(3) In HMMs, hidden states are able to generate observable signals while in kalman filters, states are unobservable due to Gaussian noise.

Similarity

(1) Both of them are first order Markovian processes. The k-th state depends on the (k-1)-th state only.

(2) The observations in both follow Gaussian distribution.

2. Explain how to estimate the coefficients a1 and a2 of the auto-regressive (AR) model, $y_t = a_1 y_{t-1} + a_2 y_{t-2} + \epsilon_t$, from observed sequence data $y_1, y_2, \ldots, y_T \in R$.

Maximum likelihood estimation

(1) We assume that ϵ_t is independent and following a $N(0, \sigma^2)$ distribution. Given $y_1, y_2, \ldots, y_T$, we want to estimate the parameter set (a_1, a_2, σ^2).

(2) First, we need to find the likelihood function for this AR(2) model:

$$p(y_1, y_2, y_3, \ldots, y_n | a_1, a_2, \sigma^2) = p(y_1, y_2, \ldots, y_T | a_1, a_2, \sigma^2) \prod_{t=T+1}^{n} p(y_t | y_{t-1}, y_{t-2}, a_1, a_2, \sigma^2)$$

here the first T observations are already known. Therefore, the equation could be rewritten as:

$$p(y_1, y_2, y_3, \ldots, y_n | a_1, a_2, \sigma^2) = \prod_{t=T+1}^{n} p(y_t | y_{t-1}, y_{t-2}, a_1, a_2, \sigma^2)$$

At the meantime, y also follows the normal distribution, so

$$p(y_t | y_{t-1}, y_{t-2}, a_1, a_2, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} exp(-\frac{(y_t - a_1 y_{t-1} - a_2 y_{t-2})^2}{2\sigma^2})$$

Thus, the log-likelihood function can be:

$$L = \sum_{T+1}^{n} log(\frac{1}{\sqrt{2\pi\sigma}} exp(-\frac{(y_t - a_1 y_{t-1} - a_2 y_{t-2})^2}{2\sigma_2})) = -\frac{n-T}{2} log(2\pi) - \frac{n-T}{2} log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=T+1}^{n} (y_t - a_1 y_{t-1} - a_2 y_{t-2})^2$$

Calculate the derivatives and let them equal to 0, we can get the result of estimation:

$$a_1 = \frac{\sum_{t=T+1}^{n} y_t y_{t-1} - a_2 \sum_{t=T+1}^{n} y_{t-1} y_{t-2}}{\sum_{t=T+1}^{n} y_{t-1}^2}$$

$$a_2 = \frac{\sum_{t=T+1}^{n} y_t y_{t-2} - a_1 \sum_{t=T+1}^{n} y_{t-2} y_{t-1}}{\sum_{t=T+1}^{n} y_{t-2}^2}$$

$$\sigma^2 = \frac{1}{n-T} \sum_{t=T+1}^{n} (y_t - a_1 y_{t-1} - a_2 y_{t-2})^2$$