Hidden Markov Model

Depu Meng

Oct. 2018

1 Basic Concepts & Examples

1.1 Basic Concepts

Definition 1.1 Assume $X = \{X_n; n \geq 1\}$ is a Markov chain over finite state space $\varPhi = \{1, 2, ..., K\}$, if states of X are unobservable, $Y = \{Y_n; n \geq 1\}$ is an observable random variable series over finite set $V = \{v_1, v_2, ..., v_L\}$ that is correlated with X, then $(X, Y) = \{(X_n, Y_n); n \geq 1\}$ is a *Hidden Markov Chain.*

Denote $\pi = \{\pi_1, \pi_2, ..., \pi_k\}$ is the initial distribution of X, $A = [a_{ij}]$ is the transition matrix of X

$$a_{ij} = P\{X_{n+1} = j | X_n = i\}, i, j \in \varPhi \tag{1}$$

is the one-step transition probality of X. Denote

$$b_{ij} = P\{Y_n = v_j | X_n = i\}, i \in \varPhi, v_j \in V \tag{2}$$

represents the probability that Y equals to v_j given X is at state i at time n. If X is homogeneous, then Y is also irrelavent with time n. Denote $B = [b_{ij}]$ as the observation probability matrix. Due to the unobservability of X, π, A, B can not be directly measured. Generally, we call parameter set $\lambda = \{\pi, A, B\}$ the math model of Hidden Markov chain (X, Y), as well as *Hidden Markov Model* (HMM).

Property 1.1 Assume N is the length of time on observation, denote $\mathbf{X} = \{X_1, X_2, ..., X_N\}$, $\mathbf{Y} = \{Y_1, Y_2, ..., Y_N\}$ as the sample series of Markov chain X and observation series Y in the time period $1 \sim N$ respectively, then the joint distribution of $\mathbf{X}$ and $\mathbf{Y}$ satisfy the following *Hidden Markov Condition.*

$$P\{\mathbf{X} = x, \mathbf{Y} = y\} = \pi_{i_1} b_{i_1 j_1} ... b_{i_{N-1} j_{N-1}} b_{i_{N-1} j_N} b_{i_N j_N}, \tag{3}$$

where $x = \{i_1, ..., i_N\}$, $y = \{v_{j_1}, ..., v_{j_N}\}$.

1.2 Examples

Coin flipping problem Stochasticly choose a coin from two coins indexed $1 and 2$, and then toss it, observe the result and repeat this process. Denote X_n as the coin chosen in t^{th} time, then $X = \{X_n; n \geq 1\}$ is a Markov chain over state space $\varPhi = \{1, 2\}$. Denote Y_n as the result of n^{th} experiment, then $Y = \{Y_n; n \geq 1\}$

is observation series that takes values from $V = \{H, T\}$, where H means the front side and T means the back side. State transition matrix and observation probability matrix are

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \tag{4}$$

where summation of each row equals to 1. For initial distribution $\pi = \{\pi_1, \pi_2\}$, we have $\pi_1 + \pi_2 = 1$. Then $\lambda = \{\pi, A, B\}$ is the corresponding HMM of Hidden Markov chain (X, Y), which contains 5 unknown parameters.

2 Basic Problems & Solutions

2.1 Basic Problems

Problem 1 For a specific observation sample series $\mathbf{Y} = \{Y_1, ..., Y_n\}$, known it is generated by one of the given HMMs, determine which model that generates the sample series is called *Pattern Recognition* problem, as well as *classification* problem.

Problem 2 From a series of observation samples $\mathbf{Y} = \{Y_1, ..., Y_n\}$ and known HMM $\lambda = \{\pi, A, B\}$, give the best estimate of hidden states is called *State Estimate* problem, as well as *Decoding* problem.

Problem 3 From a series of observation samples $\mathbf{Y} = \{Y_1, ..., Y_n\}$, give the best estimate of parameter set $\lambda = \{\pi, A, B\}$, we call it *Model Learning* problem.

2.2 Solutions

Solution to Problem 1 To solve Problem 1 is mainly to solve $P\{\mathbf{Y}|\lambda\}$, for different λ, we can compare with their corresponding $P\{\mathbf{Y}|\lambda\}$, and choose the largest one. So how to compute $P\{\mathbf{Y}|\lambda\}$?

$$P\{\mathbf{Y} = y|\lambda\} = \sum_x P\{\mathbf{Y} = y, \mathbf{X} = x|\lambda\} \tag{5}$$

$$= \sum_x \pi_{i_1} b_{i_1 j_1} ... b_{i_{N-1} j_{N-1}} b_{i_{N-1} j_N} b_{i_N j_N} \tag{6}$$

In order to make the writing more convenient, we introduce notation $\alpha_n(i)$ and $\beta_n(i)$ first.

$$\alpha_n(i) = P\{Y_1 = v_{j_1}, ..., Y_n = v_{j_n}, X_n = i|\lambda\} \tag{7}$$

Recurrence relation of $\alpha_n(i)$.

$$\alpha_n(i) = \sum_{j=1}^{K} \alpha_n(j) a_{ji} b_{i j_{n+1}}, n = 1, ..., N-1 \tag{8}$$

proof

$$\alpha_{n+1}(i) = P\{Y_1 = v_{j_1}, Y_n = v_{j_n}, Y_{n+1} = v_{j_{n+1}}, X_{n+1} = i|\lambda\} \tag{9}$$

$$= \sum_{k=1}^{K} P\{Y_1 = v_{j_1}, ..., Y_n = v_{j_n}, Y_{n+1} = v_{j_{n+1}}, X_n = k, X_{n+1} = i|\lambda\} \tag{10}$$

$$= \sum_{k=1}^{K} P_A P_B P_C \tag{11}$$

where

$$P_A = P\{Y_1 = v_{j_1}, ..., Y_n = v_{j_n}, X_n = k|\lambda\} \tag{12}$$

$$P_B = P\{X_{n+1} = i|Y_1 = v_{j_1}, ..., Y_n = v_{j_n}, X_n = k, \lambda\} \tag{13}$$

$$P_C = P\{Y_{n+1} = v_{j_{n+1}}|Y_1 = v_{j_1}, ..., Y_n = v_{j_n}, X_n = k, X_{n+1} = i, \lambda\} \tag{14}$$

Notice that for P_A, we have

$$P_A = P\{Y_1 = v_{j_1}, ..., Y_n = v_{j_n}, X_n = k|\lambda\} = \alpha_n(k) \tag{15}$$

$$\tag{16}$$

For P_B, we have

$$P_B = P\{X_{n+1} = i|Y_1 = v_{j_1}, ..., Y_n = v_{j_n}, X_n = k, \lambda\} \tag{17}$$

$$= P\{X_{n+1} = i|X_n = k, \lambda\} \tag{18}$$

$$= a_{ki} \tag{19}$$

For P_C, we have

$$P_C = P\{Y_{n+1} = v_{j_{n+1}}|Y_1 = v_{j_1}, ..., Y_n = v_{j_n}, X_n = k, X_{n+1} = i, \lambda\} \tag{20}$$

$$= P\{Y_{n+1} = v_{j_{n+1}}|X_{n+1} = i, \lambda\} \tag{21}$$

$$= b_{ij_{n+1}} \tag{22}$$

Combine the three equations together, we have

$$\alpha_{n+1}(i) = \sum_{k=1}^{K} P_A P_B P_C \tag{23}$$

$$= \sum_{k=1}^{K} \alpha_n(k) a_{ki} b_{ij_{n+1}} \tag{24}$$

□