MATH 285: Stochastic Processes

math-old.ucsd.edu/~ynemish/teaching/285

Today: Hidden Markov chains

- Homework 4 is due on Friday, February 11, 11:59 PM

Example: Occasionally Dishonest Casino

Casino has two dice: fair (F) and loaded (L).

- F: $P(i) = \frac{1}{6}$
- L: $P(1) = 0.5$, $P(i) = 0.1$ for $i \geq 2$

Casino switches the die:

- $F \to L$ with probability 0.05
- $L \to F$ with probability 0.95

As a player you don't know which die is in use, you only observe the number that is rolled.

Suppose you play the game (roll the die) 6 times and observe 1, 1, 1, 1, 1, 1.

Q: What is the most likely sequence of dice used by casino?

Hidden Markov Model

Def 16.2 A Hidden Markov Model (HMM) is a pair of stochastic processes $(X_n, Y_n)_{n\geq 0}$ where (Y_n) is a Markov chain with state space S, and $(X_n)_{n\geq 0}$ has a possibly different state space R, and the vector valued process $Z_n = (X_n, Y_n)$ is a Markov chain. For $y \in S$ and $x \in R$ the conditional probabilities

$$e_y(x) = \mathbb{P}[X_n = x \mid Y_n = y]$$

are called the emission probabilities. Let $p: S \times S \to [0,1]$ be the transition kernel for (Y_n). It is taken as an assumption that the transition kernel for (Z_n) is

$$\mathbb{P}\left[Z_{n+1} = (x', y') \mid Z_n = (x, y)\right] = p(y, y')\, e_{y'}(x')$$

Hidden Markov Model

Remarks

(1) In general (X_n) is not a Markov chain

(2) Transition kernel for (Z_n) does not depend on x; this is not true in general for Markov chains on $S \times R$

$$\mathbb{P}[X_0 = x_0, Y_0 = y_0, X_1 = x_1, \dots, Y_{n-1} = y_{n-1}, X_n = x_n, Y_n = y_n]$$

$$= \mathbb{P}[X_0 = x_0, Y_0 = y_0]\, p(y_0, y_1) e_{y_1}(x_1)\, p(y_1, y_2) e_{y_2}(x_2) \cdots p(y_{n-1}, y_n) e_{y_n}(x_n)$$

$$= \mathbb{P}[Y_0 = y_0]\, e_{y_0}(x_0)\, p(y_0, y_1) e_{y_1}(x_1)\, p(y_1, y_2) e_{y_2}(x_2) \cdots p(y_{n-1}, y_n) e_{y_n}(x_n)$$

$$= \mathbb{P}[X_0 = x_0, \dots, X_n = x_n \mid Y_0 = y_0, \dots, Y_n = y_n]\, \mathbb{P}[Y_0 = y_0, \dots, Y_n = y_n]$$

$$= \mathbb{P}[X_0 = x_0, \dots, X_n = x_n \mid Y_0 = y_0, \dots, Y_n = y_n]\, \mathbb{P}[Y_0 = y_0]\, p(y_0, y_1) \cdots p(y_{n-1}, y_n)$$

$$\Rightarrow \mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n \mid Y_0 = y_0, \dots, Y_n = y_n] = e_{y_0}(x_0) \cdots e_{y_n}(x_n)$$

Example: Occasionally Dishonest Casino (2)

Construct a HMM that models ODC

- $S = \{F, L\}$, (Y_n) MC on S with transition probabilities

$$p(F, L) = 0.05 \qquad p(L, F) = 0.95$$

- $R = \{1, 2, 3, 4, 5, 6\}$, $X_n \in R$

Emission probabilities: $e_F(i) = \frac{1}{6}$ for all $i \in R$

$$e_L(i) = \begin{cases} 0.5, & i = 1 \\ 0.1, & i \in \{2, 3, 4, 5, 6\} \end{cases}$$

- $Z_n = (X_n, Y_n)$

$$\mathbb{P}[Z_{n+1} = (j, \beta) \mid Z_n = (i, \alpha)] = p(\alpha, \beta)\, e_\beta(j)$$

The forward algorithm

Let (X_n, Y_n) be a HMM. Denote

- $x = (x_0, x_1, \dots, x_N)$ the observed sequence
- $y = (y_0, y_1, \dots, y_N)$ the state sequence
- $\mathbb{P}[x] = \mathbb{P}[X_0 = x_0, \dots, X_N = x_N]$
- $\mathbb{P}[x, y] = \mathbb{P}[X_0 = x_0, \dots, X_N = x_N, Y_0 = y_0, \dots, Y_N = y_N]$

Q: What is the probability of $(y_0, y_1, \dots, y_N)$ given that we observe $(x_0, x_1, \dots, x_N)$?

Using the above notation, we have to compute

$$\mathbb{P}[y \mid x] = \frac{\mathbb{P}[x, y]}{\mathbb{P}[x]} \qquad \mathbb{P}[x] - ?$$

We know that $\mathbb{P}[x, y] = \mathbb{P}[y_0] e_{y_0}(x_0) \, p(y_0, y_1) e_{y_1}(x_1) \cdots p(y_{N-1}, y_N) e_{y_N}(x_N)$

The forward algorithm

Direct way of computing $\mathbb{P}[x]$

$$\mathbb{P}[x] = \sum_{y \in S^{|N|+1}} \mathbb{P}[x, y]$$

Problem: computationally infeasible $\sim N|S|^N$ computations

grows exponentially fast with N

The forward algorithm allows to compute $\mathbb{P}[x]$ in polynomial time.

Fix observed sequence $x = (x_0, x_1, \dots, x_N)$. For any $y \in S$ and $n \in \{0, 1, \dots, N\}$ define the probability

$$\alpha_n(y) = \mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, Y_n = y]$$

that first n observations occured and the hidden state is y.

The forward algorithm

Then

$$\alpha_{n+1}(y') = \mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, X_{n+1} = x_{n+1}, Y_{n+1} = y']$$

$$= \sum_{y \in S} \mathbb{P}[X_0 = x_0, \dots, X_n = x_n, Y_n = y, X_{n+1} = x_{n+1}, Y_{n+1} = y']$$

Now condition on $X_0, X_1, \dots, X_n, Y_n$

$$\mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, Y_n = y, X_{n+1} = x_{n+1}, Y_{n+1} = y']$$

$$= \mathbb{P}[X_{n+1} = x_{n+1}, Y_{n+1} = y' \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, Y_n = y]$$

$$\times \mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, Y_n = y]$$

$$= \mathbb{P}[X_{n+1} = x_{n+1}, Y_{n+1} = y' \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, Y_n = y]\, \alpha_n(y)$$

Lemma 16.5 Let $Z_n = (X_n, Y_n)$ be Markov chain. Then

$$\mathbb{P}[Z_{n+1} = (x_{n+1}, y') \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n, Y_n = y]$$

$$= \mathbb{P}[Z_{n+1} = (x_{n+1}, y') \mid Z_n = (x_n, y)]$$

The forward algorithm

Therefore,

$$\alpha_{n+1}(y') = \sum_{y\in S} \mathbb{P}[Z_{n+1} = (x_{n+1}, y') \mid Z_n = (x_n, y)]\, \alpha_n(y)$$

$$= \sum_{y\in S} p(y,y')\, e_{y'}(x_{n+1})\, \alpha_n(y)$$

$$= e_{y'}(x_{n+1}) \sum_{y\in S} p(y,y')\, \alpha_n(y) \qquad (*)$$

and we can compute

$$\mathbb{P}[x] = \sum_{y\in S} \mathbb{P}[X_0 = x_0, X_1 = x_1, \dots, X_N = x_n, Y_N = y] = \sum_{y\in S} \alpha_N(y) \qquad (**)$$

Complexity of the forward algorithm:

- $\alpha_0(y) = \mathbb{P}[X_0 = x_0, Y_0 = y] = \mathbb{P}[Y_0 = y]\, e_y(x_0)$
- By $(*)$ we need $\sim 2|S|$ operations to compute $\alpha_n(y)$
- By $(**)$ we have to compute $\alpha_n(y)$ for all n, y $\sim N|S|^2$

Proof of Lemma 16.5

$$\mathbb{P}[Z_{n+1}=(x_{n+1},y') \mid X_0=x_0, X_1=x_1,\dots, X_n=x_n, Y_n=y]$$

$$= \frac{\mathbb{P}[X_0=x_0,\dots, X_{n+1}=x_{n+1}, Y_n=y, Y_{n+1}=y']}{\mathbb{P}[X_0=x_0,\dots, X_n=x_n, Y_n=y]}$$

$$\mathbb{P}[X_0=x_0, X_1=x_1,\dots, X_n=x_n, Y_n=y, X_{n+1}=x_{n+1}, Y_{n+1}=y']$$

$$= \sum_{y_0,\dots,y_{n-1}} \mathbb{P}[Z_0=(x_0,y_0),\dots, Z_n=(x_n,y), Z_{n+1}=(x_{n+1},y')]$$

$$= \sum_{y_0,\dots,y_{n-1}} \mathbb{P}[Z_{n+1}=(x_{n+1},y') \mid Z_n=(x_n,y)]\,\mathbb{P}[Z_0=(x_0,y_0),\dots, Z_n=(x_n,y)]$$

$$= \mathbb{P}[Z_{n+1}=(x_{n+1},y') \mid Z_n=(x_n,y)]\,\mathbb{P}[X_0=x_0, X_1=x_1,\dots, X_n=x_n, Y_n=y]$$

$\blacksquare$