CSE 250A HW 7

Jiping Lin A15058075

$$\begin{tikzcd} O_1 & O_2 & & O_t & O_{t+1} & & O_T \\ S_1 \arrow[u] \arrow[r] & S_2 \arrow[u] \arrow[r] & \cdots \arrow[r] & S_t \arrow[u] \arrow[r] & S_{t+1} \arrow[u] \arrow[r] & \cdots \arrow[r] & S_T \arrow[u] \end{tikzcd}$$

7.1 A HOUSE DIVIDED AGAINST ITSELF → CANNOT STAND

7.2 (a) $P(S_{t+1}=j \mid S_t=i, O_1, \cdots, O_T)$

$$= \frac{P(S_{t+1}=j, S_t=i, O_1, \cdots, O_t, O_{t+1}, \cdots, O_T)}{P(S_t=i, O_1, \cdots, O_t, O_{t+1}, \cdots, O_T)}$$

$$= \frac{P(S_t=i, O_1, \cdots, O_t) P(S_{t+1}=j \mid S_t=i, O_1, \cdots, O_t) P(O_{t+1} \mid S_{t+1}=j, S_t=i, O_1, \cdots, O_t) P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j, S_t=i, O_1, \cdots, O_{t+1})}{P(S_t=i, O_1, \cdots, O_t) P(O_{t+1}, \cdots, O_T \mid S_t=i, O_1, \cdots, O_t)}$$

[Product Rule split in 4 pts]

$$= \frac{\alpha_{it} \underbrace{P(S_{t+1}=j \mid S_t=i)}_{\text{dsep.(1)}} \underbrace{P(O_{t+1} \mid S_{t+1}=j)}_{\text{d.sep(1)(2)}} \underbrace{P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j)}_{\text{d-sep(1)}}}{\alpha_{it} \underbrace{P(O_{t+1}, \cdots, O_T \mid S_t=i)}_{\text{d.sep(1) \& (2) } (O_t \to S_{t+1})}}$$

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{it} \beta_{it}}$$

$$= \frac{a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{\beta_{it}}$$

(b) $P(S_t=i \mid S_{t+1}=j, O_1, \cdots, O_T)$

$$= \frac{P(S_t=i, S_{t+1}=j, O_1, \cdots, O_T)}{P(S_{t+1}=j, O_1, \cdots, O_{t+1}, O_{t+2}, \cdots, O_T)}$$

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{P(S_{t+1}=j, O_1, \cdots, O_{t+1}) P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j, O_1, \cdots, O_{t+1})}$$ ← Same as num. in (a)

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{j(t+1)} P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j)}$$ d.sep(1), (2)

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{j(t+1)} \beta_{j(t+1)}}$$

$$= \frac{\alpha_{it} a_{ij} b_j(O_{t+1})}{\alpha_{j(t+1)}}$$

(c) $P(S_{t-1}=i, S_t=k, S_{t+1}=j \mid O_1, \cdots, O_T)$

$$= \frac{P(S_{t-1}=i, S_t=k, S_{t+1}=j, O_1, \cdots, O_{t-1}, O_t, O_{t+1}, \cdots, O_T)}{P(O_1, \cdots, O_T)}$$

$$= \frac{P(S_{t-1}=i, O_1, \cdots, O_{t-1}) P(S_t=k, S_{t+1}=j, O_t, \cdots, O_T \mid S_{t-1}=i, O_1, \cdots, O_{t-1})}{P(O_1, \cdots, O_T)}$$

NEXT PAGE

$$= \frac{\alpha_{i(t-1)} P(S_t=k \mid S_{t-1}=i, O_1, \cdots O_{t-1}) P(S_{t+1}=j, O_t, \cdots, O_T \mid S_{t-1}=i, S_t=k, O_1, \cdots, O_{t-1})}{P(O_1, \cdots, O_T)}$$

$$\overset{\text{d.sep(1)(2)}}{=} \frac{\alpha_{i(t-1)} P(S_t=k \mid S_{t-1}=i) P(O_t \mid S_{t-1}=i, S_t=k, O_1, \cdots, O_{t-1}) P(S_{t+1}=j, O_{t+1}, \cdots, O_T \mid S_{t-1}=i, S_t=k, O_1, \cdots, O_t)}{P(O_1, \cdots, O_T)}$$

$$\overset{\text{d.sep(1)}}{=} \frac{\alpha_{i(t-1)} a_{ik} P(O_t \mid S_t=k) P(S_{t+1}=j \mid S_{t-1}=i, S_t=k, O_1, \cdots, O_t) P(O_{t+1}, \cdots, O_T \mid S_{t+1}=j, S_{t-1}=i, S_t=k, O_1, \cdots, O_t)}{P(O_1, \cdots, O_T)}$$

$$\overset{\text{d.sep(1)(2)}}{=} \frac{\alpha_{i(t-1)} a_{ik} b_k(O_t) P(S_{t+1}=j \mid S_t=k) P(O_{t+1} \mid S_{t+1}=j, S_{t-1}=i, S_t=k, O_1, \cdots, O_t) P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j, S_{t-1}=i, S_t=k, O_1, \cdots, O_{t+1})}{P(O_1, \cdots, O_T)}$$

$$= \frac{\alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} P(O_{t+1} \mid S_{t+1}=j) P(O_{t+2}, \cdots, O_T \mid S_{t+1}=j)}{P(O_1, \cdots, O_T)}$$

$$= \frac{\alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} b_j(O_{t+1}) \beta_{j(t+1)} \ (*)}{\sum_x P(O_1, \cdots, O_T, S_t=x)}$$

$$= (*) / \sum_x P(S_t=x, O_1, \cdots, O_t) P(O_{t+1}, \cdots, O_T \mid S_t=x, O_1, \cdots, O_t)$$

$$= (*) / \sum_x \alpha_{xt} P(O_{t+1}, \cdots, O_T \mid S_t=x)$$

$$= \frac{\alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} b_j(O_{t+1}) \beta_{j(t+1)}}{\sum_x \alpha_{xt} \beta_{xt}}$$

(d) $P(S_{t+1}=j \mid S_{t-1}=i, O_1, \cdots, O_T)$

$$= \frac{P(S_{t+1}=j, S_{t-1}=i, O_1, \cdots, O_T)}{P(S_{t-1}=i, O_1, \cdots, O_T)}$$

$$= \frac{\sum_k P(S_{t+1}=j, S_t=k, S_{t-1}=i, O_1, \cdots, O_T)}{P(S_{t-1}=i, O_1, \cdots, O_{t-1}) P(O_t, \cdots, O_T \mid S_{t-1}=i, O_1, \cdots, O_{t-1})} \quad \text{(c) num.}$$

$$= \frac{\sum_k \alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{i(t-1)} P(O_t, \cdots, O_T \mid S_{t-1}=i)} \quad \text{d.sep(1)}$$

$$= \frac{\sum_k \alpha_{i(t-1)} a_{ik} b_k(O_t) a_{kj} b_j(O_{t+1}) \beta_{j(t+1)}}{\alpha_{i(t-1)} \beta_{i(t-1)}}$$

7.3 FTFFTFTTFFTT

7.4 (a) $q_{jt} = P(S_t=j \mid O_1,\cdots,O_{t-1},O_t)$

$$= \frac{P(S_t=j \mid O_1,\cdots,O_{t-1})\,P(O_t \mid S_t=j,O_1,\cdots,O_{t-1})}{P(O_t \mid O_1,\cdots,O_{t-1})} \quad \ulcorner\text{Baye's Rule}\lrcorner$$

$$= \frac{P(S_t=j \mid O_1,\cdots,O_{t-1})\,P(O_t \mid S_t=j)}{P(O_t \mid O_1,\cdots,O_{t-1})} \quad \leftarrow \text{d.sep(1)}$$

$\leftarrow$ denom.

$= \sum_i P(S_{t-1}=i, S_t=j \mid O_1,\cdots,O_{t-1})\, b_j(O_t)$ / denom.

$= \sum_i P(S_{t-1}=i \mid O_1,\cdots,O_{t-1})\, P(S_t=j \mid S_{t-1}=i, O_1,\cdots,O_{t-1})\, b_j(O_t)$ / denom.

$= b_j(O_t) \sum_i q_{i(t-1)}\, P(S_t=j \mid S_{t-1}=i)$ / denom $\leftarrow$ d.sep(1)(2)

$= b_j(O_t) \sum_i q_{i(t-1)}\, a_{ij}$

denom $= P(O_t \mid O_1,\cdots,O_{t-1}) = \sum_j P(S_t=j, O_t \mid O_1,\cdots,O_{t-1})$

$= \sum_j P(S_t=j \mid O_1,\cdots,O_{t-1})\, P(O_t \mid S_t=j, O_1,\cdots,O_{t-1})$

$= \sum_j b_j(O_t) \sum_i q_{i(t-1)}\, a_{ij}$ (from num.)

So $$q_{jt} = \frac{b_j(O_t)\sum_i q_{i(t-1)} a_{ij}}{\sum_{ij} b_j(O_t)\, q_{i(t-1)} a_{ij}}$$

(b) $$P(X_t \mid Y_1,\cdots,Y_t) = \frac{P(X_t \mid Y_1,\cdots,Y_{t-1})\, P(Y_t \mid X_t, Y_1,\cdots,Y_{t-1})}{P(Y_t \mid Y_1,\cdots,Y_{t-1})}$$

Num $= P(X_t \mid Y_1,\cdots,Y_{t-1})\, P(Y_t \mid X_t)$ $\leftarrow$ d.sep(1)

$= P(Y_t \mid X_t) \int P(X_t, X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}$ (Marg. on X_{t-1})

$= P(Y_t \mid X_t) \int P(X_t \mid X_{t-1}, Y_1,\cdots,Y_{t-1})\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}$

$= P(Y_t \mid X_t) \int \underset{\text{d.sep(1)}}{P(X_t \mid X_{t-1})}\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}$

Denom $= \int P(X_t, Y_t \mid Y_1,\cdots,Y_{t-1})\, dX_t$

$= \int \underbrace{P(X_t \mid Y_1,\cdots,Y_{t-1})\, P(Y_t \mid X_t, Y_1,\cdots,Y_{t-1})}_{\text{num.}}\, dX_t$

$= \int P(Y_t \mid X_t)\left(\int P(X_t \mid X_{t-1})\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}\right) dX_t$.

it is easier than other distributions where we have to compute each cond. prob. and find integral.

So $$P(X_t \mid Y_1,\cdots,Y_t) = \text{Num./denom.} = \frac{P(Y_t \mid X_t)\int P(X_t \mid X_{t-1})\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}}{\int P(Y_t \mid X_t)\left(\int P(X_t \mid X_{t-1})\, P(X_{t-1} \mid Y_1,\cdots,Y_{t-1})\, dX_{t-1}\right) dX_t}$$

If X, Y are Gaussian, then Y|X is also gaussian, then since the integral of Gaussian is easy to compute $\lrcorner$

7.5 (a) $P(Y_1=j, O_1=o_1)$

$= \sum_i P(Y_1=j, X_1=i, O_1=o_1)$

$= \sum_i P(Y_1=j) P(X_1=i \mid Y_1=j) P(O_1=o_1 \mid X_1=i, Y_1=j)$

$= \sum_i P(X_1=i, Y_1=j) b_{ij}(o_1)$

$= \sum_i P(X_1=i) \underbrace{P(Y_1=j)}_{\text{d-sep(3)}} b_{ij}(o_1)$

$= \sum_i P(X_1=i) \pi_j b_{ij}(o_1)$

(b) Assume $t>1$,

$\alpha_{jt} = P(o_1, \cdots, o_t, Y_t=j)$

$= \sum_k \sum_i P(Y_{t-1}=k, Y_t=j, X_t=i, o_1, \cdots, o_t)$

$= \sum_k \sum_i P(o_1, \cdots, o_{t-1}, Y_{t-1}=k) P(X_t=i \mid Y_{t-1}=k, o_1, \cdots, o_{t-1})$
$P(Y_t=j \mid X_t=i, Y_{t-1}=k, o_1, \cdots, o_{t-1})$
$P(o_t \mid X_t=i, Y_{t-1}=k, Y_t=j, o_1, \cdots, o_{t-1})$

$= \sum_k \sum_i \alpha_{k(t-1)} \underbrace{P(X_t=i \mid Y_{t-1}=k)}_{\text{d-sep(2)}} \underbrace{P(Y_t=j)}_{\text{d-sep(3)}} \underbrace{P(o_t \mid X_t=i, Y_t=j)}_{\text{d-sep(1)}}$

$= \sum_k \sum_i \alpha_{k(t-1)} a_{ki} \pi_j b_{ij}(o_t)$
recursion

where $t=1$ case is part (a)

(c) $P(o_1, \cdots, o_T) = \sum_x P(o_1, \cdots, o_T, Y_T=j) = \sum_x \alpha_{xT}$

(d) $P(o_1, \cdots, o_T) = \sum_x \alpha_{xT} = \sum_x \sum_k \sum_i \alpha_{k(T-1)} a_{ki} \pi_x b_{ix}(o_T)$

α_{k1} takes $O(n_x)$ since sum is from $i=1$ to n_x

$\alpha_{kT} = O(\underbrace{n_y O(\alpha_{k(T-1)})}_{\text{sum from } k=1 \text{ to } n_y \text{, recurse from } T \text{ to } 1}) \Rightarrow \alpha_{kT}$ takes $O((T-1) n_x n_y)$

then $\sum_x \alpha_{xt}$ takes $n_y O(\alpha_{xt})$ so the final complexity is $O(n_x n_y^2 (T-1))$

HW7 Code

November 18, 2021

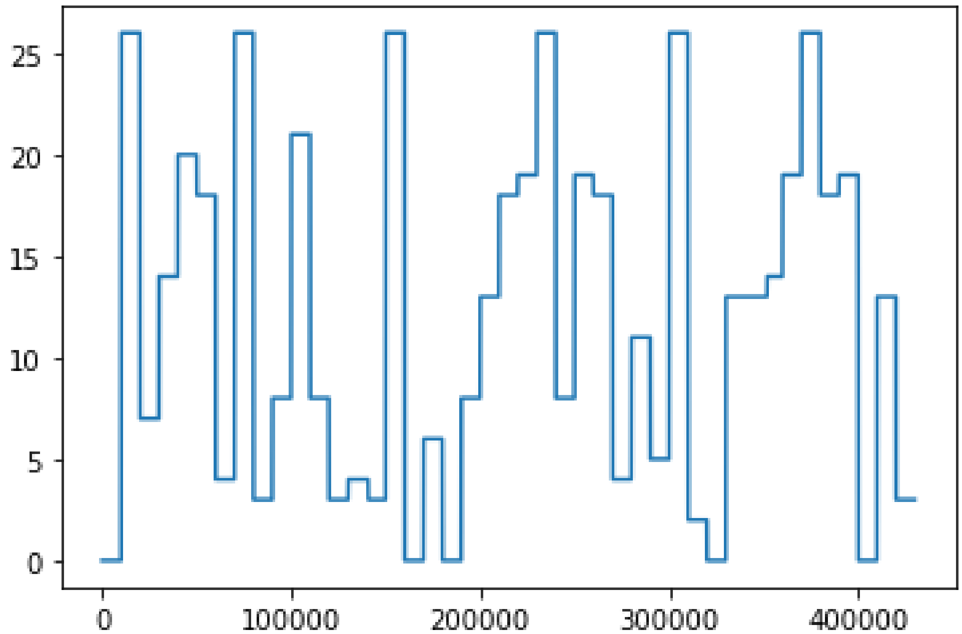
[1]:

```
import string

import numpy as np
import matplotlib.pyplot as plt
b = np.loadtxt('emissionMatrix.txt')
pi = np.loadtxt('initialStateDistribution.txt')
obs = np.loadtxt('observations.txt').astype(int)
a = np.loadtxt('transitionMatrix.txt')
log_a = np.log(a)
T = len(obs)
n = len(pi)
l = np.zeros((n, T))
phi = np.zeros((T, n))
for j in range(l.shape[1]):
    for i in range(l.shape[0]):
        if j == 0:
            l[i][j] = np.log(pi[i]) + np.log(b[i][obs[0]])
        if j > 0:
            l_it = l[:, j-1]
            a_ij = log_a[:, i]
            first = np.max(l_it + a_ij)
            phi[j][i] = np.argmax(l_it + a_ij)
            l[i][j] = first + np.log(b[i][obs[j]])
s = np.zeros(len(obs))
s[-1] = int(np.argmax(l[:, -1].flatten()))
for t in range(len(obs) - 2, -1, -1):
    s[t] = phi[t+1][int(s[t+1])]
x = range(0, len(obs))
plt.plot(x, s)
plt.show()

alphabet = list('abcdefghijklmnopqrstuvwxyz ')
result = []
for i in range(len(s)):
    if i == 0:
        result.append(alphabet[int(s[0])])
        continue
    if s[i] == s[i - 1]:
```

```
        if alphabet[int(s[i - 1])] == 'n' and alphabet[int(s[i - 2])] == 'a':
            result.append(alphabet[int(s[i])])
        else:
            continue
    else:
        result.append(alphabet[int(s[i])])
result = ''.join(result)
print(result)
```




```
a house divided against itself cannot standd
```

[ ]: