In [1]:

```
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:

```
Afile = 'transitionMatrix.txt'
Bfile = 'emissionMatrix.txt'
PIfile = 'initialStateDistribution.txt'
obsfile = 'observations.txt'
```

In [3]:

```
def loadFile(path):
    data = []
    with open(path) as file:
        for line in file:
            data.append([float(item) for item in line.strip().strip('\n').split()])
    return np.asarray(data)
```

In [4]:

```
A = loadFile(Afile)
B = loadFile(Bfile)
PI = loadFile(PIfile)
obs = np.apply_along_axis(int, 0, loadFile(obsfile))
```

In [6]:

```
obs
```

Out[6]:

```
array([0, 0, 0, ..., 0, 0, 0])
```

In [7]:

```
n = A.shape[0]
m = B.shape[1]
T = obs.shape[0]
```

In [8]:

```
print(n, m, T)
```

```
27 2 430000
```

In [9]:

```
lstar = np.empty((n, T))
```

In [10]:

```
hidden_state = np.empty_like(obs)
```

In [11]:

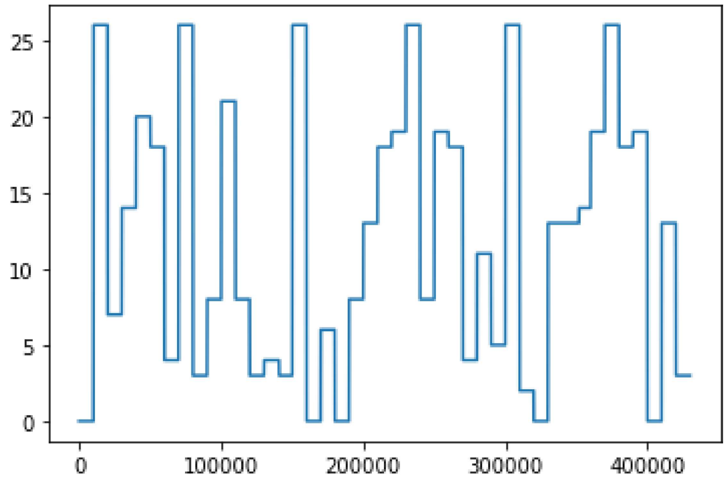
```
for t in range(T):
    for i in range(n):
        if t == 0:
            lstar[i][0] = np.log(PI[i]) + np.log(B[i][obs[0]])
        else:
            max_val = -np.inf
            for j in range(n):
                max_val = max(max_val, lstar[j][t - 1] + np.log(A[j][i]))
            lstar[i][t] = max_val + np.log(B[i][obs[t]])
```

In [70]:

```
for t in range(T - 1, 0, -1):
    current_max = -np.inf
    current_i = 0
    for i in range(n):
        if t == T - 1:
            if current_max < lstar[i][t]:
                current_max = lstar[i][t]
                hidden_state[t] = i
        else:
            temp = (lstar[i][t] + np.log(A[i][hidden_state[t + 1]]))
            if current_max < temp:
                current_max = temp
                hidden_state[t] = i
```

In [113]:

```
plt.plot(hidden_state)
plt.show()
```



a house divided against itself cannot stand

7.2

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

$\Rightarrow P(S_{t+1}=j \mid S_t=i, O_{t+1}, \ldots O_T)$ $\boxed{C.I.}$

$\Rightarrow \dfrac{P(O_{t+1} \ldots O_T \mid S_{t+1}=j, S_t=i) P(S_{t+1}=j \mid S_t=i)}{P(O_{t+1}, \ldots O_T \mid S_t)}$ $\boxed{Bayes}$

$\Rightarrow \dfrac{P(O_{t+1} \ldots O_T \mid S_{t+1}=j)\, a_{ij}}{\beta_{it}}$ $\boxed{C.I.}$

$\Rightarrow \dfrac{P(O_{t+2}, \ldots O_T \mid S_{t+1}=j) P(O_{t+1} \mid S_{t+1}, O_{t+2} \ldots O_T)\, a_{ij}}{\beta_{it}}$ $\boxed{P.R.}$

$\Rightarrow \dfrac{\beta_{jt+1}\, P(O_{t+1} \mid S_{t+1}=j)\, a_{ij}}{\beta_{it}}$ $\boxed{C.I.}$

$\Rightarrow \boxed{\dfrac{\beta_{jt+1}\, b_j(O_{t+1})\, a_{ij}}{\beta_{it}}}$

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

$$\Rightarrow \frac{P(S_{t+1}=j \mid S_t = i, O_1 \dots O_T)\, P(S_t = i \mid O_1 \dots O_T)}{P(S_{t+1} = j \mid O_1 \dots O_T)}$$ $\boxed{\text{Bayes}}$

$$\Rightarrow \frac{a_{ij}\, b_j(O_{t+1})\, \beta_{jt+1}}{\beta_{it}} \times \frac{P(S_t, O_1 \dots O_T)\, \cancel{P(O_1 \dots O_T)}}{\cancel{P(O_1 \dots O_T)}\, P(S_{t+1}, O_1 \dots O_T)}$$

$\boxed{\text{From (a) \& P.R.}}$

$$\Rightarrow \frac{a_{ij}\, b_j(O_{t+1})\, \beta_{jt+1}}{\beta_{it}} \times \frac{P(S_t, O_1 \dots O_t)\, P(O_{t+1} \dots O_T \mid S_t)}{P(S_{t+1}, O_1 \dots O_{t+1})\, P(O_{t+2} \dots O_T \mid S_{t+1})}$$

$\boxed{\text{P.R. \& C.I}}$

$$\Rightarrow a_{ij}\, b_j \frac{(O_{t+1})\, \cancel{\beta_{jt+1}}}{\cancel{\beta_{it}}} \frac{\alpha_{it}\, \cancel{\beta_{it}}}{\alpha_{jt+1}\, \cancel{\beta_{jt+1}}}$$

$$\Rightarrow \boxed{\frac{a_{ij}\, b_j(O_{t+1})\, \alpha_{it}}{\alpha_{jt+1}}}$$

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

$\Rightarrow P(S_{t+1} \mid O_1 \ldots O_T, S_{t-1}, S_t) P(S_t \mid O_1 \ldots O_T, S_{t-1})$

$\times P(S_{t-1} \mid O_1 \ldots O_T)$ $\boxed{P.R.}$

$\Rightarrow P(S_{t+1} \mid S_t, O_{t+1}, \ldots O_T) P(S_t \mid S_{t-1}, O_t \ldots O_T) \times$

$P(S_{t-1} \mid O_1 \ldots O_T)$ $\boxed{C.I.}$

$\ldots (i)$

- $P(S_{t+1} \mid S_t, O_{t+1} \ldots O_T) = \dfrac{P(O_{t+1} \ldots O_T \mid S_{t+1}, S_t) P(S_{t+1} \mid S_t)}{P(O_{t+1} \ldots O_T \mid S_t)}$

$\boxed{Bayes}$

$\Rightarrow = \dfrac{P(O_{t+1} \ldots O_T \mid S_{t+1}) a_{kj}}{\beta_{kt}}$ $\boxed{C.I.}$

$= \dfrac{P(O_{t+2} \ldots O_T \mid S_{t+1}, O_{t+1}) P(O_{t+1} \mid S_{t+1}) a_{kj}}{\beta_{kt}}$

$\boxed{P.R.}$

$= \dfrac{\beta_{jt+1} b_j(O_{t+1}) a_{kj}}{\beta_{kt}}$ $\ldots (ii)$ $\boxed{C.I}$

- Similarily;

$$P(S_t \mid S_{t-1}, O_t \ldots O_T) = \frac{\beta_{kt}\, b_k(O_t)\, a_{ik}}{\beta_{it-1}} \quad \ldots (iii)$$

Finally;

$$P(S_{t-1} \mid O_1 \ldots O_T) = \frac{P(S_{t-1}, O_1 \ldots O_T)}{P(O_1 \ldots O_T)} \quad \boxed{P.R.}$$

$$\cancel{= P(O_t, \ldots O_T \mid S_t, O_1 \ldots O_{t-1})\, P(S_t,}$$

$$= \frac{P(O_t \ldots O_T \mid S_{t-1}, O_1 \ldots O_{t-1})\, P(O_1 \ldots O_{t-1}, S_{t-1})}{\sum_K P(S_{t-1} = k, O_1 \ldots O_T)} \quad \boxed{\begin{matrix} P.R \\ \& \\ marg \end{matrix}}$$

$$= \frac{\alpha_{it-1}\, \beta_{it-1}}{\sum_k \alpha_{kt-1}\, \beta_{kt-1}} \quad \ldots (iv)$$

Put (ii), (iii) & (iv) in (i)

$$\Rightarrow \boxed{\frac{\alpha_{it-1}\, \beta_{jt+1}\, a_{kj}\, a_{ik}\, b_j(O_{t+1})\, b_k(O_t)}{\sum_k \alpha_{kt-1}\, \beta_{kt-1}}}$$

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

$\Rightarrow \sum_k P(S_{t+1}=j, S_t=k \mid S_{t-1}=i, O_1 \ldots O_T)$ $\boxed{\text{marg}}$

$\Rightarrow \sum_k P(S_{t+1} \mid S_t, S_{t-1}, O_1 \ldots O_T) P(S_t \mid S_{t-1}, O_1 \ldots O_T)$

$\Rightarrow \sum_k P(S_{t+1} \mid S_t \, O_{t+1} \ldots O_T) P(S_t \mid S_{t-1} \, O_t \ldots O_T)$

$\Rightarrow$ From (a)

$$\Rightarrow \boxed{\sum_K \frac{\beta_{j\,t+1} \, b_j(O_{t+1}) \, b_K(O_t) \, a_{Kj} \, a_{iK}}{\beta_{i\,t-1}}}$$

7.3

a) F
b) T
c) F
d) F
e) T
f) F
g) T
h) T
i) F
j) F
k) T
l) T

7.4

$q_{jt} = P(S_t = j \mid O_1, \ldots O_t)$

a)

$= \sum_i P(S_t = j, S_{t-1} = i \mid O_1 \ldots O_t)$ $\boxed{\text{marg}}$

$= \sum_i \dfrac{P(S_t = j, S_{t-1} = i, O_t \mid O_1 \ldots O_{t-1})}{P(O_t \mid O_1 \ldots O_{t-1})}$ $\boxed{\text{P.R.}}$

$$\Rightarrow \frac{\sum_i P(S_t=j, O_t | S_{t-1}=i) P(S_{t-1}=i | O_1 \ldots O_{t-1})}{\sum_{ij} P(S_t=j, S_{t-1}=i, O_t | O_1 \ldots O_{t-1})}$$ $\boxed{\text{P.R.}}$ $\boxed{\text{marg}}$

$$\Rightarrow \frac{\sum_i P(O_t | S_t=j) P(S_t=j | S_{t-1}=i)\, q_{it-1}}{\sum_{ij} P(S_t=j, O_t | S_{t-1}) P(S_{t-1} | O_1 \ldots O_{t-1})}$$ $\boxed{\text{P.R. \& C.I.}}$

$$\Rightarrow \frac{\sum_i b_j(O_t)\, a_{ij}\, q_{it-1}}{\sum_{ij} b_j(O_t)\, a_{ij}\, q_{it-1}}$$

$$\Rightarrow \boxed{\begin{gathered} \frac{1}{Z_t} b_j(O_t) \sum_i a_{ij}\, q_{it-1} \\ \text{where, } Z_t = \sum_{ij} b_j(O_t)\, a_{ij}\, q_{it-1} \end{gathered}}$$

b) $P(x_t | y_1, \ldots y_t)$

$$\Rightarrow \int dx_{t-1}\, P(x_t, x_{t-1} | y_1 \ldots y_t) \qquad \boxed{\text{marg}}$$

$$\Rightarrow \frac{\int dx_{t-1}\, P(x_t, x_{t-1}, y_t | y_1, \ldots y_{t-1})}{P(y_t | y_1 \ldots y_{t-1})} \qquad \boxed{\text{P.R.}}$$

$$\Rightarrow \frac{\int dx_{t-1}\, P(x_t, y_t | x_{t-1}) P(x_{t-1} | y_1 \ldots y_{t-1})}{\int dx_t \int dx_{t-1}\, P(x_t, x_{t-1}, y_t | y_1 \ldots y_{t-1})}$$

[P.R. & C.I. for numerator
marg. for denominator]

$$\Rightarrow \frac{\int dx_{t-1}\, P(y_t | x_t) P(x_t | x_{t-1})\, P(x_{t-1} | y_1 \ldots y_{t-1})}{\int dx_t \int dx_{t-1}\, P(x_t, y_t | x_{t-1}) P(x_{t-1} | y_1 \ldots y_{t-1})}$$

[P.R. for numerator
P.R. & C.I for denominator]

$$\Rightarrow \frac{P(y_t|x_t)\int dx_{t-1}\, P(x_t|x_{t-1})\, P(x_{t-1}|y_1 \ldots y_{t-1})}{\int dx_t\, P(y_t|x_t)\int dx_{t-1}\, P(x_t|x_{t-1})\, P(x_{t-1}|y_1 \ldots y_{t-1})}$$

[C.I. for denominator]

$$\Rightarrow \frac{1}{Z_t} P(y_t|x_t)\int dx_{t-1}\, P(x_t|x_{t-1})\, P(x_{t-1}|y_1 \ldots y_{t-1})$$

where;

$$Z_t = \int dx_t\, P(y_t|x_t)\int dx_{t-1}\, P(x_t|x_{t-1})\, P(x_{t-1}|y_1 \ldots y_{t-1})$$

Integrating over distributions other than gaussians might not be trivial and thus will be for all but gaussian random variables.

7.5

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

$\Rightarrow \sum_i P(X_1 = i, Y_1 = j, O_1 = o_1)$ $\boxed{\text{marg}}$

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

$\boxed{\text{P.R.}}$

$\Rightarrow \sum_i b_{ij}(o_1) P(X_1 = i) P(Y_1 = j)$ $\boxed{\text{C.I.}}$

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

b) $\alpha_{jt} = P(O_1, \ldots O_t, Y_t = j)$

$\Rightarrow \sum_k P(O_1, \ldots O_t, Y_{t-1} = k, Y_t = j)$ $\boxed{\text{marg}}$

$\Rightarrow \sum_k P(O_1, \ldots O_{t-1}, Y_{t-1}) P(O_t, Y_t = j | Y_{t-1}, O_1 \ldots O_{t-1})$

$\boxed{\text{P.R.}}$

$\Rightarrow \sum_k \alpha_{kt-1} P(O_t, Y_t = j | Y_{t-1} = k)$

$\boxed{\text{C.I.}}$

$$\Rightarrow \sum_k \alpha_{k\,t-1} \sum_i P(O_t, X_t = i, Y_t = j \mid Y_{t-1} = k)$$

marg

$$\Rightarrow \sum_k \alpha_{k\,t-1} \sum_i P(O_t, Y_t = j \mid X_t = i) P(X_t = i \mid Y_{t-1} = k)$$

P.R.

$$\Rightarrow \sum_k \alpha_{k\,t-1} \sum_i P(O_t \mid X_t = i, Y_t = j) P(Y_t = j \mid X_t = i) a_{ki}$$

P.R.

$$\Rightarrow \sum_k \alpha_{k\,t-1} \sum_i b_{ij}(O_t)\, \pi_j\, a_{ki}$$

$$\Rightarrow \boxed{\sum_k \alpha_{k\,t-1} \sum_i b_{ij}(O_t)\, \pi_j\, a_{ki}} = \alpha_{jt}$$

c) $P(O_1, \ldots O_T) = \sum_j P(O_1 \ldots O_T, Y_T = j)$

marg

$$\cancel{= \sum_j \alpha_j T}$$

$$= \boxed{\sum_j \alpha_{jT}}$$

d) No. of values for $X_t \rightarrow n_x$

" " " $Y_t \rightarrow n_y$

Total observations $= T$

For all intermediary computations of α_{jt}, we have;

$$\alpha_{jt} = \sum_k \alpha_{kt-1} \sum_i b_{ij}(O_t) a_{ki} \pi_j$$

$$= \pi_j \sum_k \alpha_{kt-1} \sum_i b_{ij}(O_t) a_{ki}$$

$$= \pi_j \sum_i b_{ij}(O_t) \sum_k \alpha_{kt-1} a_{ki}$$

k iterates over n_y values & 'i' iterates over n_x values

$\Rightarrow$ $O(n_x n_y)$ is complexity for this step.

But since, $\sum_k \alpha_{kt-1} a_{ki}$ can be calculated in the outer loop, we need not repeat it for all rows of α_{jt} calculation

This will be repeated for T columns and thus the effective complexity is;

$$\boxed{O(T n_x n_y)}$$