تمرین سری ۱۰ درس BSP - رادین خیام - 9910579

سؤال ۱-

الف)

$$P(q_2 = s_3) = \sum_{i=1}^{4} \pi_i \times a_{i,3}$$

ب)

$$P(q_t = s_3 \mid q_{t-2} = s_4) = \sum_{i=1}^{4} \underbrace{P(q_{t-1} = s_i \mid q_{t-2} = s_4)}_{a_{4,i}} \times a_{i,3}$$

$$= a_{4,1} \times a_{1,3} + a_{4,2} \times a_{2,3} + a_{4,3} \times a_{3,3} + a_{4,4} \times a_{4,3}$$

این عبارت فرق دارد با $P(q_t = s_3)$ دلیل واضحی هم دارد، اگر قرار بود که $P(q_t = s_3) = P(q_t = s_3 \mid q_{t-2} = s_4)$ باشند این بدان معنی بود که $P(q_t = s_3)$ و $P(q_{t-2} = s_4)$ از هم مستقل می باشند که این موضوع درست نیست چون q_t و q_{t-1} به هم مربوط هستند و q_{t-1} و q_{t-2} هم به هم مربوط می باشند پس نمی‌شود که q_t و q_{t-2} مستقل باشند.

سؤال 2 -

الف)

$$O = \{ \overbrace{s_i, s_i, \dots\dots, s_i}^{d}, s_j \neq s_i \}$$

$$A = \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}$$

$$, P(O \mid \text{Model}, q_1 = s_i) = (a_{ii})^{d-1}(1 - a_{ii}) = p_i(d)$$

$$\bar{d}_i = \sum_{d=1}^{\infty} d\, p_i(d) = \sum_{d=1}^{\infty} d\,(a_{ii})^{d-1}(1 - a_{ii}) = \frac{1}{1 - a_{ii}}$$

$$\longrightarrow \boxed{\bar{d}_1 = 1} \quad , \quad \boxed{\bar{d}_2 = 2}$$

ب)

$$P(q_t = s_2 \mid q_{t-3} = s_2)$$

حالت های ممکن :

$$\{s_2, s_1, s_1, s_2\} \{s_2, s_1, s_2, s_2\} \{s_2, s_2, s_1, s_2\}$$

$$\{s_2, s_2, s_2, s_2\}$$

$$a_{2,1} \times \overset{0}{\cancel{a_{1,1}}} \times a_{1,2} + a_{2,1} \times a_{1,2} \times a_{2,2} + a_{2,2} \times a_{2,1} \times a_{1,2} + a_{2,2} \times a_{2,2} \times a_{2,2}$$

$$= 0.5 \times 1 \times 0.5 + 0.5 \times 0.5 \times 1 + 0.5 \times 0.5 \times 0.5 = \boxed{0.625}$$

سوال 3-

$$P(O|\lambda) = \sum_{q_1,q_2,q_3} \pi_{q_1} b_{q_1}(O_1)\, a_{q_1 q_2}\, b_{q_2}(O_2)\, a_{q_2 q_3}\, b_{q_3}(O_3)$$

* جمع سیگما روی تمام حالات q_i ها زده می شود.

* HMM ارگادیک هست پس a_{ij} همگی ناصفر هستند.

حال فرض می‌کنیم که حکم سوال درست باشد

$$P(O=000|\lambda) = \sum_{q_1,q_2,q_3} \pi_{q_1} b_{q_1}(0)\, a_{q_1 q_2}\, b_{q_2}(0)\, a_{q_2 q_3}\, b_{q_3}(0) = 0$$

① $\pi_2 = 0,\ b_1(0) = 0$ or ② $\pi_1 = 0,\ b_2(0) = 0$ or ③ $b_1(0) = b_2(0) = 0$

$$P(O=111|\lambda) = \sum_{q_1,q_2,q_3} \pi_{q_1} b_{q_1}(1)\, a_{q_1 q_2}\, b_{q_2}(1)\, a_{q_2 q_3}\, b_{q_3}(1) = 0$$

① $\pi_2 = 0,\ b_1(1) = 0$ or ② $\pi_1 = 0,\ b_2(1) = 0$ or ③ $b_1(1) = b_2(1) = 0$

* دقت شود که $b_1(0) + b_1(1) = 1$ و $b_2(0) + b_2(1) = 1$.

← با توجه به این جواب هایی که شامل ③ ③ باشند رد می شوند.
تنها می توانیم جواب های ② ① و ② ① را انتخاب کنیم.

$\pi_1 = \pi_2 = 0\ ,\ b_1(0) = b_2(1) = 0$

که این باعث می‌شود که احتمال همه مشاهدات صفر بشود که منطقی نیست.

برای همین این دو جواب هم رد می‌شوند و نتیجه می‌دهد که نمی‌شود احتمال آن دو مشاهده همزمان صفر شود.

سؤال 4 -

الف) $a_{i,j}=0 \ , \ i\neq j$ (*)

1) Initialization:

$$\delta_1(i) = \pi_i b_i(O_1), \qquad 1 \le i \le N \tag{32a}$$

$$\psi_1(i) = 0. \tag{32b}$$

2) Recursion:

$$\delta_t(j) = \max_{1\le i\le N} [\delta_{t-1}(i)a_{ij}]b_j(O_t), \qquad 2 \le t \le T, \quad 1 \le j \le N \tag{33a}$$

(*) $\delta_t(j) = \delta_{t-1}(j)\, a_{jj}\, b_j(O_t)$

$$\psi_t(j) = \operatorname*{argmax}_{1\le i\le N} [\delta_{t-1}(i)a_{ij}], \qquad 2 \le t \le T, \quad 1 \le j \le N. \tag{33b}$$

(*) $\psi_t(j) = j \ , \ 2\leq t\leq T$ (**)

3) Termination: (34a)

$$P^* = \max_{1\le i\le N} [\delta_T(i)]$$

$$q_T^* = \operatorname*{argmax}_{1\le i\le N} [\delta_T(i)]. \tag{34b}$$

* حال فرض می‌کنیم الگوریتم جلو رفته است و در این مرحله q_T^* پیدا شده است.

4) Path (state sequence) backtracking:

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \qquad t = T-1, T-2, \cdots, 1. \tag{35}$$

$q_t^* = \psi_{t+1}(q_{t+1}^*) \ , \ t=T-1,T-2,\dots,1$

(**) $\boxed{q_t^* = q_T^* \ , \ t=T-1,T-2,\dots,1}$

بنابراین نیازی به مرحله backtracking نمی‌باشد و همه استیت ها برابر با همان q_T^* می‌باشند.

ب) به ازای هر i ← $(*)\ \forall j: a_{ij} = cte$

$$a_{11} = a_{12} = \dots = a_{1N} = c$$

از طرفی میدانیم که اگر روی j ها جمع بزنیم، حاصل باید که یک بشود ←

$$a_{11} = a_{12} = \dots = a_{1N} = \frac{1}{N}$$

به همین ترتیب می بینیم که:

$$a_{21} = a_{22} = \dots = a_{2N} = \frac{1}{N}$$

و این چنین که همه $\boxed{a_{ij} = \frac{1}{N}}$

1) Initialization:

$$\delta_1(i) = \pi_i b_i(O_1), \qquad 1 \le i \le N \tag{32a}$$

$$\psi_1(i) = 0. \tag{32b}$$

2) Recursion:

$$\delta_t(j) = \max_{1 \le i \le N} [\delta_{t-1}(i) \cancel{a_{ij}}^{\frac{1}{N}}] b_j(O_t), \qquad 2 \le t \le T, \quad 1 \le j \le N \tag{33a}$$

$\frac{1}{N} =$ عدد ثابت

$$\psi_t(j) = \underset{1 \le i \le N}{\operatorname{argmax}} [\delta_{t-1}(i) a_{ij}], \qquad 2 \le t \le T, \quad 1 \le j \le N. \tag{33b}$$

→ $\psi_t(j) = \underset{1 \le i \le N}{\operatorname{argmax}} [\delta_{t-1}(i)]$

3) Termination: (34a)

$$P^* = \max_{1 \le i \le N} [\delta_T(i)]$$

$$q_T^* = \underset{1 \le i \le N}{\operatorname{argmax}} [\delta_T(i)]. \tag{34b}$$

4) Path (state sequence) backtracking:

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \qquad t = T-1, T-2, \cdots, 1. \tag{35}$$

→ $q_t^* = \psi_{t+1}(q_{t+1}^*) = \underset{1 \le i \le N}{\operatorname{argmax}} [\delta_t(i)]$

عملاً کاری که تو این مرحله 2 و وقتی $\delta_t(i)$ بدست میاد باید ببینیم کجا ماکسیمم هست ← این میشه q_t^* ما.

سؤال 5 -

الف)

$$O = \{O_1, O_2, O_3, O_4, O_5\} = \{S, R, R, S, R\}$$

$$Q = \{q_1, q_2, q_3, q_4, q_5\} = \{M, H, L, L, L\}$$

$$P(O|Q;\lambda) = b_1(1) \times b_2(2) \times b_3(2) \times b_3(1) \times b_3(2)$$

$$= \frac{1}{2} \times \frac{1}{4} \times \frac{3}{4} \times \frac{1}{4} \times \frac{3}{4} = \frac{9}{512} \simeq \boxed{0.0176}$$

$$P(Q;\lambda) = \pi_1 \times a_{12} \times a_{23} \times a_{33} \times a_{33} = 0.5 \times 0.4 \times 0.1 \times 0.5 \times 0.5 = \frac{1}{200}$$

$$P(O,Q|\lambda) = P(O|Q;\lambda)\, P(Q;\lambda) = \frac{9}{512} \times \frac{1}{200} = \frac{9}{102400}$$

حال می‌خواهیم $P(O;\lambda)$ را محاسبه کنیم، برای اینکار الگوریتم زیر را در متلب پیاده سازی می‌کنیم:

1) Initialization:

$$\alpha_1(i) = \pi_i b_i(O_1), \qquad 1 \le i \le N. \tag{19}$$

2) Induction:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}), \qquad 1 \le t \le T-1, \quad 1 \le j \le N. \tag{20}$$

3) Termination:

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i). \tag{21}$$

```
clc;
alpha = zeros(T,N);
% Initialization
for i = 1:N
    alpha(1,i) = P(i) * B(i, O(1));
end
% Induction
for t = 1:T-1
    for j =1:N
        tmp = 0;
        for i=1:N
            tmp = tmp + alpha(t,i)*A(i,j);
        end
        alpha(t+1,j) = tmp*B(j,O(t+1));
    end
end
% Termination
Prob = sum(alpha(T,:));
disp(['P(O;lambda) = ',num2str(Prob)]);
```

```
P(O;lambda) = 0.025416
```

$$\longrightarrow P(O;\lambda) = 0.0254$$

$$P(Q|O;\lambda) = \frac{P(O,Q;\lambda)}{P(O;\lambda)}$$

$$= \frac{\frac{9}{102400}}{\frac{254}{10000}} = \boxed{0.0034}$$

ب)

```
function [prob] = prob_calc(Q, O, P, A, B)
    T = length(Q);
    prob = P(Q(1));
    for i = 1:T-1
        prob = prob*B(Q(i),O(i))*A(Q(i),Q(i+1));
    end
    prob = prob*B(T),O(T));
end
```

* با در نظر گرفتن تمام حالت‌ها

```
P(O;lambda) = 0.025416
best Q = 1   3   3   3   3
P(Q_best,O ; Lambda) = 0.00098877
```

```
%% Considering all states
clc;
states = 1:3;
[q1, q2, q3, q4, q5] = ndgrid(states, states, states, states, states);
all_Qs = [q1(:), q2(:), q3(:), q4(:), q5(:)];

prob_OQ = zeros(1,243);
for i = 1:243
    Q_sample = all_Qs(i,:);
    prob_OQ(i) = prob_calc(Q_sample, O, P, A, B);
end
P_O = sum(prob_OQ);
disp(['P(O;lambda) = ',num2str(P_O)]);

[maxValue,maxIndex] = max(prob_OQ);
disp(['best Q = ',num2str(all_Qs(maxIndex,:))]);
disp(['P(Q_best,O ; Lambda) = ',num2str(maxValue)]);
```

ساخت تمام حالت‌ها برای Q

محاسبه $P(O,Q;\lambda)$ به ازای مقادیر مختلف Q

حاصل جمع تمامی حالات برابر است با $P(O;\lambda)$

پیدا کردن Q ای که $P(O,Q;\lambda)$ یا به طور متناظر $P(Q|O;\lambda)$ در آن ماکسیمم می‌شود.

* روش پیشرو:

```
clc;
alpha = zeros(T,N);
% Initialization
for i = 1:N
    alpha(1,i) = P(i) * B(i, O(1));
end
% Induction
for t = 1:T-1
    for j =1:N
        tmp = 0;
        for i=1:N
            tmp = tmp + alpha(t,i)*A(i,j);
        end
        alpha(t+1,j) = tmp*B(j,O(t+1));
    end
end
% Termination
Prob = sum(alpha(T,:));
disp(['P(O;lambda) = ',num2str(Prob)]);
```

1) Initialization:

$$\alpha_1(i) = \pi_i b_i(O_1), \qquad 1 \le i \le N. \tag{19}$$

2) Induction:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}), \qquad 1 \le t \le T-1, \quad 1 \le j \le N. \tag{20}$$

3) Termination:

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i). \tag{21}$$

```
P(O;lambda) = 0.025416
```

alpha =

0.2500	0.1500	0.0750
0.0825	0.0456	0.0956
0.0406	0.0163	0.0578
0.0209	0.0227	0.0107
0.0098	0.0052	0.0104

* با روش پسرو :

```
%% Backward Algorithm
clc;
beta = zeros(T,N);
% Initialization
beta(T,:) = 1;
% Induction
for t = T-1:-1:1
    for i =1:N
        tmp = 0;
        for j=1:N
            tmp = tmp + A(i,j)*B(j,O(t+1))*beta(t+1,j);
        end
        beta(t,i) = tmp;
    end
end
% Termination
Prob = 0;
for j = 1:N
    Prob = Prob+ P(j)*B(j,O(1))*beta(1,j);
end
disp(['P(O;lambda) = ',num2str(Prob)]);
```

1) Initialization:

$$\beta_T(i) = 1, \qquad 1 \le i \le N. \tag{24}$$

2) Induction:

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1})\, \beta_{t+1}(j),$$

$$t = T-1, T-2, \cdots, 1, \; 1 \le i \le N. \tag{25}$$

3) Termination

$$P(O|\lambda) = \sum_{j=1}^{N} \pi_j \, b_j(O(1)) \, \beta_1(j)$$

```
P(O;lambda) = 0.025416
```

```
beta =

    0.0544    0.0428    0.0719
    0.1064    0.0948    0.1288
    0.2363    0.2600    0.2000
    0.4750    0.4000    0.6000
    1.0000    1.0000    1.0000
```

* روش ترکیبی پیشرو - پسرو :

```
%% Forward - Backward Algorithm
clc;
t = 2;
Prob = 0;
for i = 1:N
    Prob = Prob + beta(t,i)*alpha(t,i);
end
disp(['P(O;lambda) = ',num2str(Prob)]);
```

```
P(O;lambda) = 0.025416
```

```
>> alpha(2,:)

ans =

    0.0825    0.0456    0.0956

>> beta(2,:)

ans =

    0.1064    0.0948    0.1288
```

پ)

```
%% Viterbi Algorithm
clc;
delta = zeros(T,N);
psi = zeros(T,N);

% Initialization
for i = 1:N
    delta(1,i) = P(i)*B(i,O(1));
    psi(1,i) = 0;
end
% Recursion
for t = 2:T
    for j=1:N
        delta(t,j) = max(delta(t-1,:).*A(:,j)')*B(j,O(t));
        [~, psi(t,j)] = max(delta(t-1,:).*A(:,j)');
    end
end
% Termination
P_star = max(delta(T,:));
q_star = zeros(1,T);
[~,q_star(T)] = max(delta(T,:));
% backtracking
for t=T-1:-1:1
    q_star(t) = psi(t+1,q_star(t+1));
end
disp(['Best Q = ',num2str(q_star)]);
```

```
Best Q = 1   3   3   3   3
```

* نتیجه مطابق با قسمت ب می باشد.

ت)

$$\xrightarrow{} P(O_1=S,\ O_2=R\ ;\lambda)$$

(تست

$$P(O_1,O_2)=\sum_{q_1,q_2}\pi_{q_1}\,b_{q_1}(O_1)\,a_{q_1q_2}\,b_{q_2}(O_2)$$

محاسبه با متلب

$$=\sum_{q_1,q_2}\pi_{q_1}\,b_{q_1}(S)\,a_{q_1q_2}\,b_{q_2}(R)=\boxed{0.2238}$$

$P(O_1=S\ ;\lambda)$ یا

جمع به ازای تمام حالات

$$P(O_1)=\sum_{q_1}\pi_{q_1}\,b_{q_1}(O_1)=\pi_1b_1(1)+\pi_2b_2(1)+\pi_3b_3(1)$$

$P(O_2=R\ ;\lambda)$

$$=\frac{1}{2}\times\frac{1}{2}+\frac{\cancel{2}\,1}{10}\times\frac{3}{\cancel{4}\,2}+\frac{3}{10}\times\frac{1}{4}=\frac{19}{40}\simeq\boxed{0.475}$$

$$P(O_2)=P(O_1=S,\ O_2=R\ ;\lambda)+P(O_1=R,\ O_2=R\ ;\lambda)$$

$$=0.2238+0.2738=\boxed{0.4976}$$

$$P(O_1=S\mid O_2=R\ ;\lambda)=\frac{P(O_1=S,\ O_2=R\ ;\lambda)}{P(O_2=R\ ;\lambda)}=\frac{0.2238}{0.4976}\simeq\boxed{0.45}$$

$$P(O_2=R\mid O_1=S\ ;\lambda)=\frac{P(O_1=S,\ O_2=R\ ;\lambda)}{P(O_1=S\ ;\lambda)}=\frac{0.2238}{0.475}\simeq\boxed{0.47}$$

$$P(O_2=R\mid q_1=L\ ;\lambda)=\sum_{3,q_2}\pi_3\,a_{3q_2}\,b_{q_2}(R)$$

$$=\pi_3a_{31}b_1(2)+\pi_3a_{32}b_2(2)+\pi_3a_{33}b_3(2)$$

$$=\frac{1}{2}\left(\frac{4}{10}\times\frac{1}{2}+\frac{1}{10}\times\frac{1}{4}+\frac{5}{10}\times\frac{3}{4}\right)=\frac{12}{40}=\frac{3}{10}=\boxed{0.3}$$

$$P(O_2=R\mid q_1=L,\ O_1=S\ ;\lambda)=P(O_1=S\ ;\lambda)\times P(O_2=R\mid (q_1=L \text{ And } O_1=S)\ ;\lambda)$$

$$=0.475\times\underbrace{\cancel{\pi_3}}_{\frac{3}{10}}\times\underbrace{b_{\cancel{2}}(1)}_{\frac{1}{\cancel{4}}}\ \underbrace{\left(a_{31}b_1(2)+a_{32}b_2(2)+a_{33}b_3(2)\right)}_{\frac{\cancel{24}\,6}{40}}=\boxed{0.02}$$