Introduction to NLP (CS7.401)

Written Assignment (28^{th} January, 2022)

Abhinav S Menon

We have $\lambda = (A, B, \pi)$ defined as

$$A = \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \end{bmatrix},$$
$$B = \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.4 & 0.5 \\ 0.6 & 0.1 & 0.3 \end{bmatrix},$$
$$\pi = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}.$$

Question 1

We need to find $P(RRGG \mid \lambda)$ using the forward procedure. The initial α values are:

$$\alpha_1(1) = \pi_1 \cdot B_{11} = \frac{1}{3} \cdot 0.3 = 0.1$$
$$\alpha_1(2) = \pi_2 \cdot B_{21} = \frac{1}{3} \cdot 0.1 = 0.033$$
$$\alpha_1(3) = \pi_3 \cdot B_{31} = \frac{1}{3} \cdot 0.6 = 0.2.$$

since the first observation is R, given by the first column of B.

For the next set,

$$\begin{aligned} \alpha_2(1) &= \left[\sum_{i=1}^{3} \alpha_1(i) A_{i1}\right] B_{11} \\ &= [(0.1)(0.1) + (0.033)(0.6) + (0.2)(0.3)] \cdot 0.3 \\ &= 0.09 \end{aligned}$$

$$\begin{aligned}
\alpha_2(2) &= \left[\sum_{i=1}^{3} \alpha_1(i) A_{i2}\right] B_{21} \\
&= [(0.1)(0.4) + (0.033)(0.2) + (0.2)(0.4)] \cdot 0.1 \\
&= 0.186
\end{aligned}$$

$$\begin{aligned}
\alpha_2(3) &= \left[\sum_{i=1}^{3} \alpha_1(i) A_{i3}\right] B_{31} \\
&= [(0.1)(0.5) + (0.033)(0.2) + (0.2)(0.3)] \cdot 0.6 \\
&= 0.176
\end{aligned}$$

For the next set,

$$\begin{aligned}
\alpha_3(1) &= \left[\sum_{i=1}^{3} \alpha_2(i) A_{i1}\right] B_{12} \\
&= [(0.09)(0.1) + (0.186)(0.6) + (0.176)(0.3)] \cdot 0.5 \\
&= 0.087
\end{aligned}$$

$$\begin{aligned}
\alpha_3(2) &= \left[\sum_{i=1}^{3} \alpha_2(i) A_{i2}\right] B_{22} \\
&= [(0.09)(0.4) + (0.186)(0.2) + (0.176)(0.4)] \cdot 0.4 \\
&= 0.191
\end{aligned}$$

$$\begin{aligned}
\alpha_3(3) &= \left[\sum_{i=1}^{3} \alpha_2(i) A_{i3}\right] B_{32} \\
&= [(0.09)(0.5) + (0.186)(0.2) + (0.176)(0.3)] \cdot 0.1 \\
&= 0.047
\end{aligned}$$

For the last set,

$$\begin{aligned}
\alpha_4(1) &= \left[\sum_{i=1}^{3} \alpha_3(i) A_{i1}\right] B_{12} \\
&= [(0.087)(0.1) + (0.191)(0.6) + (0.047)(0.3)] \cdot 0.5 \\
&= 0.0685
\end{aligned}$$

$$\begin{aligned}
\alpha_4(2) &= \left[\sum_{i=1}^{3} \alpha_3(i) A_{i2}\right] B_{22} \\
&= [(0.087)(0.4) + (0.191)(0.2) + (0.047)(0.4)] \cdot 0.4 \\
&= 0.242
\end{aligned}$$

$$
\begin{aligned}
\alpha_4(3) &= \left[\sum_{i=1}^{3} \alpha_3(i) A_{i3}\right] B_{32} \\
&= [(0.087)(0.5) + (0.191)(0.2) + (0.047)(0.3)] \cdot 0.1 \\
&= 0.057
\end{aligned}
$$

Thus, the total probability is

$$
P(RRGG \mid \lambda) = 0.0685 + 0.242 + 0.057 = 0.3675
$$

Question 2

We need to find the best state sequence using the Viterbi algorithm. The initial values of δ are the same as those of α:

$$
\begin{aligned}
\delta_1(1) &= 0.1 \\
\delta_1(2) &= 0.033 \\
\delta_1(3) &= 0.2
\end{aligned}
$$

and

$$
\psi_1(1) = \psi_1(2) = \psi_1(3) = 0.
$$

At the next level,

$$
\begin{aligned}
\delta_2(1) &= \max_{1 \le i \le 3} [\delta_1(i) A_{i1}] B_{11} \\
&= \max\{(0.1)(0.1), (0.033)(0.6), (0.2)(0.3)\} \cdot 0.3 \\
&= 0.018 \\
\psi_2(1) &= \operatorname*{argmax}_{1 \le i \le 3} [\delta_1(i) A_{i1}] \\
&= 3.
\end{aligned}
$$

$$
\begin{aligned}
\delta_2(2) &= \max_{1 \le i \le 3} [\delta_1(i) A_{i2}] B_{21} \\
&= \max\{(0.1)(0.4), (0.033)(0.2), (0.2)(0.4)\} \cdot 0.1 \\
&= 0.008 \\
\psi_2(2) &= \operatorname*{argmax}_{1 \le i \le 3} [\delta_1(i) A_{i2}] \\
&= 3.
\end{aligned}
$$

$$
\begin{aligned}
\delta_2(3) &= \max_{1 \le i \le 3} [\delta_1(i) A_{i3}] B_{31} \\
&= \max\{(0.1)(0.5), (0.033)(0.2), (0.2)(0.3)\} \cdot 0.6 \\
&= 0.036 \\
\psi_2(3) &= \operatorname*{argmax}_{1 \le i \le 3} [\delta_1(i) A_{i3}] \\
&= 3.
\end{aligned}
$$

At the next level,

$$
\begin{aligned}
\delta_3(1) &= \max_{1\leq i\leq 3}[\delta_2(i)A_{i1}]B_{12} \\
&= \max\{(0.018)(0.1),(0.008)(0.6),(0.036)(0.3)\}\cdot 0.5 \\
&= 0.0054 \\
\psi_3(1) &= \operatorname*{argmax}_{1\leq i\leq 3}[\delta_2(i)A_{i1}] \\
&= 3.
\end{aligned}
$$

$$
\begin{aligned}
\delta_3(2) &= \max_{1\leq i\leq 3}[\delta_2(i)A_{i2}]B_{22} \\
&= \max\{(0.018)(0.4),(0.008)(0.2),(0.036)(0.4)\}\cdot 0.4 \\
&= 0.00576 \\
\psi_3(2) &= \operatorname*{argmax}_{1\leq i\leq 3}[\delta_2(i)A_{i2}] \\
&= 3.
\end{aligned}
$$

$$
\begin{aligned}
\delta_3(3) &= \max_{1\leq i\leq 3}[\delta_2(i)A_{i3}]B_{32} \\
&= \max\{(0.018)(0.5),(0.008)(0.2),(0.036)(0.3)\}\cdot 0.1 \\
&= 0.00108 \\
\psi_3(3) &= \operatorname*{argmax}_{1\leq i\leq 3}[\delta_2(i)A_{i3}] \\
&= 3.
\end{aligned}
$$

At the last level,

$$
\begin{aligned}
\delta_4(1) &= \max_{1\leq i\leq 3}[\delta_3(i)A_{i1}]B_{12} \\
&= \max\{(0.0054)(0.1),(0.00576)(0.6),(0.00108)(0.3)\}\cdot 0.5 \\
&= 0.001728 \\
\psi_4(1) &= \operatorname*{argmax}_{1\leq i\leq 3}[\delta_3(i)A_{i1}] \\
&= 2.
\end{aligned}
$$

$$
\begin{aligned}
\delta_4(2) &= \max_{1\leq i\leq 3}[\delta_3(i)A_{i2}]B_{22} \\
&= \max\{(0.0054)(0.4),(0.00576)(0.2),(0.00108)(0.4)\}\cdot 0.4 \\
&= 0.0004608 \\
\psi_4(2) &= \operatorname*{argmax}_{1\leq i\leq 3}[\delta_3(i)A_{i3}] \\
&= 2.
\end{aligned}
$$

$$
\begin{aligned}
\delta_4(3) &= \max_{1\le i\le 3}[\delta_3(i)A_{i3}]B_{32} \\
&= \max\{(0.0054)(0.5), (0.00576)(0.2), (0.00108)(0.3)\} \cdot 0.1 \\
&= 0.0001152 \\
\psi_4(3) &= \operatorname*{argmax}_{1\le i\le 3}[\delta_3(i)A_{i3}] \\
&= 2.
\end{aligned}
$$

Thus, we have

$$P^* = \max_i \delta_4(i) = 0.001728$$

which is the probability of the path given by

$$
\begin{aligned}
q_4^* &= \operatorname*{argmax}_{1\le i\le 3} \delta_4(i) = 1 \\
q_3^* &= \psi_4(1) = 2 \\
q_2^* &= \psi_3(2) = 3 \\
q_1^* &= \psi_2(3) = 3,
\end{aligned}
$$

i.e., [3,3,2,1].