Hidden Markov Models (Modelos de Markov Escondidos)

Seja $\{X_n, n=1,2,\ldots\}$ uma cadeia de Markov com matriz de transição $P=[P_{ij}]$ e probabilidade inicial $p_i = P(X_1=i)$. Suponha que exista um conjunto finito S de sinais que é emitido cada vez que a cadeia de Markov entra em um estado. Assume-se que se S_n representa o n-ésimo sinal emitido então:

$$P(S_1 = s \mid X_1 = j) = p(s|j)$$

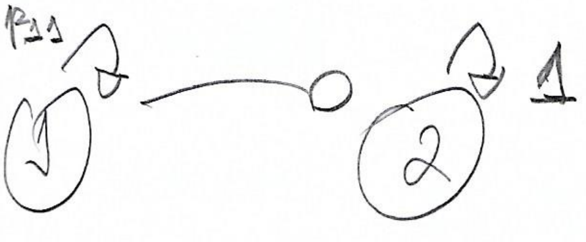
$$P(S_n = s \mid X_1, S_1, \ldots, X_{n-1}, S_{n-1}, X_n = j) = p(s|j)$$

$\rightarrow S_1, S_2 \ldots$ são observados
$\rightarrow X_1, X_2, \ldots$ não são observados
$\Big\}$ Cadeia de Markov escondida.

<u>Exemplo</u>: Máquina pode estar em 2 estados;

1 $\rightarrow$ Perfeitas condições

2 $\rightarrow$ Condições impróprias



Máquina é inspecionada e 2 possíveis resultados podem ser obtidos:

G $\rightarrow$ Bom estado

B $\rightarrow$ Mal estado

Temos então:

$$p(G|1), \quad p(B|1)$$

$$p(G|2), \quad p(G|2).$$

(3)

Seja $S^n = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$ o vetor aleatório com os n primeiras observações. Definimos

$$F_n(y) = P(S^n = s^n, X_n = y), \quad s^n = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$$

$$P(X_n = y \mid S^n = s^n) = \frac{P(S^n = s^n, X_n = y)}{P(S^n = s^n)}$$

$$= \frac{F_n(y)}{\sum_{s} F_n(s)}$$

$$F_n(j) = P(S^{n-1} = s^{n-1}, S_n = s_n, X_n = j) =$$

$$\sum_i P(S^{n-1} = s^{n-1}, S_n = s_n, X_{n-1} = i, X_n = j) =$$

$$\sum_i P(S_n = s_n, X_n = j \mid S^{n-1} = s^{n-1}, X_{n-1} = i) \cdot \underbrace{P(S^{n-1} = s^{n-1}, X_{n-1} = i)}_{F_{n-1}(i)} =$$

$$\sum_i P(S_n = s_n, X_n = j \mid X_{n-1} = i)\, F_{n-1}(i) =$$

$$\sum_i F_{n-1}(i)\, p_{ij}\, p(s_n \mid j)$$

pois

$$P(S_n = s_n, X_n = j \mid X_{n-1} = i) = \frac{P(S_n = s_n, X_n = j, X_{n-1} = i)}{P(X_{n-1} = i)} =$$

$$P(S_n = s_n \mid X_n = j, X_{n-1} = i)\, P(X_n = j \mid X_{n-1} = i) =$$

$$p(s_n \mid j)\, p_{ij}$$

Conclusão:

$$F_n(j) = \sum_i F_{n-1}(i)\, p_{ij}\, p(s_n|j), \quad n=2,3,\ldots$$

$$F_1(i) = P(X_1=i, S_1=s_1) = P(S_1=s_1|X_1=i)\,P(X_1=i)$$
$$= p_i\, p(s_1|i)$$

Exemplo: Considere no exemplo anterior:

$p_1 = 0,8$, $p_{11} = 0,9$, $\begin{cases} p(G|1) = 0,99, \\ p(G|2) = 0,10 \end{cases}$

Foram observados: $S_1 = G$, $S_2 = B$, $S_3 = G$.

(i) A prob. da máquina estar no estado 1 quando a 3ª inspeção foi realizada.

(ii) A prob. de $X_4 = 1$.

(iii) A prob. de $S_4 = G$.

6

$$\Omega^3 = \begin{pmatrix} G \\ B \\ G \end{pmatrix} \qquad F_n(j) = P(S^n = s^n, X_n = j)$$

$$F_1(1) \xrightarrow{} P(S_1 = G, X_1 = 1)$$

$$F_1(1) = p_1 \cdot p(G|1) = 0{,}8 \cdot 0{,}99 = 0{,}792$$

$$F_1(2) \xrightarrow{} P(S_1 = G, X_1 = 2)$$

$$F_1(2) = p_2\, P(G|2) = 0{,}2 \cdot 0{,}10 = 0{,}02$$

$$P(S^2 = s^2, X_2 = 1)$$

$$F_2(1) = \sum_{i=1}^{2} F_1(i)\, p_{i1}\, P(B|1)$$

$$= 0{,}792 \cdot 0{,}9 \cdot 0{,}01 = 0{,}007128$$

$$F_2(2) = \sum_{i=1}^{2} F_1(i) \cdot p_{i2}\, P(B|2)$$

$$= 0{,}792 \cdot 0{,}1 \cdot 0{,}90 + 0{,}02 \cdot 1 \cdot 0{,}9$$

$$= 0{,}08928$$

$=2^3, X_3 = \gamma)$

(7)

$$F_3(1) = \sum_{i=1}^{2} F_2(i)\, p_{i1}\, P(\sigma|1)$$

$$= 0{,}007128 \cdot 0{,}9 \cdot 0{,}99 = 0{,}006351$$

$$F_3(2) = \sum_{i=1}^{2} F_2(i)\, p_{i2}\, P(\sigma|2)$$

$$= 0{,}007128 \cdot 0{,}1 \cdot 0{,}1 + 0{,}08926 \cdot 1 \cdot 0{,}1$$

$$= 0{,}0089928$$

(i) $$P(X_3 = 1 \mid S^3 = 2^3) = \frac{F_3(1)}{F_3(1) + F_3(2)}$$

$$= \frac{0{,}006351}{0{,}006351 + 0{,}0089928} = 0{,}4137$$

$$P(X_4=1 \mid S^3=s^3) = \frac{P(X_4=1,\ X_3=1,\ S^3=s^3) + P(X_4=1, X_3=2, S^3=s^3)}{P(S^3=s^3)}$$

$$= P(X_4=1 \mid X_3=1, S^3=s^3) \cdot P(X_3=1 \mid S^3=s^3) +$$

$$P(X_4=1 \mid X_3=2,\ S^3=s^3) \cdot P(X_3=2 \mid S^3=s^3) =$$

$$= p_{11} \cdot 0{,}4137 + \cancel{p_{21} \cdot 0{,}5863}^{\;0\ \ 0}$$

$$= 0{,}9 \cdot 0{,}4137 = 0{,}37233.$$

(iii) $$P(S_4=G \mid S^3=s^3) = \frac{P(S_4=G,\ S^3=s^3, X_4=1) + P(S_4=G, S^3=s^3, X_4=2)}{P(S^3=s^3)}$$

$$= P(S_4=G \mid X_4=1,\ \cancel{S^3=s^3}) \cdot P(X_4=1 \mid S^3=s^3) +$$

$$P(S_4=G \mid X_4=2,\ \cancel{S^3=s^3}) \cdot P(X_4=2 \mid S^3=s^3)$$

$$P(G|1) \cdot \overset{0{,}3723}{\cancel{0{,}4137}} + P(G|2) \cdot (1-\overset{0{,}3723}{\cancel{0{,}4137}})$$

$$0{,}99 \cdot \underset{0{,}37223}{\cancel{0{,}4137}} + 0{,}10 \cdot \underset{\substack{(1-0{,}3723)\\ 0{,}6277}}{\cancel{0{,}5863}} = \underset{0{,}43313}{\cancel{0{,}4682}}$$

Estimação de Estados - Algoritmo de Viterbi:

Desejamos calcular

$$\max_{x^n} P\left(X^n = x^n \mid S^n = s^n\right)$$

onde

$$X^n = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}, \quad x^n = \begin{pmatrix} i_1 \\ \vdots \\ i_n \end{pmatrix}.$$

Como

$$P\left(X^n = x^n \mid S^n = s^n\right) = \frac{P\left(X^n = x^n, S^n = s^n\right)}{P\left(S^n = s^n\right)} \longrightarrow \text{não depende de } x^n$$

segue que o problema de maximização acima é equivalente a

$$\max_{x^n} P\left(X^n = x^n, S^n = s^n\right).$$

ra $t \le n$, defina

$$V_t(y) := \max_{i_1,\dots,i_{t-1}} P\left(X^{t-1} = \begin{pmatrix} i_1 \\ \vdots \\ i_{t-1} \end{pmatrix}, X_t = y, S^t = s^t\right).$$

Temos que

$$V_t(y) = \max_{i} \max_{i_1,\dots,i_{t-2}} P\left(X^{t-2} = \begin{pmatrix} i_1 \\ \vdots \\ i_{t-2} \end{pmatrix}, X_{t-1} = i, X_t = y, S^t = s^t\right)$$

$$= \max_{i} \max_{i_1,\dots,i_{t-2}} P\left(X^{t-2} = \begin{pmatrix} i_1 \\ \vdots \\ i_{t-2} \end{pmatrix}, X_{t-1} = i, S^{t-1} = s^{t-1}, X_t = y, S_t = s_t\right)$$

$$= \max_{i} \max_{i_1,\dots,i_{t-2}} P\left(X_t = y, S_t = s_t \,\middle|\, X^{t-2} = \begin{pmatrix} i_1 \\ \vdots \\ i_{t-2} \end{pmatrix}, X_{t-1} = i, \cancel{S^{t-1} = s^{t-1}}\right) \cdot$$

$$P\left(X^{t-2} = \begin{pmatrix} i_1 \\ \vdots \\ i_{t-2} \end{pmatrix}, X_{t-1} = i, S^{t-1} = s^{t-1}\right) =$$

$$= \max_{i} \max_{i_1,\dots,i_{t-2}} \underbrace{P\left(X_t = y, S_t = s_t \,\middle|\, X_{t-1} = i\right)}_{p_{iy}\, p(s_t \mid y)} P\left(X^{t-2} = \begin{pmatrix} i_1 \\ \vdots \\ i_{t-2} \end{pmatrix},\right.$$

$$\left. X_{t-1} = i, S^{t-1} = s^{t-1}\right) =$$

11

$$= \max_i p_{ij}\, p(s_n|j) \underbrace{\max_{i_1,\dots,i_{n-2}} P\left(X^{n-2} = \begin{pmatrix} i_1 \\ \vdots \\ i_{n-2} \end{pmatrix}, X_{n-1} = i, S^{n-1} = s^{n-1}\right)}_{V_{n-1}(i)}$$

$$= p(s_n|j) \max_i p_{ij} V_{n-1}(i)$$

Logo, Forward Algorithm →

$$\begin{cases} V_n(j) = p(s_n|j) \max_i p_{ij} V_{n-1}(i), & n \geq 2 \\ V_1(j) = P(X_1 = j, S_1 = s_1) = p(s_1|j)\, p_j \end{cases}$$

$$V_2(j) = p(s_2|j) \max_i p_{ij} V_1(j)$$

$$V_3(j) = p(s_3|j) \max_i p_{ij} V_2(j) \quad \text{---}$$

Obtenção da sequência de estados de otimização:

Direção Reversa: Backward Algorithm ←

$$j_n = \arg\max_j V_n(j)$$

Em geral, para $k < n$

$$i_k(j) = \arg\max_i p_{ij} V_k(i) \quad \cdot \text{ ponto}$$

$$V_k(j) = p(s_k|j) \cdot \max_i p_{ij} V_{k-1}(i) = p(s_k|j) \cdot p_{i_{k-1}(j)j} V_{k-1}(i_{k-1}(j))$$

$$\max_{i_1,\dots,i_n} P\left(X^n = \begin{pmatrix} i_1 \\ \vdots \\ i_n \end{pmatrix}, S^n = s^n\right) =$$

$$\max_j V_n(j) = V_n(j_n) =$$

$$\max_{i_1,\dots,i_{n-1}} P\left(X^n = \begin{pmatrix} i_1 \\ \vdots \\ i_{n-1} \\ j_n \end{pmatrix}, S^n = s^n\right) =$$

$$p(s_n | j_n) \, p_{i_{n-1}(j_n) j_n} V_{n-1}\left(i_{n-1}(j_n)\right)$$

Logo, a sequência ótima é:

$$j_n,\ i_{n-1}(j_n),\ i_{n-2}\left(i_{n-1}(j_n)\right), \dots, i_1(\dots).$$

Algoritmo de Viterbi!

0,9 B $\quad \frac{0,1}{0}$ $\quad$ 1 $\quad$ 2 $\quad S_1$ $\quad p_1 = 0,8 \quad p(G|1) = 0,99$

$p_2 = 0,2 \quad p(G|2) = 0,10$

13 $\quad S_1 = G$, $S_2 = B$, $S_3 = G$

No exemplo,

$n=1$

$$V_1(1) = p_1\, p(G|1) = 0,8 \cdot 0,99 = 0,792$$

$$V_1(2) = p_2\, p(G|2) = 0,2 \cdot 0,1 = 0,02$$

$n=2$

$V_n(j) = p(A_n|j) \cdot \max_i p_{ij} V_{n-1}(i)$

$$V_2(1) = p(B|1) \max_{i=1,2} p_{i1} V_1(i)$$

$$= 0,01 \cdot \max\{\overset{p_{11}\ V_1(1)}{0,9 \cdot 0,792},\ \overset{\cancel{p_{21}}^{\,0}\ V_1(2)}{0}\}$$

$$= 0,01 \cdot 0,9 \cdot 0,792 = 0,007128$$

$$V_2(2) = p(B|2) \max_{i=1,2} p_{i2} V_1(i)$$

$$= 0,9 \max\{\overset{p_{12}\ V_1(1)}{0,1 \cdot 0,792},\ \overset{p_{22}\ V_1(2)}{1 \cdot 0,02}\}$$

$$= 0,9 \max\{0,0792,\ 0,02\} = 0,07128$$

$n=3$

$$V_3(1) = p(G|1) \max_{i=1,2} p_{i1} V_2(i)$$

$$= 0,99 \cdot \max\{\overset{p_{11}\ V_2(1)}{0,9 \cdot 0,007128},\ \overset{\cancel{p_{21}}^{\,0}\ V_2(2)}{0}\}$$

$$= 0,006351$$

$$V_3(2) = p(G|2) \max_{i=1,2} p_{i2} V_2(i)$$

$$= 0,10 \max\{\overset{p_{12}\ V_2(1)}{0,1 \cdot 0,007128},\ \overset{p_{22}\ V_2(2)}{1 \cdot 0,003168}\}$$

$$= 0,10 \max\{0,0007128,\ 0,07128\} = 0,007128$$

$$i_3 = \arg\max\{V_3(1), V_3(2)\} =$$

$$= \arg\max\{\overset{1}{\underline{0{,}006351}}, \overset{2}{0{,}007128}\}$$

$$= 2$$

$$i_2(2) = \arg\max\{P_{12}V_2(1), P_{22}V_2(2)\}$$

$$= \arg\max\{0{,}1 \cdot 0{,}007128, 1 \cdot 0{,}07128\}$$

$$= 2$$

$$i_1(2) = \arg\max\{P_{12}V_1(1), P_{22}V_1(2)\}$$

$$= \arg\max\{0{,}1 \cdot 0{,}792, 1 \cdot 0{,}02\}$$

$$= \arg\max\{0{,}0792, 0{,}02\}$$

$$= \underline{1}$$

$$P(X^n = x^n, S^n = s^n) =$$

$$P\left(X^{n-2}=x^{n-2}, S^{n-1}=s^{n-1}, X_{n-1}=x_{n-1}, X_n = x_n, S_n = s_n\right) =$$

$$P\left(S_n = s_n \mid X_n = x_n, X_{n-1} = x_{n-1}, \cancel{X^{n-2}=x^{n-2}, S^{n-1}=s^{n-1}}\right) \cdot$$

$$P\left(X_n = x_n, X_{n-1}=x_{n-1}, X^{n-2}=x^{n-2}, S^{n-1}=s^{n-1}\right) =$$

$$p(s_n \mid x_n)\, P\left(X_n = x_n \mid X_{n-1} = x_{n-1}, \cancel{X^{n-2}=x^{n-2}, S^{n-1}=s^{n-1}}\right) \cdot$$

$$P\left(X^{n-1}=x^{n-1}, S^{n-1}=s^{n-1}\right) = p_{x_{n-1}x_n}\, p(s_n \mid x_n)\, P\left(X^{n-1}=x^{n-1},\right.$$

$$\left. S^{n-1}=s^{n-1}\right) = \prod_{k=1}^{n} p(s_k \mid x_k)\, p_{x_{k-1}x_k}$$

$$P(X_1 = x_1, S_1 = s_1) = P(S_1 = s_1 \mid X_1 = x_1)\, P(X_1 = x_1)$$

$$= p(s_1 \mid x_1)\, p_{x_1} \quad , \quad p_{x_0 x_1} := p_{x_1}$$

$$\max_{x^n} P\left(X^n = x^n, S^n = s^n\right) \Leftrightarrow$$

$$\max_{\{x_1, \ldots, x_n\}} \sum_{k=1}^{n} \ln\left(p_{x_{k-1}x_k}\, p(s_k \mid x_k)\right)$$