MATH 285: Stochastic Processes

math-old.ucsd.edu/~ynemish/teaching/285

Today: HMM. Viterbi algorithm

- Homework 4 is due on Friday, February 11, 11:59 PM

Hidden Markov Model

(Y_n) is a MC on S and transition probabilities $p(i,j)$

(X_n) is a stochastic process (non necessarily Markov) with state space $\mathbb{R}$ and $\mathbb{P}[X_n = x \mid Y_n = y] = e_y(x)$

$Z_n = (X_n, Y_n)$ is a MC with transition probabilities

$$\mathbb{P}[Z_{n+1} = (x', y') \mid Z_n = (x, y)] = p(y, y')\, e_{y'}(x')$$

- $x = (x_0, x_1, \dots, x_N)$ the observed sequence
- $y = (y_0, y_1, \dots, y_N)$ the state sequence
- $\mathbb{P}[x] = \mathbb{P}[X_0 = x_0, \dots, X_N = x_N]$
- $\mathbb{P}[x, y] = \mathbb{P}[X_0 = x_0, \dots, X_N = x_N, Y_0 = y_0, \dots, Y_N = y_N]$

Q: What is the probability that the hidden states are $(y_0, y_1, \dots, y_N)$ given that we observe $(x_0, x_1, \dots, x_N)$?

The forward algorithm

$\mathbb{P}[y|x] = \dfrac{\mathbb{P}[x,y]}{\mathbb{P}[x]}$ How to efficiently compute $\mathbb{P}[x]$?

- Initialization:

  For $y \in S$, set $\alpha_0(y) = \mathbb{P}[X_0 = y_0, Y_0 = y_0] = \mathbb{P}[Y_0 = y_0]\, e_{y_0}(x_0)$

- Recursion:

  For $y' \in S$ and $0 < n < N$ set $\alpha_{n+1}(y') = e_{y'}(x_{n+1}) \sum_{y \in S} \alpha_n(y)\, p(y, y')$

- Termination:

  $\mathbb{P}[x] = \sum_{y \in S} \alpha_N(y)$

Requires $O(N|S|^2)$ operations

Q: Given x, find y that maximizes $\mathbb{P}[y|x]$.

Most likely trajectory

Motivation: signal processing, speach recognition,
error correcting codes

Y_n – signal (uncontaminated)

X_n – signal with random noise

Receive the sequence $(x_0, x_1, \dots, x_N)$

What is the best guess for the values of $(y_0, y_1, \dots, y_N)$?

Mathematically: compute $y^* = \operatorname{argmax}_y \mathbb{P}[y|x]$, so that

$$\mathbb{P}[y^*|x] = \max_y \mathbb{P}[y|x]$$

- y^* always exists (finite state space)
- y^* is not necessarily unique

Computational complexity

Direct calculation:

- $\mathbb{P}[y|x]$ for fixed y $\quad O(N|S|^2)$ operations
- Repeat for all $y \in S^N$ $\quad |S|^N$ times
- Select the maximizer

In total $O(N|S|^{N+2})$ operations, grows exponentially in N

Viterbi algorithm:

- Recursive algorithmthat allows to compute y^*
- Complexity grows polynomially in N
- $\max\limits_{y} \mathbb{P}[y|x] = \frac{1}{\mathbb{P}[x]} \max\limits_{y} \mathbb{P}[x,y]$
- Define $V_n(y) := \max\limits_{y_0,\dots,y_{n-1}} \mathbb{P}[X_0=x_0,\dots,X_n=x_n, Y_0=y_0,\dots,Y_{n-1}=y_{n-1}, Y_n=y]$

Viterbi algorithm

$$V_n(y) := \max_{y_0,\dots,y_{n-1}\in S} \mathbb{P}[X_0=x_0,\dots,X_n=x_n,\ Y_0=y_0,\dots,Y_{n-1}=y_{n-1},\ Y_n=y]$$

Then $\max_{Y} \mathbb{P}[X,Y] = \max_{y\in S} V_N(y)$

Idea
- compute $\max_Y \mathbb{P}[X,Y]$ recursively
- backtrack to find the maximizing sequence

$$\begin{aligned}
&\mathbb{P}[X_0=x_0,\dots,X_n=x_n,X_{n+1}=x_{n+1},Y_0=y_0,\dots,Y_{n-1}=y_{n-1},Y_n=y,Y_{n+1}=y']\\
&= \mathbb{P}[Z_0=(x_0,y_0),Z_1=(x_1,y_1),\dots,Z_n=(x_n,y),Z_{n+1}=(x_{n+1},y')]\\
&= \mathbb{P}[Z_0=(x_0,y_0),Z_1=(x_1,y_1),\dots,Z_n=(x_n,y)]\,\mathbb{P}[Z_{n+1}=(x_{n+1},y')\mid Z_n=(x_n,y)]\\
&= \mathbb{P}[X_0=x_0,\dots,X_n=x_n,Y_0=y_0,\dots,Y_n=y]\,p(y,y')\,e_{y'}(x_{n+1})
\end{aligned}$$

Viterbi algorithm

$$V_{n+1}(y') = \max_{y_0,\dots,y} \mathbb{P}[X_0 = x_0, \dots, X_{n+1} = x_{n+1}, Y_0 = y_0, \dots, Y_n = y, Y_{n+1} = y']$$

$$= \max_{y_0,\dots,y} \mathbb{P}[X_0 = x_0, \dots, X_n = x_n, Y_0 = y_0, \dots, Y_n = y]\, p(y, y')\, e_{y'}(x_{n+1})$$

$$= e_{y'}(x_{n+1}) \max_{y} \left(p(y, y') \max_{y_0,\dots,y_{n-1}} \mathbb{P}[X_0 = x_0, \dots, X_n = x_n, Y_0 = y_0, \dots, Y_n = y] \right)$$

$$= e_{y'}(x_{n+1}) \max_{y} \left(p(y, y')\, V_n(y) \right)$$

This allows to compute $\max_{Y} \mathbb{P}[x, Y]$ recursively

- $V_0(y) = \mathbb{P}[X_0 = x_0, Y_0 = y] = \mathbb{P}[Y_0 = y]\, e_y(x_0)$
- $V_{n+1}(y') = e_{y'}(x_{n+1}) \max_{y} \left(p(y, y')\, V_n(y) \right)$
- $\max_{Y} \mathbb{P}[x, Y] = \max_{y} V_N(y)$

Viterbi algorithm

Backtracking: keep track of the element that maximizes $p(y,y')V_n(y)$:

- for $y,y' \in S$ define $W_{n+1}(y,y') := e_{y'}(x_{n+1})\,p(y,y')\,V_n(y)$
- for all $y' \in S$ find y that maximizes $W_{n+1}(y,y')$

$$\psi_n^*(y') := \underset{y}{\operatorname{argmax}}\ W_{n+1}(y,y')$$

- in particular $V_{n+1}(y') = W_{n+1}(\psi_n^*(y'), y')$
- set $y_N^* = \underset{y}{\operatorname{argmax}}\ V_N(y)$, $\quad y_n^* = \psi_n^*(y_{n+1}^*)$
- then $\max\limits_{Y} \mathbb{P}[X,Y] = V_N(y_N^*) = e_{y_N^*}(x_N) \max\limits_{y} p(y,y_N^*)V_{N-1}(y)$

$$= e_{y_N^*}(x_N)\,p(y_{N-1}^*, y_N^*)\,V_{N-1}(y_{N-1}^*) = \dots$$

$$\dots = e_{y_N^*}(x_N)\,p(y_{N-1}^*, y_N^*)\,e_{y_{N-1}^*}(x_{N-1})\,p(y_{N-2}^*, y_{N-1}^*) \cdots e_{y_1^*}(x_1)\,p(y_0^*, y_1^*)\,V_0(y_0^*)$$

Viterbi algorithm $\qquad O(N|S|^2)$

Initialization: $\qquad |S|$ operations

For $y \in S$, set $V_0(y) = \mathbb{P}[X_0 = x_0, Y_0 = y_0] = \mathbb{P}[Y_0 = y_0]\, e_{y_0}(x_0)$

Recursion: $\qquad O(2|S|^2 N + N)$

For $y, y' \in S$ and $0 < n \le N$ set $W_{n+1}(y, y') = V_n(y)\, p(y, y')\, e_{y'}(x_{n+1})$

Then compute $\psi_n^*(y') = \operatorname*{argmax}_y W_{n+1}(y, y')$

Set $V_{n+1}(y') = W_{n+1}(\psi_n^*(y'), y')$

Termination:

$\max_Y \mathbb{P}[X, Y] = \max_y V_N(y)$, define $y_N^* = \operatorname*{argmax}_y V_N(y)$

Backtracking

For $0 \le k < N$, set $y_k^* = \psi_k^*(y_{k+1}^*)$

There may be more than one maximizer