H26 1

(1) $W_i^{(t)}$ と $W_i^{(t-1)}$ の関係

題意より

$$\begin{cases} W_i^{(t)} = \prod_{j=1}^{t} P_i^{(j)}(x_j) & \cdots ① \\ W_i^{(t-1)} = \prod_{j=1}^{t-1} P_i^{(j)}(x_j) & \cdots ② \end{cases}$$

と表せるので、

①を②で両辺割ることで、

$$\frac{W_i^{(t)}}{W_i^{(t-1)}} = P_i^{(t)}(x_t) \qquad W_i^{(t)} = P_i^{(t)}(x_t)\,W_i^{(t-1)}$$

$V_i^{(t+1)}$ の計算式

重み付けの比率は

$$V_1^{(t+1)} : V_2^{(t+1)} \cdots : V_N^{(t+1)} = W_1^{(t)} : W_2^{(t)} : \cdots : W_N^{(t)}$$ より

$$V_i^{(t+1)} = \frac{W_i^{(t)}}{\sum_{j=1}^{N} W_j^{(t)}} = \frac{P_i^{(t)}(x_t)\,W_i^{(t-1)}}{\sum_{j=1}^{N} P_j^{(t)}(x_t)\,W_j^{(t-1)}}$$ と求まる

(2) 以下、各tにおける各予報士 i ($1 \leq i \leq N$) の尤度を保管するための長さNの配列をW と定義し、利用する。(但し、これはtにあわせて動的に更新されるものとする)

このもとで、t日目におけるシニア予報士の予測確率分布を各tについて計算する時間計算量 O(NT) のアルゴリズムの疑似コードを次ページに示す。

なお、コード中の //(コメントアウト)に、その内容の補足を記す。

```
W[i] = 1  (1≤i≤N)  //初期化

for t in {1..T}
    wsum = 0
    for i in {1..N}  // t日目の予測直前の各予報士の尤度の総和
        wsum = wsum + W[i]
    p0, p1 = 0, 0   // t日目におけるシニア予報士の P(X=0), P(X=1) の値
    for i in {1..N}
        weight = W[i] / wsum      // 各予報士の予測を重み付き和
        p0 = p0 + p_i^(t)(0) * weight
        p1 = p1 + p_i^(t)(1) * weight

    for i in {1..N}       // t日目の実データより W を更新
        W[i] = W[i] * p_i^(t)(x_t)
```

上記のアルゴリズムでは

$W[i] :=$ t日目に用いられる予報士 i の尤度 $(1 \leq i \leq N)$

が任意の t に対しループ不変式として成り立っていることに注意する。

この時、この総和を $O(N)$ で求めておけば、各 t において、任意の予報士 i の重みを $O(1)$ で計算できる。$(= W[i]/wsum)$

したがって、各 t において シニア予報士の予測確率分布も $O(N)$ で計算することができる。

t日目の最後では、予測後に得られた実データを利用して、W の配列を $O(N)$ で更新することで、ループ不変式を維持することに注意する。

計算量について評価する。

各 t において、幅 N のループを高々定数回 回しているため、$O(NT)$ で動作する。

(3) 定義より

$$\hat{p}^{(t)}(x_t) = \sum_{i=1}^{N} v_i^{(t)} p_i^{(t)}(x_i) \cdots ①$$

また、 $v_i^{(t)} = \dfrac{w_i^{(t-1)}}{\sum_{k=1}^{N} w_k^{(t-1)}} \cdots ②$ となる。

つまり

$$
\begin{aligned}
Loss(x^T) &= -\sum_{t=1}^{T} \log \hat{p}^{(t)}(x_t) \\
&= -\sum_{t=1}^{T} \left\{ \log \sum_{i=1}^{N} \frac{w_i^{(t-1)} p_i^{(t)}(x_i)}{\sum_{k=1}^{N} w_k^{(t-1)}} \right\} \\
&= -\sum_{t=1}^{T} \left\{ \log \frac{\sum_{i=1}^{N} w_i^{(t)}}{\sum_{k=1}^{N} w_k^{(t-1)}} \right\} \quad (\because (1)\text{の } w_i^{(t-1)} \text{ と } w_i^{(t)} \text{ の関係式}) \\
&= -\sum_{t=1}^{T} \log\left(\sum_{i=1}^{N} w_i^{(t)}\right) + \sum_{t=1}^{T} \log\left(\sum_{k=1}^{N} w_k^{(t-1)}\right) \\
&= \log \sum_{k=1}^{N} w_k^{(0)} - \log \sum_{i=1}^{N} w_i^{(T)} \quad (\because \text{差分は2項を除き相殺}) \\
&= \log N - \log \sum_{i=1}^{N} w_i^{(T)} \\
&= \log N - \log \sum_{i=1}^{N} \left(\prod_{t=1}^{T} p_i^{(t)}(x_t) \right)
\end{aligned}
$$

(但し、0日目までのデータ列に対する各予報士の尤度は $p_i^{(0)}(x_0) = 1$ より 1で与えられることを利用している。)

(4) i番目の予報士のT日間の累積予想損失を $Loss(i)$ とすると、

$$Loss(i) = -\sum_{t=1}^{T} \log p_i^{(t)}(x_t) = -\log \prod_{t=1}^{T} p_i^{(t)}(x_t) = -\log w_i^{(T)} \cdots ③$$

である。

この時 ③及び、(3)の結果より任意の i_0 ($1 \leq i_0 \leq N$) に対して

$$\mathrm{Loss}(i_0) + \log N - \mathrm{Loss}(x^T) = \log N - \log W_{i_0}^{(T)} - \left(\log N - \log \sum_{i=1}^{T} w_i^{(T)}\right)$$

$$= \log\left(\sum_{i=1}^{T} w_i^{(T)}\right) - \log W_{i_0}^{(T)}$$

$$\geqq 0.$$

$$\therefore \mathrm{Loss}(i_0) + \log N \geqq \mathrm{Loss}(x^T)$$

上式は i_0 に依存しないので.

$$\therefore \min_{1 \leq i \leq N} \mathrm{Loss}(i) + \log N \geqq \mathrm{Loss}(x^T) \quad \blacksquare$$