知的情報処理システム特論　第13回

二宮　崇

今日の講義の予定

- HMMの前向き後向きアルゴリズムによるパラメータ学習

- 教科書
  - 北研二(著) 辻井潤一(編) 言語と計算4 確率的言語モデル 東大出版会
  - Christopher M. Bishop “PATTERN RECOGNITION AND MACHINE LEARNING” Springer, 2006

HMMの教師無し学習
Unsupervised Learning of HMMs

- パラメータ推定
  - 訓練データ (入力)

訓練データ$: x_1 x_2 \cdots x_n$

$x_i:$ 文を表す記号列(単語列)。$x_i = o_{i1} o_{i2} o_{i3} \cdots o_{iT_i}$とする。

$T_i : x_i$の記号列長

  - パラメータ (出力)

$$
\begin{aligned}
\pi, a, b &= \mathop{\arg\max}_{\pi,a,b} \prod_{i=1}^{n} p(x_i) \\
&= \mathop{\arg\max}_{\pi,a,b} \prod_{i=1}^{n} p(o_{i1} o_{i2} \cdots o_{iT_i}) \\
&= \mathop{\arg\max}_{\pi,a,b} \prod_{i=1}^{n} \sum_{q_1 \in Q, q_2 \in Q, \cdots q_{T_i} \in Q} p(q_1 o_{i1} q_2 o_{i2} \cdots q_{T_i} o_{iT_i})
\end{aligned}
$$

HMMのEMアルゴリズム

- パラメータ更新式
  - π, a, b をπ', a', b' に更新

$$\pi'_q = \frac{\sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta)[q_1 = q]}{n}$$

$$= \frac{\sum_{i=1}^{n} \sum_{q_2 \in Q, \cdots, q_{T_i} \in Q} p(q q_2 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta)}{n}$$

$$= \frac{\text{先頭が}q\text{となる回数の期待値}}{n}$$

HMMのEMアルゴリズム

- パラメータ更新式
  - π, a, b をπ', a', b' に更新

$$a'_{q,r} = \frac{\displaystyle\sum_{i=1}^{n}\sum_{q_1\in Q,\cdots,q_{T_i}\in Q} p(q_1\cdots q_{T_i}\mid o_1\cdots o_{T_i};\theta)\left(\sum_{t=2}^{T_i}[q=q_{t-1}][r=q_t]\right)}{\displaystyle\sum_{r'\in Q}\sum_{i=1}^{n}\sum_{q_1\in Q,\cdots,q_{T_i}\in Q} p(q_1\cdots q_{T_i}\mid o_1\cdots o_{T_i};\theta)\left(\sum_{t=2}^{T_i}[q=q_{t-1}][r'=q_t]\right)}$$

$$= \frac{\displaystyle\sum_{i=1}^{n}\sum_{q_1\in Q,\cdots,q_{T_i}\in Q} p(q_1\cdots q_{T_i}\mid o_1\cdots o_{T_i};\theta)C(q,r;q_1\cdots q_{T_i})}{\displaystyle\sum_{r'\in Q}\sum_{i=1}^{n}\sum_{q_1\in Q,\cdots,q_{T_i}\in Q} p(q_1\cdots q_{T_i}\mid o_1\cdots o_{T_i};\theta)C(q,r';q_1\cdots q_{T_i})}$$

$$= \frac{\text{状態}q\text{から状態}r\text{へ遷移する回数の期待値}}{\text{状態}q\text{から遷移する回数の期待値}}$$

$C(q,r;q_1\cdots q_T)$　$q_1\cdots q_T$の中でqからrに遷移する回数

HMMのEMアルゴリズム

- パラメータ更新式
  - π, a, b を π', a', b' に更新

$$b'_{q,o} = \frac{\sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta) \left(\sum_{t=1}^{T_i} [q = q_t][o = o_t] \right)}{\sum_{o' \in \Sigma} \sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta) \left(\sum_{t=1}^{T_i} [q = q_t][o' = o_t] \right)}$$

$$= \frac{\sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta) C(q, o; q_1 \cdots q_{T_i}, o_1 \cdots o_{T_i})}{\sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta) C(q; q_1 \cdots q_{T_i})}$$

$$= \frac{\text{状態}q\text{に滞在し記号}o\text{を出力する回数の期待値}}{\text{状態}q\text{に滞在する回数の期待値}}$$

$C(q; q_1 \cdots q_T)$　$q_1 \cdots q_T$ の中で q が出現した回数

$C(q, o; q_1 \cdots q_T, o_1 \cdots o_T)$　$q_1 \cdots q_T o_1 \cdots o_T$ の中で q から o を出力する回数

HMMのためのEMアルゴリズムの問題点

- 期待値を計算するために全ての可能な隠れ状態を列挙すると文長に対し、指数爆発的計算量が必要

解決策：前向き後向きアルゴリズム。隠れ状態を列挙することなく、この期待値を計算する動的計画法。

前向きアルゴリズムと後ろ向きアルゴリズム

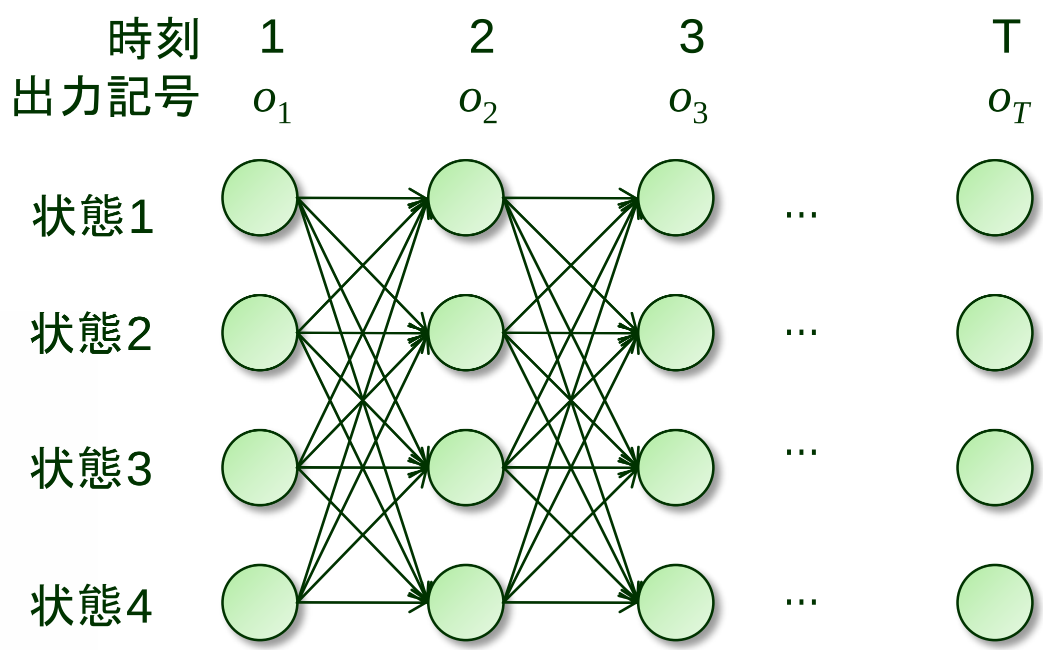
前向きアルゴリズム (1/3)

- 動的計画法

$\alpha(t, q)$: o_1 … o_tを出力し、o_tが出力される時(時刻t)に状態qである確率

$$\alpha(t,q) = \sum_{q_1 \in Q, \cdots, q_{t-1} \in Q} \pi_{q_1} b_{q_1,o_1} \left(\prod_{u=2}^{t-1} a_{q_{u-1},q_u} b_{q_u,o_u} \right) a_{q_{t-1},q} b_{q,o_t}$$

全てのt,qに対し、$\alpha(t, q)$を求めれば良い

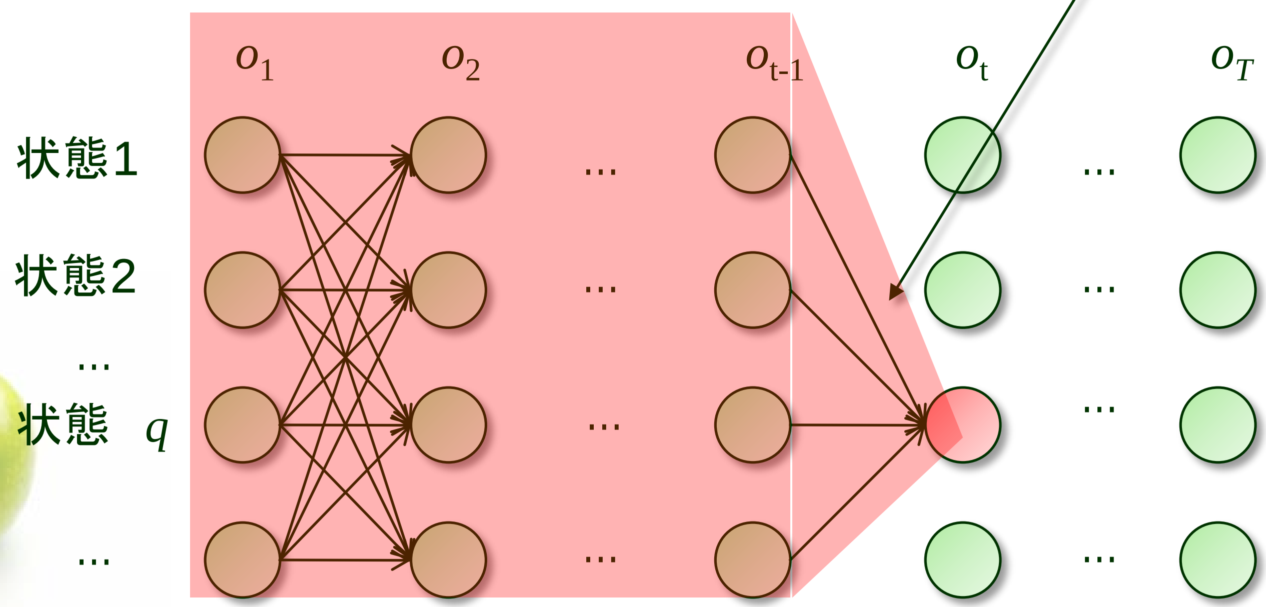


トレリス

前向きアルゴリズム (2/3)

$$
\begin{aligned}
\alpha(t,q) &= \sum_{q_1\in Q,\cdots,q_{t-1}\in Q} \pi_{q_1} b_{q_1,o_1} \left(\prod_{u=2}^{t-1} a_{q_{u-1},q_u} b_{q_u,o_u} \right) a_{q_{t-1},q} b_{q,o_t} \\
&= \sum_{q_{t-1}\in Q} \left\{ \sum_{q_1\in Q,\cdots,q_{t-2}\in Q} \left\{ \pi_{q_1} b_{q_1,o_1} \left(\prod_{u=2}^{t-2} a_{q_{u-1},q_u} b_{q_u,o_u} \right) a_{q_{t-2},q_{t-1}} b_{q_{t-1},o_{t-1}} \right\} a_{q_{t-1},q} b_{q,o_t} \right\} \\
&= \sum_{q_{t-1}\in Q} \alpha(t-1,q_{t-1}) a_{q_{t-1},q} b_{q,o_t}
\end{aligned}
$$

全ての遷移確率の和



前向きアルゴリズム (3/3)

$\alpha[1,q] := \pi[q]b[q,o_1]$ (for all q)

for t =2 to T

  for $q \in Q$

    $\alpha[t, q] := \Sigma_{q' \in Q} \{\alpha[t\text{-}1, q']a[q',q]b[q, o_t]\}$

後ろ向きアルゴリズム (1/3)

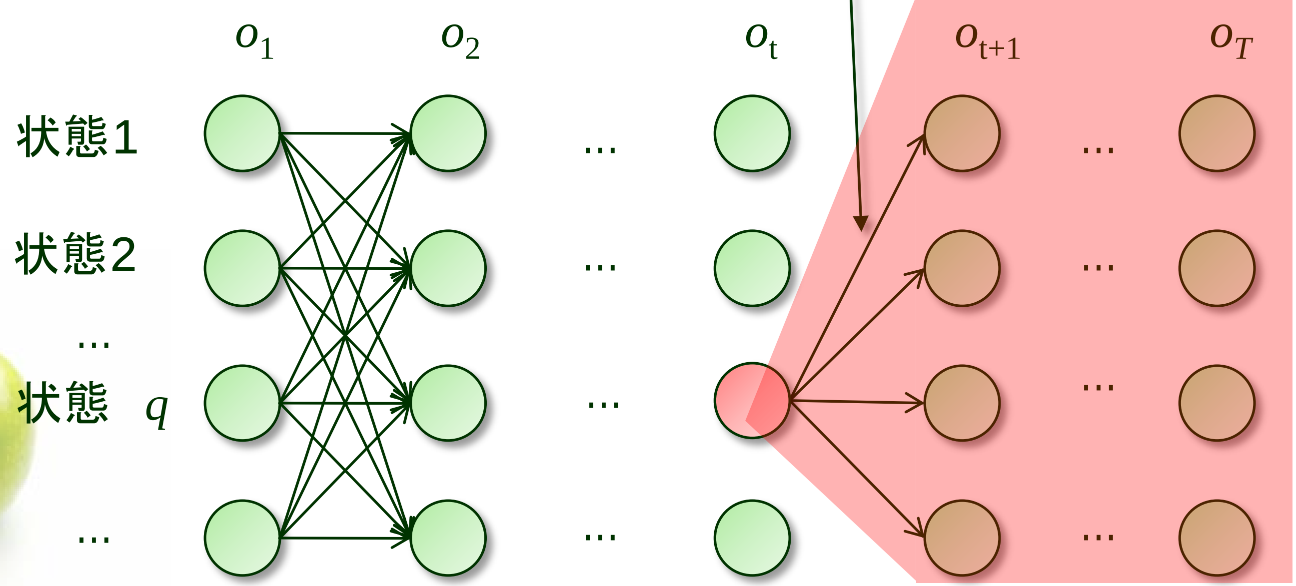
- 前向きアルゴリズムを逆方向に適用
  - 文末から文頭に向かって実行
  - 向きが逆になることを除けば前向きアルゴリズムとまったく同じ
- $\beta(t, q)$:時刻tに状態qから始める状態遷移によって$o_{t+1}\ldots o_T$まで出力する確率

$$\beta(t,q) = \sum_{q_{t+1}\in Q,\cdots,q_T\in Q} a_{q,q_{t+1}} b_{q_{t+1},o_{t+1}} \prod_{u=t+2}^{T} a_{q_{u-1},q_u} b_{q_u,o_u}$$

後ろ向きアルゴリズム (2/3)

$$
\begin{aligned}
\beta(t,q) &= \sum_{q_{t+1}\in Q,\cdots,q_T\in Q} a_{q,q_{t+1}} b_{q_{t+1},o_{t+1}} \prod_{u=t+2}^{T} a_{q_{u-1},q_u} b_{q_u,o_u} \\
&= \sum_{q_{t+1}\in Q} \left\{ a_{q,q_{t+1}} b_{q_{t+1},o_{t+1}} \sum_{q_{t+2}\in Q,\cdots,q_T\in Q} \left\{ a_{q_{t+1},q_{t+2}} b_{q_{t+2},o_{t+2}} \prod_{u=t+3}^{T} a_{q_{u-1},q_u} b_{q_u,o_u} \right\} \right\} \\
&= \sum_{q_{t+1}\in Q} a_{q,q_{t+1}} b_{q_{t+1},o_{t+1}} \beta(t+1,q_{t+1})
\end{aligned}
$$

全ての遷移確率の和



後ろ向きアルゴリズム (3/3)

$\beta[T,q] := 1$ (for all q)

for $t := T\text{-}1$ to 1

  for $q \in Q$

    $\beta[t, q] := \Sigma_{q' \in Q} \{a[q, q']b[q', o_{t+1}]\beta[t+1, q']\}$

前向き後ろ向きアルゴリズムの導出

前向き後ろ向きアルゴリズム: 生成確率

- 生成確率

$$p(o_1 \cdots o_{T_i}) = \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 o_1 \cdots q_{T_i} o_{T_i}) = \sum_{q \in Q} \alpha(T_i, q)$$

前向き後ろ向きアルゴリズム: π’

$$\pi'_q = \frac{1}{n}\sum_{i=1}^{n} \sum_{q_1\in Q,\cdots,q_{T_i}\in Q} p(q_1\cdots q_{T_i} \mid o_1\cdots o_{T_i};\theta)[q_1 = q]$$

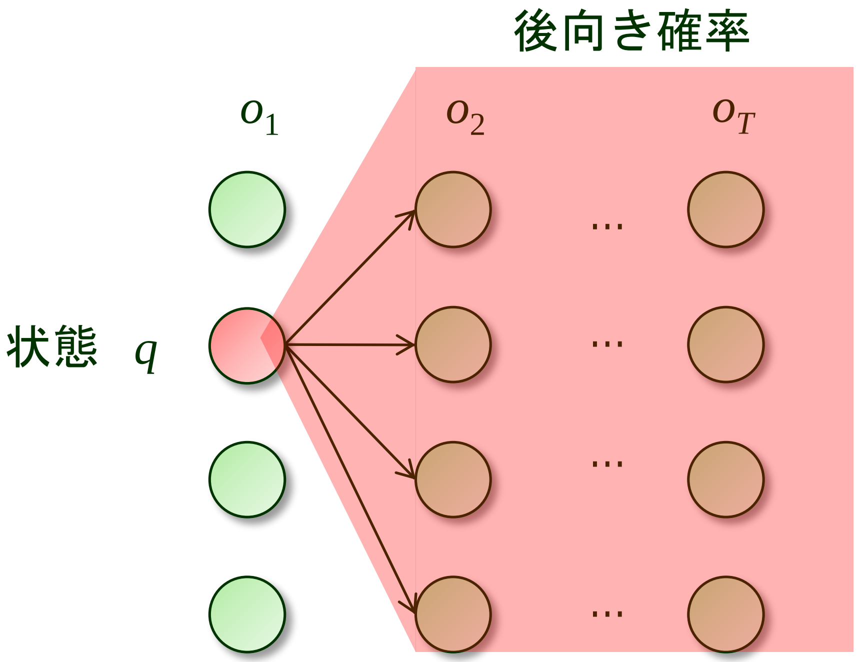
$$= \frac{1}{n}\sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{q_2\in Q,\cdots,q_{T_i}\in Q} \pi_q b_{q,o_1} a_{q,q_2} b_{q_2,o_2} \prod_{t=3}^{T_i} a_{q_{t-1},q_t} \prod_{t=3}^{T_i} b_{q_t,o_t}$$

$$= \frac{1}{n}\sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \pi_q b_{q,o_1} \beta(1,q)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \frac{\alpha(1,q)\beta(1,q)}{\sum_{q'\in Q}\alpha(T_i,q')}$$

前向き後ろ向きアルゴリズム: π'

$$\pi'_q = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{p(o_1\cdots o_{T_i};\theta)}\pi_q b_{q,o_1}\beta(1,q) = \frac{1}{n}\sum_{i=1}^{n}\frac{\alpha(1,q)\beta(1,q)}{\sum_{q'\in Q}\alpha(T_i,q')}$$



前向き後ろ向きアルゴリズム: a'

$$a'_{q,r}\text{の分子部分} = \sum_{i=1}^{n} \sum_{q_1\in Q,\cdots,q_{T_i}\in Q} p(q_1\cdots q_{T_i} \mid o_1\cdots o_{T_i};\theta)\left(\sum_{t=2}^{T_i}[q=q_{t-1}][r=q_t]\right)$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{q_1\in Q,\cdots,q_{T_i}\in Q} \pi_{q_1}\prod_{u=2}^{T_i} a_{q_{u-1},q_u}\prod_{u=1}^{T_i} b_{q_u,o_u}\left(\sum_{t=2}^{T_i}[q=q_{t-1}][r=q_t]\right)$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=2}^{T_i}\sum_{q_1\in Q,\cdots,q_{T_i}\in Q} \pi_{q_1}\prod_{u=2}^{T_i} a_{q_{u-1},q_u}\prod_{u=1}^{T_i} b_{q_u,o_u}[q=q_{t-1}][r=q_t]$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=2}^{T_i}\sum_{q_1\in Q,\cdots,q_{T_i}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-2} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-2},q_{t-1}} b_{q_{t-1},o_{t-1}} a_{q_{t-1},q_t} b_{q_t,o_t} a_{q_t,q_{t+1}} b_{q_{t+1},o_{t+1}}\left(\prod_{u=t+2}^{T_i} a_{q_{u-1},q_u} b_{q_u,o_u}\right)[q=q_{t-1}][r=q_t]$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=2}^{T_i}\sum_{q_1\in Q,\cdots,q_{t-2}\in Q}\sum_{q_{t+1}\in Q,\cdots,q_{T_i}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-2} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-2},q} b_{q,o_{t-1}} a_{q,r} b_{r,o_t} a_{r,q_{t+1}} b_{q_{t+1},o_{t+1}}\left(\prod_{u=t+2}^{T_i} a_{q_{u-1},q_u} b_{q_u,o_u}\right)$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=2}^{T_i}\sum_{q_1\in Q,\cdots,q_{t-2}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-2} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-2},q} b_{q,o_{t-1}} a_{q,r} b_{r,o_t} \sum_{q_{t+1}\in Q,\cdots,q_{T_i}\in Q} a_{r,q_{t+1}} b_{q_{t+1},o_{t+1}}\left(\prod_{u=t+2}^{T_i} a_{q_{u-1},q_u} b_{q_u,o_u}\right)$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=2}^{T_i}\sum_{q_1\in Q,\cdots,q_{t-2}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-2} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-2},q} b_{q,o_{t-1}} a_{q,r} b_{r,o_t} \beta(t,r)$$

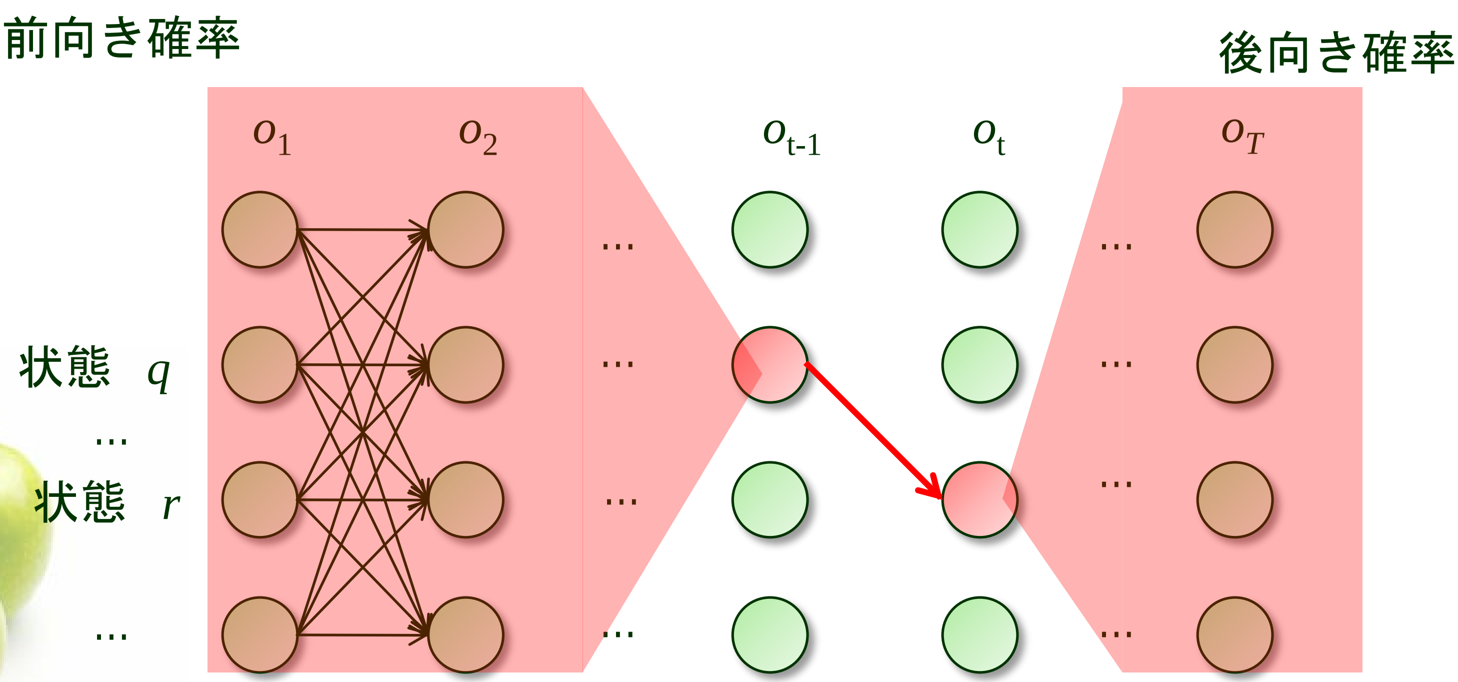
$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=2}^{T_i} a_{q,r} b_{r,o_t} \beta(t,r) \sum_{q_1\in Q,\cdots,q_{t-2}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-2} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-2},q} b_{q,o_{t-1}}$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=2}^{T_i} \alpha(t-1,q) a_{q,r} b_{r,o_t} \beta(t,r)$$

前向き後ろ向きアルゴリズム: a'

$\gamma(t,q,r)$: 記号列$o_1 \ldots o_T$に対し、時刻$t-1$から時刻tに状態qから状態rに遷移する確率

$$\gamma(t,q,r)=\frac{\alpha(t-1,q)a_{q,r}b_{r,o_t}\beta(t,r)}{\sum_{q'\in Q}\alpha(T,q')}$$



前向き後ろ向きアルゴリズム: a'

$$a'_{q,r} = \frac{\displaystyle\sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta) \left(\sum_{t=2}^{T_i} [q = q_{t-1}][r = q_t] \right)}{\displaystyle\sum_{r' \in Q} \sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta) \left(\sum_{t=2}^{T_i} [q = q_{t-1}][r' = q_t] \right)}$$

$$= \frac{\displaystyle\sum_{i=1}^{n} \sum_{t=2}^{T_i} \gamma(t, q, r)}{\displaystyle\sum_{r' \in Q} \sum_{i=1}^{n} \sum_{t=2}^{T_i} \gamma(t, q, r')}$$

前向き後ろ向きアルゴリズム: *b'*

$$b'_{q,o}\text{の分子部分} = \sum_{i=1}^{n} \sum_{q_1\in Q,\cdots,q_{T_i}\in Q} p(q_1\cdots q_{T_i} \mid o_1\cdots o_{T_i};\theta)\left(\sum_{t=1}^{T_i}[q=q_t][o=o_t]\right)$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=1}^{T_i} \sum_{q_1\in Q,\cdots,q_{T_i}\in Q} \pi_{q_1} \prod_{u=2}^{T_i} a_{q_{u-1},q_u} \prod_{u=1}^{T_i} b_{q_u,o_u} [q=q_t][o=o_t]$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=1}^{T_i} \sum_{q_1\in Q,\cdots,q_{T_i}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-1} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-1},q_t} b_{q_t,o_t} a_{q_t,q_{t+1}} b_{q_{t+1},o_{t+1}} \left(\prod_{t=u+2}^{T_i} a_{q_{t-1},q_t} b_{q_t,o_t}\right)[q=q_t][o=o_t]$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=1}^{T_i} \sum_{q_1\in Q,\cdots,q_{t-1}\in Q} \sum_{q_{t+1}\in Q,\cdots,q_{T_i}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-1} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-1},q} b_{q,o_t} a_{q,q_{t+1}} b_{q_{t+1},o_{t+1}} \left(\prod_{t=u+2}^{T_i} a_{q_{t-1},q_t} b_{q_t,o_t}\right)[o=o_t]$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=1}^{T_i} \sum_{q_1\in Q,\cdots,q_{t-1}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-1} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-1},q} b_{q,o_t}[o=o_t] \sum_{q_{t+1}\in Q,\cdots,q_{T_i}\in Q} a_{q,q_{t+1}} b_{q_{t+1},o_{t+1}} \left(\prod_{t=u+2}^{T_i} a_{q_{t-1},q_t} b_{q_t,o_t}\right)$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=1}^{T_i} \sum_{q_1\in Q,\cdots,q_{t-1}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-1} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-1},q} b_{q,o_t}[o=o_t]\beta(t,q)$$

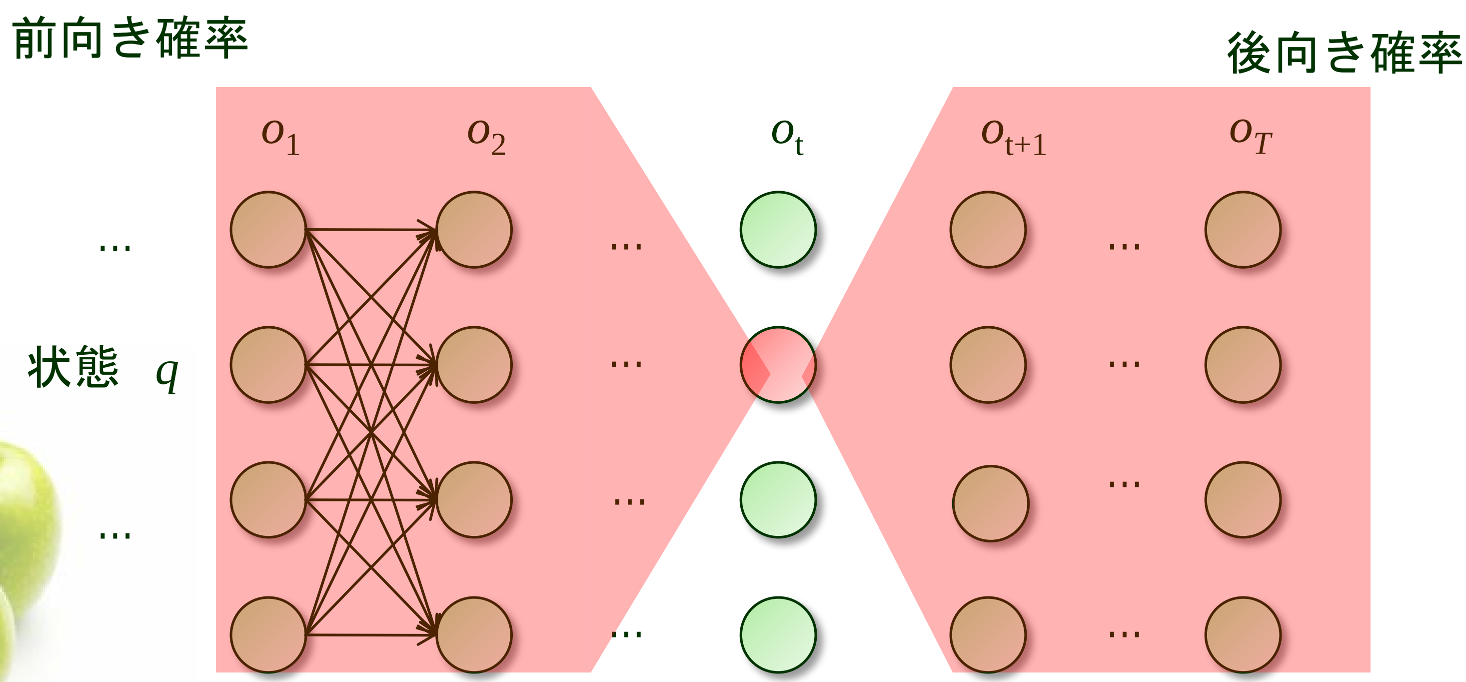
$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=1}^{T_i} [o=o_t]\beta(t,q) \sum_{q_1\in Q,\cdots,q_{t-1}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-1} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-1},q} b_{q,o_t}$$

$$= \sum_{i=1}^{n} \frac{1}{p(o_1\cdots o_{T_i};\theta)} \sum_{t=1}^{T_i} [o=o_t]\alpha(t,q)\beta(t,q)$$

前向き後ろ向きアルゴリズム: *b'*

$\gamma(t,q)$: 記号列$o_1 \ldots o_T$に対し、時刻tに状態qに滞在する確率

$$\gamma(t,q) = \frac{\alpha(t,q)\beta(t,q)}{\sum_{q' \in Q} \alpha(T,q')}$$



前向き後ろ向きアルゴリズム: b'

$$b'_{q,o} = \frac{\sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta) \left(\sum_{t=1}^{T_i} [q = q_t][o = o_t] \right)}{\sum_{o' \in \Sigma} \sum_{i=1}^{n} \sum_{q_1 \in Q, \cdots, q_{T_i} \in Q} p(q_1 \cdots q_{T_i} \mid o_1 \cdots o_{T_i}; \theta) \left(\sum_{t=1}^{T_i} [q = q_t][o' = o_t] \right)}$$

$$= \frac{\sum_{i=1}^{n} \sum_{t=1}^{T_i} [o = o_t] \gamma(t, q)}{\sum_{o' \in \Sigma} \sum_{i=1}^{n} \sum_{t=1}^{T_i} [o' = o_t] \gamma(t, q)}$$

$$= \frac{\sum_{i=1}^{n} \sum_{t=1}^{T_i} [o = o_t] \gamma(t, q)}{\sum_{i=1}^{n} \sum_{t=1}^{T_i} \gamma(t, q)}$$

ちなみに、 $$\pi'_q = \frac{1}{n} \sum_{i=1}^{n} \frac{\alpha(1, q) \beta(1, q)}{\sum_{q' \in Q} \alpha(T_i, q')} = \frac{1}{n} \gamma(1, q)$$

前向き後向きアルゴリズム まとめ(1/2)

$\alpha(t, q)$: o_1 … o_tを出力し、o_tが出力される時(時刻t)に状態qである確率

$$\begin{aligned}\alpha(t,q) &= \sum_{q_1\in Q,\cdots,q_{t-1}\in Q} \pi_{q_1} b_{q_1,o_1}\left(\prod_{u=2}^{t-1} a_{q_{u-1},q_u} b_{q_u,o_u}\right) a_{q_{t-1},q} b_{q,o_t} \\ &= \sum_{q_{t-1}\in Q} \alpha(t-1,q_{t-1}) a_{q_{t-1},q} b_{q,o_t}\end{aligned}$$

$\beta(t, q)$:時刻tに状態qから始める状態遷移によってo_{t+1}…o_Tまで出力する確率

$$\begin{aligned}\beta(t,q) &= \sum_{q_{t+1}\in Q,\cdots,q_T\in Q} a_{q,q_{t+1}} b_{q_{t+1},o_{t+1}} \prod_{u=t+2}^{T} a_{q_{u-1},q_u} b_{q_u,o_u} \\ &= \sum_{q_{t+1}\in Q} a_{q,q_{t+1}} b_{q_{t+1},o_{t+1}} \beta(t+1,q_{t+1})\end{aligned}$$

$$\gamma(t,q) = \frac{\alpha(t,q)\beta(t,q)}{\sum_{q'\in Q}\alpha(T,q')} \qquad \gamma(t,q,r) = \frac{\alpha(t-1,q)a_{q,r}b_{r,o_t}\beta(t,r)}{\sum_{q'\in Q}\alpha(T,q')}$$

前向き後向きアルゴリズム まとめ(2/2)

あるパラメータπ, a, bが与えられているとする

- αを求める(前向きアルゴリズム)
- βを求める(後向きアルゴリズム)
  - それぞれ動的計画法。計算量はO(T)　※ビタビアルゴリズムも計算量はO(T)。ビタビアルゴリズムのおよそ3倍の計算量。
- 次のようにπ, a, bを更新する

$$\pi'_q = \frac{1}{n}(\text{先頭が}q\text{となる回数の期待値}) = \frac{1}{n}\sum_{i=1}^{n}\gamma(1,q)$$

$$a'_{q,r} = \frac{\text{状態}q\text{から状態}r\text{へ遷移する回数の期待値}}{\text{状態}q\text{から遷移する回数の期待値}} = \frac{\sum_{i=1}^{n}\sum_{t=2}^{T_i}\gamma(t,q,r)}{\sum_{r'\in Q}\sum_{i=1}^{n}\sum_{t=2}^{T_i}\gamma(t,q,r')}$$

$$b'_{q,o} = \frac{\text{状態}q\text{に滞在し記号}o\text{を出力する回数の期待値}}{\text{状態}q\text{に滞在する回数の期待値}} = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T_i}[o=o_t]\gamma(t,q)}{\sum_{i=1}^{n}\sum_{t=1}^{T_i}\gamma(t,q)}$$

まとめ

- HMMの前向き後向きアルゴリズム
  - 前向き後向きアルゴリズム
    - 動的計画法
      - 計算量はO(T)　※ビタビアルゴリズムも計算量はO(T)。ビタビアルゴリズムのおよそ3倍の計算量。
    - 隠れ状態を全て並べあげることなく、トレリス上で期待値を計算できる
      - ある状態からある状態への遷移回数の期待値
      - ある状態からある記号を出力する回数の期待値
- 今までの講義資料

http://aiweb.cs.ehime-u.ac.jp/~ninomiya/iips/