- راه حل دوم مسئله دوم:

نام این الگوریتم ویتربی (Viterbi) می باشد.

تعریف زیر را در نظر بگیرید:

$$\delta_t(i) = \max_{q_1,q_2,\cdots,q_{t-1}} P[q_1,q_2,\cdots,q_{t-1},q_t = i,o_1,o_2,\cdots,o_t \mid \lambda]$$
$$1 \le i \le N$$

P محتمل ترین دنباله حالات با این شرط است که در زمان t در حالت i باشیم و مشاهدات o1,...ot را دیده باشیم.

از فرمول δ مشخص است که یک ماکزیمم گیری رو دنباله های پیشین انجام می شود.

پس فرمول بازگشتی زیر را داریم:

$$\delta_{t+1}(j) = [\max_i \delta_t(i)a_{ij}].b_j(o_{t+1})$$

کل روند الگوریتم ویتربی به صورت زیر است:

- مقداردهی اولیه:

$$\delta_1(i) = \Pi_i b_i(o_1), 1 \le i \le N$$
$$\psi_1(i) = 0$$

از Ψ برای ذخیره مسیر استفاده می شود.

$\psi_t(i)$ محتمل ترین حالت قبل از حالت i در زمان t می باشد.

- حلقه:

فرمول های بازگشتی زیر مهم ترین قسمت الگوریتم هستند:

$$\delta_t(j) = \max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}]b_j(o_t)$$
$$\psi_t(j) = \arg\max_{1\le i\le N}[\delta_{t-1}(i)a_{ij}]$$
$$2 \le t \le T, 1 \le j \le N$$

- پایان:

$$p^* = \max_{1\le i\le N}[\delta_T(i)]$$
$$q_T^* = \arg\max_{1\le i\le N}[\delta_T(i)]$$

- بازگشت به عقب برای پیمایش معکوس مسیر:

$$q_t^* = \psi_{t+1}(q_{t+1}^*), t = T-1, T-2, \cdots, 1$$

1– مقدمه

اهداف درس:

در این فصل با آموزش مدل مخفی مارکوف آشنا می شویم

آموزش هم برای حالت گسسته و هم پیوسته بررسی خواهد شد

2– مسئله سوم برای HMMهای گسسته

تعریف مسئله سوم:

فرض کنید یک مدل λ و یک دنباله مشاهدات O داریم، چگونه می توان پارمترهای مدل را تنظیم کرد که $P(O|\lambda)$ بیشینه شود (به عبارتی آموزش مدل از روی مشاهدات)؟

تخمین پارامترها بوسیله الگوریتم بیشینه سازی Expectation انجام می شود.

قبل از شروع بحث تعریف زیر را در نظر بگیرید:

$$\xi_t(i,j) = P(q_t = i, q_{t+1} = j \mid o, \lambda) = \frac{P(o, q_t = i, q_{t+1} = j \mid \lambda)}{P(o \mid \lambda)} = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}$$

این مقدار احتمال پرش از حالت i به حالت j در زمان t را نشان می دهد.

نتیجه می شود که $\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j)$ یرا اگر همه j ها را در نظر بگیریم به احتمال وجود در حالت i در زمان t می رسیم که همان γ است.

- $\sum_{t=1}^{T-1} \gamma_t(i)$: مقدار expected (از لحاظ آماری) تعداد پرش ها از حالت i
- $\sum_{t=1}^{T} \xi_t(i,j)$: مقدار expected (از لحاظ آماری) تعداد پرش ها از حالت i به حالت j

با تعریف یک مقدار expectation و بیشینه کردن آن به فرمول ها زیر می رسیم (برای دیدن اثبات فرمول های زیر به کتاب درسی مراجعه کنید):

$$\bar{\Pi}_i = \gamma_1(i) \qquad \bar{a}_{ij} = \frac{\sum_{t=1}^{T} \xi_t(i,j)}{\sum_{t=1}^{T} \gamma_t(i)} \qquad \bar{b}_j(k) = \frac{\sum_{t=1}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)}_{o_t = V_k}$$

فرمول های بالا برای آپدیت پارامترها (a و b و π) در هر تکرار از آموزش به کار می روند.

$$Q(\lambda' \mid \lambda) = \sum_{q} P(o, q \mid \lambda') \log P(o, q \mid \lambda)$$