
بسم الله الرحمن الرحيم

Prediction of Gold Price Using Hidden Markov Model

ارایه دهنده:
مونا رستگار

استاد راهنما:
جناب آقای دکتر علی کتانفروش

دانشگاه شهید بهشتی

بهمن ۱۴۰۱

چشم‌انداز

۱	مقدمه	۲
۲	زنجیره‌ی مارکوف	۴
۳	مدل‌های مارکوف قابل مشاهده	۸
۴	مدل پنهان مارکوف	۱۴
۵	مسائل سه‌گانه	۲۲
۶	سایر منابع	۵۱

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۱ مقدمه

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۱.۱ پیش‌گفتار

تا کنون فرض می‌شد که نمونه‌ها « متغیرهای تصادفی مستقل با توزیع یکسان (iid) » هستند.

- مزیت این فرض سادگی محاسبه‌ی درست‌نمایی است.
- در عین حال، برای برخی کاربردها که نمونه‌های متوالی دارند، این پیشفرض پذیرفتنی نیست.
  - به عنوان مثال حروف یک کلمه وابستگی دارند، به عنوان مثال در زبان انگلیسی حرف h با احتمال یکسانی بعد از حرف‌های t و x ظاهر نمی‌شود.
  - بازشناسی صدا نیز مربوط به شناسایی واج‌هایی است که به یکدیگر وابسته هستند و تنها توالی مشخصی از این واج‌ها معتبر هستند، در سطحی بالاتر هر ترتیبی از کلمه‌ها نیز مجاز نیستند.
- یک «فرآیند تصادفی پارامتری» (Parametric random process) می‌تواند توالی نمونه‌ها را تولید کند.

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۲ زنجیره‌ی مارکوف

۱.۲ فرآیندهای گسسته‌ی مارکوف (Discrete Markov Process)

سیستمی را در نظر بگیرید که در هر لحظه از زمان در یکی از N حالت مشخص شده باشد.

$$S_1, S_2, \ldots, S_N \tag{۱}$$

حالت سیستم در زمان t با q_t نمایش داده می‌شود:

$$q_t = S_i \quad \Longrightarrow \quad \text{در زمان } t \text{ سیستم در حالت } S_i \text{ می‌باشد.} \tag{۲}$$

احتمال تغییر حالت سیستم به حالتی دیگر با توجه به حالت‌های قبلی سیستم تعیین می‌شود:

$$P(q_{t+1} = S_j \quad | \quad q_t = S_i, \quad q_{t-1} = S_k, \quad \ldots) \tag{۳}$$

۲.۲ فرآیندهای گسسته‌ی مارکوف (ادامه...) (Discrete Markov Process)

برای حالت خاصی از مدل مارکوف، حالت در زمان $t+1$ تنها به حالت در زمان t بستگی دارد، که به آن «مدل مارکوف مرتبه‌ی اول» (First-order Markov Model) می‌گویند.

$$P(q_{t+1} = S_j | q_t = S_i, q_{t-1} = S_k, \dots) = P(q_{t+1} = S_j | q_t = S_i) \qquad (۴)$$

با فرض این که «احتمال گذرا» (transition) مستقل از زمان باشد:

$$a_{ij} = (q_{t+1} = S_j | q_t = S_i) \qquad a_{ij} \geq 0 \quad \text{and} \quad \sum_{j=1}^{N} a_{ij} = 1 \qquad (۵)$$

احتمال اولیه، احتمال این است که اولین حالت S_i باشد:

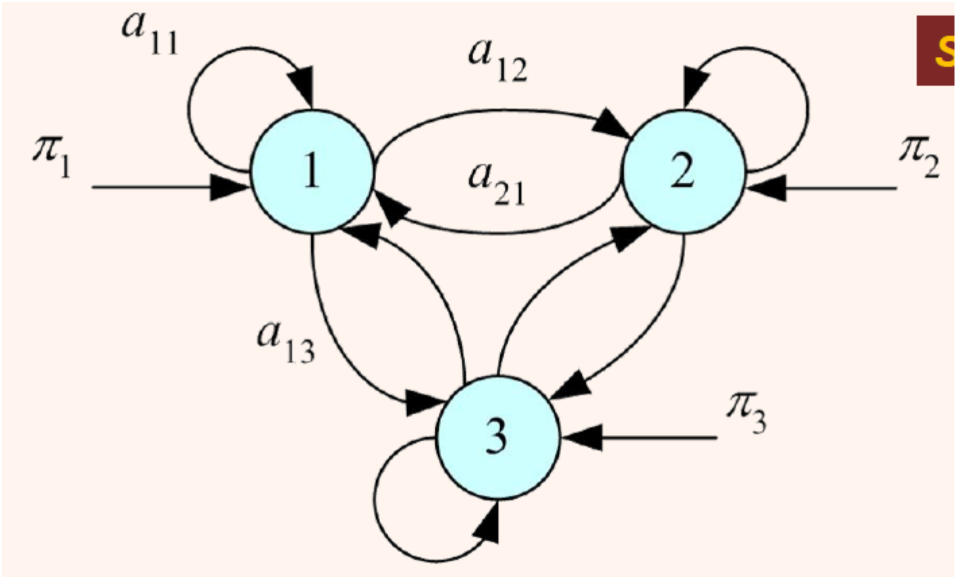
$$\pi_i = P(q_i = S_i) \qquad \sum_{i=1}^{N} \pi_i = 1 \qquad (۶)$$

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۳.۲ فرآیندهای گسسته‌ی مارکوف (ادامه...) (Discrete Markov Process)

Stochastic automaton



- ❑ $\mathbf{A} = [a_{ij}]$ یک ماتریس با ابعاد $N \times N$ است که جمع عناصر هر سطر آن برابر یک می‌شود.
- ❑ $\Pi = [\pi_i]$ برداری Nتایی است که حاصل جمع تمام عناصر آن برابر یک است.

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۳ مدل‌های مارکوف قابل مشاهده

۱.۳ مدل‌های مارکوف قابل مشاهده (Observable Markov Model)

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

گزاره

در یک «مدل مارکوف قابل مشاهده»، در زمان t می‌دانیم که q_t کدام حالت را نشان می‌دهد.

ـ خروجی فرآیند، بر چسب حالت فعلی است؛ هر حالت متناظر با مشاهده‌ی یک رخداد فیزیکی می‌باشد.

نکته

«دنباله‌ی مشاهدات[a]»، O، در اینجا معادل ترتیب حالت‌های مشاهده شده است، که احتمال رخداد آن به صورت زیر محاسبه می‌شود:

$$P(O=Q|A,\Pi)=p(q_1)\prod_{t=2}^{T}p(q_t|q_{t-1})=\pi_{q_1}a_{q_1q_2}\dots a_{q_{T-1}q_T} \quad (۷)$$

[a] Observation sequence

۲.۳ مثال ۱

مثال

- هر حالت بیانگر وضعیت جوی در یک زمان مشخص در روز (مثلا ظهر) می‌باشد:
  - حالت ۱: وجود بارندگی
  - حالت ۲: هوای ابری
  - حالت ۳: هوای آفتابی
- ماتریس انتقال حالت:

$$A = \{a_{ij}\} = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} \quad (۸)$$

- با فرض این که در روز اول هوا آفتابی باشد، احتمال این که هفت روز بعد، آفتابی_ آفتابی_ بارانی_ بارانی_ آفتابی_ ابری_ آفتابی باشد:

$$O = \{S_3, S_3, S_3, S_1, S_1, S_3, S_2, S_3\} \quad (۹)$$

$$\begin{aligned} P(O|Model) =& P\{S_3, S_3, S_3, S_1, S_1, S_3, S_2, S_3\} = P(S_3)P(S_3|S_3)\dots \\ & \dots P(S_3|S_3)P(S_1|S_3)P(S_1|S_1)P(S_3|S_1)P(S_2|S_3)P(S_3|S_2) \\ =& \pi_3 a_{33} a_{33} a_{31} a_{11} a_{13} a_{32} a_{23} \\ =& 1.(0.8).(0.8).(0.1).(0.4).(0.3).(0.1).(0.2) = 1.536 \times 10^{-4} \end{aligned}$$

مثال

- احتمال باقی ماندن مدل در یک حالت به اندازه زمان d:

$$P_i(d) \equiv P(O|Model, q_1 = S_i) = (a_{ii})^{d-1}(1 - a_i) \qquad (۱۰)$$

- به طور متوسط چند روز پیاپی هوا آفتابی است؟

$$\begin{aligned} E[d_i] &= \sum_{d=1}^{\infty} d_{P_i}(d) = \sum_{d=1}^{\infty} d(a_{ii})^{d-1}(1 - a_{ii}) \\ &= (1 - a_{ii}) \sum_{d=1}^{\infty} d(a_{ii})^{d-1} = \frac{1}{1 - a_{ii}} \end{aligned} \qquad (۱۱)$$

به عنوان نمونه در مثال فوق انتظار می‌رود به طور متوسط پنج روز پیاپی هوا آفتابی، ۲/۵ روز ابری و تنها ۱/۶۷ روز متوالی هوا بارانی باشد.

$$A = \{a_{ij}\} = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$

۳.۳ مثال ۲

مثال

- فرض کنید N گلدان در اختیار داریم که در هر یک توپ‌هایی هم رنگ موجود است.
  - قرمز: S_1
  - آبی: S_2
  - سبز: S_3
- q_t رنگ توپی که در زمان t برداشته شده است، را نمایش می‌دهد.

$$\Pi = [0.5, 0.2, 0.3]^T \tag{۱۲}$$

$$A = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix} \tag{۱۳}$$

$$O = \{S_1, S_1, S_3, S_3\} \tag{۱۴}$$

$$\begin{aligned} P(O|A,\Pi) &= P(S_1) \cdot P(S_1|S_1) \cdot P(S_1|S_3) \cdot P(S_3|S_3) \\ &= \pi_1 \cdot a_{11} \cdot a_{13} \cdot a_{33} \\ &= 0.5 \cdot 0.4 \cdot 0.3 \cdot 0.8 = 0.048 \end{aligned} \tag{۱۵}$$

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

مثال

- فرض کنید در این مثال یک سری (K) مشاهده با طول T موجود است، سیستم می‌تواند پارامترهای Π و A را یاد بگیرد. اگر q_t^k حالت سیستم در زمان t در دنباله‌ی kام باشد، احتمال حالت اولیه را می‌توان به صورت زیر تقریب زد:

$$\hat{\pi}_i = \frac{\#\{\text{ sequences starting with } S_i\}}{\text{sequences}} = \frac{\sum_k 1(q_1^k = S_i)}{K} \qquad (۱۶)$$

- و احتمال گذار:

$$\begin{aligned}\hat{\pi}_i =& \frac{\#\{\text{ transitions from } S_i to S_j\}}{\#\{\text{ transitions from } S_i\}} \\ =& \frac{\sum_k \sum_{t=1}^{T-1} 1(q_t^k = S_i \text{ and } q_{t+1}^k = S_j)}{\sum_k \sum_{t=1}^{T-1} 1(q_t^k = S_i)}\end{aligned} \qquad (۱۷)$$

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۴ مدل پنهان مارکوف

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۱.۴ مدل پنهان مارکوف (Hidden Markov Model)

گزاره

مدل مارکوف قابل مشاهده برای استفاده عملی بسیار محدود می‌باشد.

نکته

در «مدل مارکوف پنهان (HMM)»، حالت‌های سیستم را نمی‌توان مشاهده نمود بلکه در هر حالت، خروجی مشاهده شده، احتمال حضور سیستم در یک حالت خاص را با تابعی احتمالاتی بیان می‌کند.

نکته

با فرض اینکه در حالت‌های مختلف خروجی سیستم از مجموعه‌ی زیر می‌باشد:

$$\{v_1, v_2, \ldots, v_M\} \tag{۱۸}$$

«احتمال مشاهده[a]» به صورت زیر به دست می‌آید:

$$b_j(m) \equiv P(O_t = v_m | q_t = S_j) \tag{۱۹}$$

[a] Observation (emission) probability

دنباله‌ی حالت‌های سیستم قابل مشاهده نیست.

- این همان نکته‌ای است که باعث شده است چنین سیستمی **پنهان** نامیده شود.
- ولی با توجه به دنباله‌ی مشاهدات، می‌توان آن را حدس زد و یا به بیان بهتر احتمال آن را محاسبه نمود.
- باید توجه داشت که به ازای هر «دنباله‌ی مشاهدات» تعداد زیادی دنباله‌ی حالت موجود است که می‌تواند همان دنباله‌ی مشاهدات را تولید نماید ولی با **احتمال‌های متفاوت**.
- در مدل پنهان مارکوف علاوه بر حرکت تصادفی بین حالت‌ها، خروجی مشاهده شده هم تصادفی است.
  - مدل ماركوف پنهان در واقع نوعي مدل ماركوف تو در تو است.
  - بدين ترتيب كه مدل ماركوف اصلي انتقال بين حالات را نشان مى‌دهد و در هر حالت، مشاهده با توجه به يك مدل ماركوف وابسته به آن حالت انجام مى‌شود.
- اولین مشکل تعیین تعداد حالات و تخصیص آن به دنباله‌ی مشاهدات است.

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

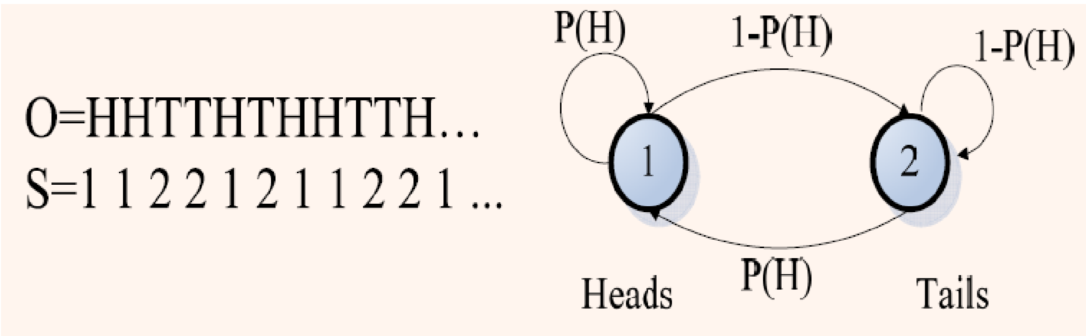
۲.۴ مثال ۳

مثال

- فرض کنید شخصی در پس یک مانع یک (یا چند) سکه را پرتاب می‌کند و بدون این که نحوه‌ی عملکردش معین باشد، تنها نتیجه‌ی پرتاب را نمایش می‌دهد:

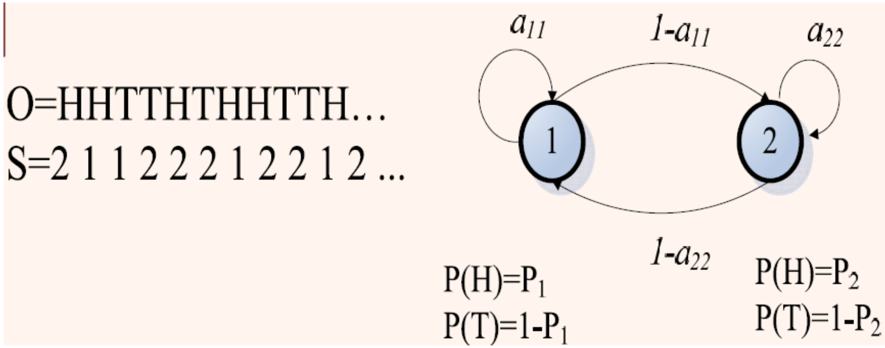
$$\begin{aligned} O &= o_1 o_2 o_3 \dots o_1 \\ &= \mathcal{H}\mathcal{H}\mathcal{T} \dots \mathcal{T}\mathcal{T}\mathcal{H} \end{aligned} \tag{۲۰}$$

- چگونه می‌توان این فرآیند را با زنجیره مارکف مدل کرد؟

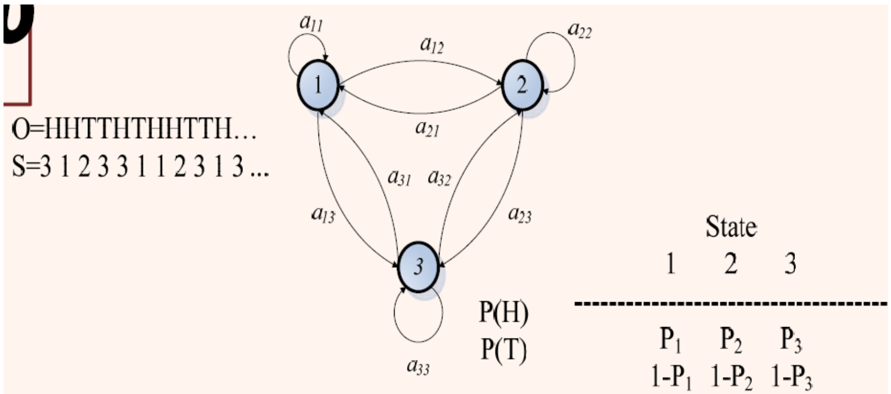


- دو حالت در نظر گرفته و هر حالت بیانگر يك روي سكه باشد.
- تنها پارامتر مجهول p(H) است (البته به جز احتمال اولیه).
- «مدل مارکوف قابل مشاهده» است!

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع



- دو حالت در نظر گرفته و هر حالت بیان گر خروجي يك روي سكه است.
- چهار پارامتر مجهول وجود دارد(البته به جز احتمال اوليه).
- «مدل ماركوف قابل پنهان» است!



- نه پارامتر مجهول وجود دارد(البته به جز احتمال اوليه).
- هر چه مدل بزرگ‌تر شود قابليت مدل كردن آن بهتر خواهد بود.
- اما در عین حال ممکن است مشکل overfitting رخ دهد.

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۳.۴ مثال ۴

مثال

- در مثال توپ و گلدان، مدل مارکوف پنهان معادل حالتي است كه در هر گلدان توپ‌هايي با رنگ‌هاي متفاوت داشته باشيم.
- در اینجا $b_j(m)$ معادل خارج کردن توپی با رنگ m از گلدان jام می‌باشد.
- اين بار نيز دنباله‌اي از رنگ ها موجود است با اين تفاوت كه نمي‌دانيم كه توپ ها متعلق به كدام گلدان هستند.

$$o = \{red, red, green, blue, yellow\}$$

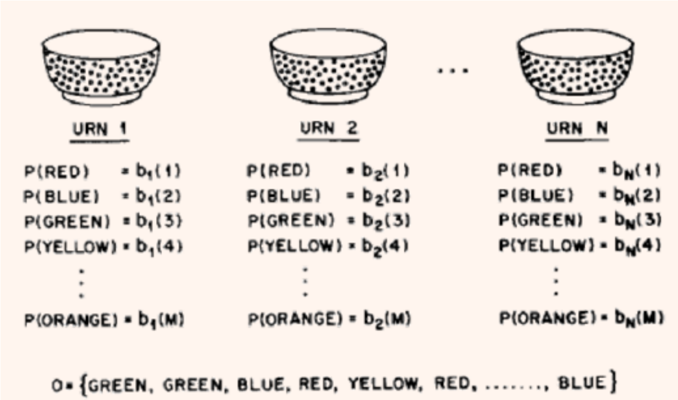
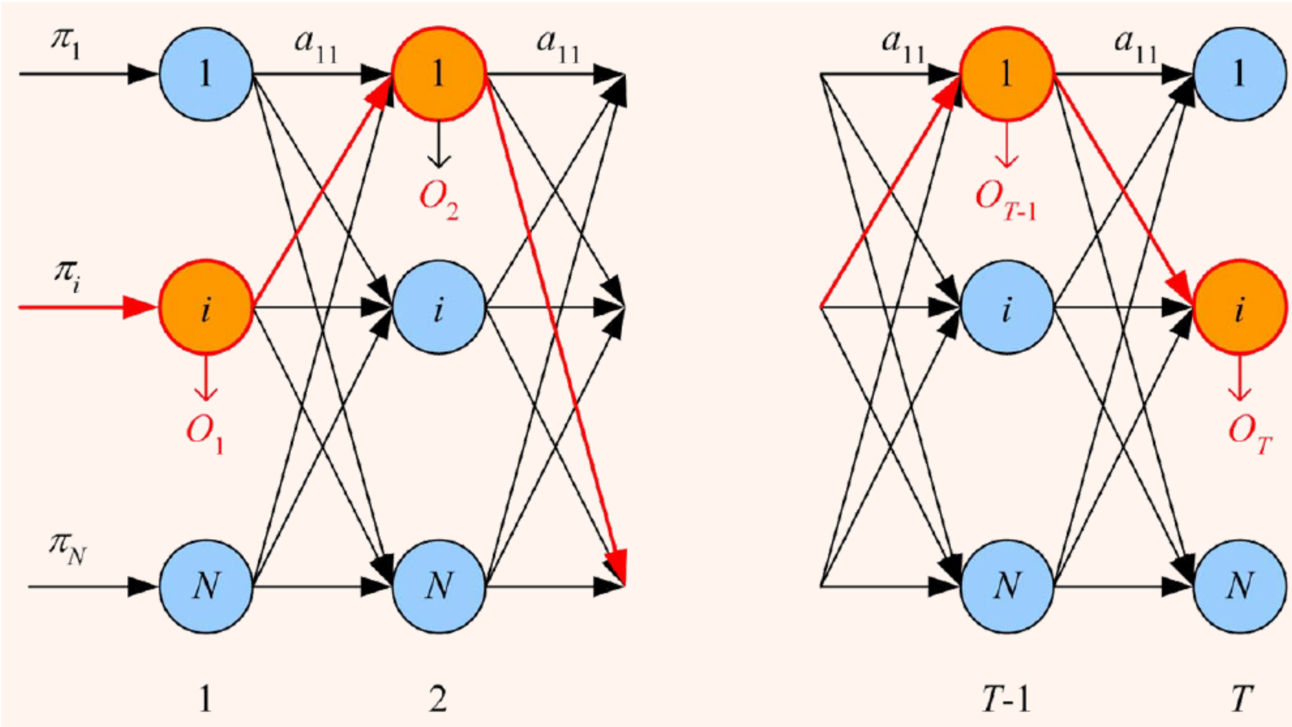


Fig. 3. An N-state urn and ball model which illustrates the general case of a discrete symbol HMM.

Jack Ferguson and his colleagues

HMM Unfolded in Time



مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۴.۴ مؤلفه‌های مدل پنهان مارکوف

تذکر

- تعداد حالت‌ها (N): $S = \{S_1, S_2, \dots S_N\}$
- تعداد نمادهای قابل مشاهده (M): $V = \{v_1, v_2, \dots v_M\}$
- احتمال گذار: $A = [a_{ij}]$ where $a_{ij} \equiv P(q_{t+1} = S_j | q_t = S_i)$
- احتمال مشاهده: $B = [b_j(m)]$ where $b_j(m) \equiv P(O_t = v_m | q_t = S_j)$
- احتمالات حالت اولیه: $\Pi = [\pi_i]$ where $\pi_i \equiv P(q_1 = S_i)$

نتیجه

- بنابراین یك مدل مارکوف پنهان را می‌توان به صورت زیر نشان داد:

$$\lambda = (A, B, \Pi) \tag{۲۱}$$

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۵ مسائل سه‌گانه

سه مسأله‌ی پایه مرتبط با HMM

۱ــ **ارزیابی**[۱]: تعیین **کارآمد** احتمال رخداد یک دنباله از مشاهدات.

$$P(O|\lambda) = ?$$

۲ــ **تعیین محتمل‌ترین حالت یک دنباله از مشاهدات**[۲]

$$P(Q^*|O,\lambda) = \max_Q P(Q|O,\lambda)$$

۳ــ **یادگیری مدل**[۳]

$$\text{Given } X = \{O^k\}_k, \text{ find } \lambda^* \text{ such that } P(X|O,\lambda^*) = \max_\lambda P(X|\lambda)$$

[1] Evaluation
[2] State sequence
[3] Learning

۱.۵ مسأله‌ی ارزیابی $[P(O|\lambda) = ?]$

با این فرض که دنباله‌ي زیر مشاهده شده است:

$$O = \{o_1 o_2 \dots o_T\} \tag{۲۲}$$

اگر دنباله‌ي حالت‌ها، مشخص باشد:

$$O = \{q_1 q_2 \dots q_T\} \tag{۲۳}$$

احتمال رخداد این دنباله از مشاهدات در حالت‌هاي مشخص شده به ترتیب زیر به دست مي‌آید:

$$P(O|Q,\lambda) = \prod_{t=1}^{T} P(o_t|q_t,\lambda) = b_{q_1}(o_1) b_{q_2}(o_2) \dots b_{q_T}(o_T) \tag{۲۴}$$

تذکر

تنها اطلاعاتي که در این حالت داریم، خروجي سیستم است و هیچ اطلاعاتي از حالات سیستم نداریم.

$$P(O|Q,\lambda) = \prod_{t=1}^{T} P(o_t|q_t,\lambda) = b_{q_1}(o_1)b_{q_2}(o_2)\dots b_{q_T}(o_T) \quad (۲۵)$$

تذکر

احتمال رخداد دنباله حالات سیستم:

$$P(Q,\lambda) = P(q_1)\prod_{t=2}^{T} P(q_t|q_{t-1}) = \pi_{q1} a_{q1q2}\dots a_{q_T-1q_T} \quad (۲۶)$$

نتیجه

$$\begin{aligned} P(O|Q,\lambda) =& P(q_1)\prod_{t=2}^{T} P(q_t|q_{t-1})\prod_{t=1}^{T} P(o_t|q_t,\lambda) \\ =& \pi_{q_1} b_{q_1}(o_1) a_{q_1q_2} b_{q_2}(o_2) a_{q_2q_3}\dots a_{q_T-1q_T} b_{q_T}(o_T) \end{aligned} \quad (۲۷)$$

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

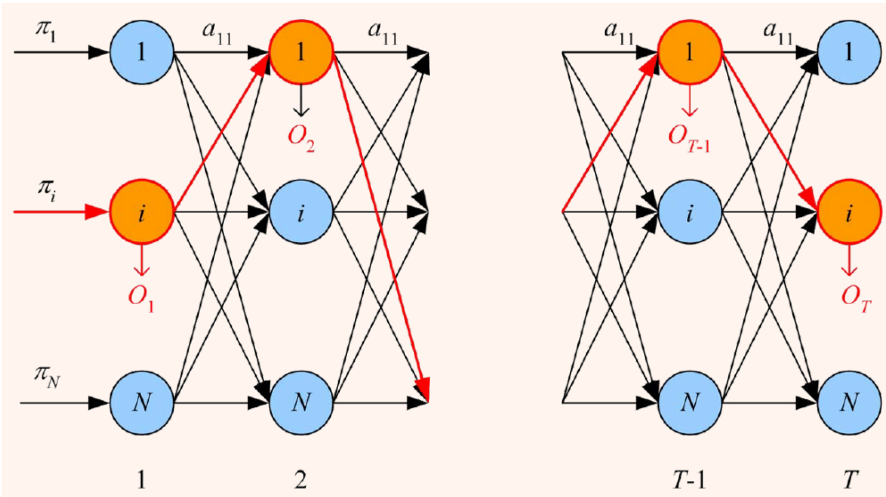
$$P(O|Q,\lambda) = \pi_{q_1} b_{q_1}(o_1) a_{q_1 q_2} b_{q_2}(o_2) a_{q_2 q_3} \ldots a_{q_{T-1} q_T} b_{q_T}(o_T) \qquad (۲۸)$$

با در اختيار داشتن رابطەي فوق و محاسبەي تابع احتمال حاشيه‌اي O خواهيم داشت:

$$P(O|\lambda) = \sum_{\text{all possible } Q} P(O|Q,\lambda) \qquad O(2TN^T) \qquad (۲۹)$$

كه البته اين شيوه عملي نيست!

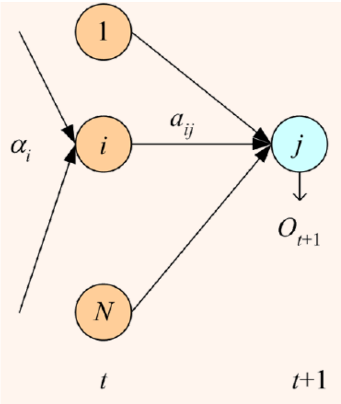
چرا كه N^T شيوه ي مختلف براي حالت ها وجود دارد.



Forward backward Procedure

Forward variable:	$\alpha_t(i) \equiv P(O_1 \dots O_t, q_t = S_i \mid \lambda)$	(۳۰)
Backward variable:	$\beta_t(i) \equiv P(O_{t+1} \dots O_T \mid q_t = S_i, \lambda)$	(۳۱)

- Initialization: $\alpha_1(i) = \pi_i b_i(O_1)$
- Induction (Recursion): $\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1})$
- Termination: $P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$



$$O(N^2T)$$

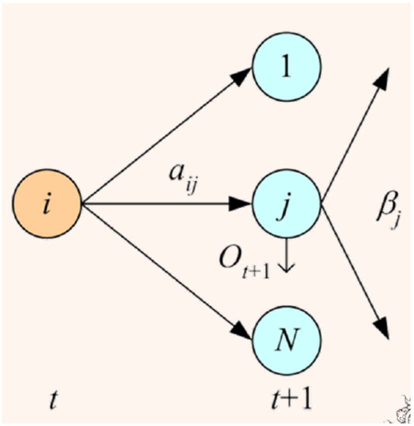
مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

ریز محاسبات (Forward backward Procedure)

$$
\begin{aligned}
\alpha_{t+1}(j) =& P(o_1 \ldots o_{t+1}, q_{t+1} = S_j|\lambda) \\
=& P(o_1 \ldots o_{t+1}|q_{t+1} = S_j, \lambda)P(q_{t+1} = S_j|\lambda) \\
=& P(o_1 \ldots o_t|q_{t+1} = S_j, \lambda)P(o_{t+1}|q_{t+1} = S_j, \lambda)P(q_{t+1} = S_j|\lambda) \\
=& P(o_1 \ldots o_t, q_{t+1} = S_j|\lambda)P(o_{t+1}|q_{t+1} = S_j, \lambda) \\
=& P(o_{t+1}|q_{t+1} = S_j, \lambda)\sum_i P(o_1 \ldots o_t, q_t = S_i, q_{t+1} = S_j|\lambda) \\
=& P(o_{t+1}|q_{t+1} = S_j, \lambda)\sum_i P(o_1 \ldots o_t, q_{t+1} = S_j|q_t = S_i, \lambda)P(q_t = S_i|\lambda) \\
=& P(o_{t+1}|q_{t+1} = S_j, \lambda)\sum_i P(o_1 \ldots o_t, q_t = S_i, \lambda)P(q_{t+1} = S_j|q_t = S_i, \lambda) \\
& P(q_t = S_i|\lambda) \\
=& P(o_{t+1}|q_{t+1} = S_j, \lambda)\sum_i P(o_1 \ldots o_t, q_t = S_i|\lambda)P(q_{t+1} = S_j|q_t = S_i, \lambda) \\
=& P(o_{t+1}|q_{t+1} = S_j, \lambda)\sum_i \alpha_t(i)P(q_{t+1} = S_j|q_t = S_i, \lambda) \\
=& \left[\sum_{i=1}^{N} \alpha_t(i)a_{ij}\right] b_j(O_{t+1})
\end{aligned}
$$

Backward variable: $\beta_t(i) \equiv P(O_{t+1} \dots O_T | q_t = S_i, \lambda)$ (۳۲)

- Initialization: $\beta_T(i) = 1$
- Recursion: $\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$



نکته: هر دو متغیر محصول ضرب اعداد بسیار کوچکي هستند و امکان پاریز آن‌ها وجود دارد، براي پرهیز از این مشکل توصیه میشود در هر مرحله نتایج نرمال شوند. $c_t = \sum_j \alpha_t(j)$

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

ریز محاسبات (Backward variable)

$$
\begin{aligned}
\beta_t(i) \equiv & P(o_{t+1}\dots o_T|q_t = S_i, \lambda)\\
= & \sum_j P(o_{t+1}\dots o_T, q_{t+1} = S_j|q_t = S_i, \lambda)\\
= & \sum_j P(o_{t+1}\dots o_T|q_{t+1} = S_j, q_t = S_i, \lambda)P(q_{t+1} = S_j|q_t = S_i, \lambda)\\
= & \sum_j P(o_{t+1}|q_{t+1} = S_j, q_t = S_i, \lambda)\\
& \quad P(o_{t+2}\dots o_T|q_{t+1} = S_j, q_t = S_i, \lambda)P(q_{t+1} = S_j|q_t = S_i, \lambda)\\
= & \sum_j P(o_{t+1}|q_{t+1} = S_j, \lambda)\\
& \quad P(o_{t+2}\dots o_T|q_{t+1} = S_j, \lambda)P(q_{t+1} = S_j|q_t = S_i, \lambda)\\
= & \sum_{j=1}^{N} a_{ij} b_j(o_{t+1})\beta_{t+1}(j)
\end{aligned}
$$

مقدمه

زنجیره‌ی مارکوف

مدل‌های مارکوف قابل مشاهده

مدل پنهان مارکوف

مسائل سه‌گانه

سایر منابع

۲.۵ یافتن دنباله حالات $P(Q^*|O,\lambda) = \mathbf{max}_Q P(Q|O,\lambda)$

- با در اختیار داشتن خصوصیات یك مدل مارکوف پنهان λ و یک دنباله از مشاهدات:

$$O = \{o_1 o_2 \dots o_T\} \tag{۳۳}$$

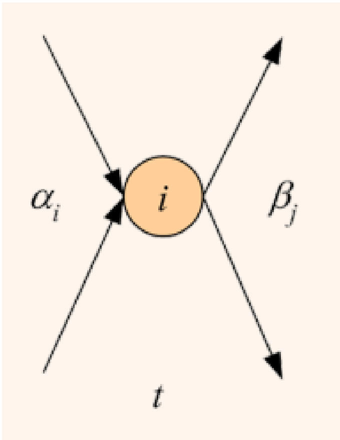
- در پي دنبالهاي از حالتها هستیم که با بیشترین احتمال دنبالهي مشاهدات مورد نظر را تولید کند:

$$O = \{q_1 q_2 \dots q_T\} \tag{۳۴}$$

- یک راه محاسبهی تمام حالات ممکن و انتخاب مسیر با بیشترین احتمال است!

مقدمه
زنجیرهی مارکوف
مدلهای مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سهگانه
سایر منابع

$$\begin{aligned}\gamma_t(i) \equiv & P(q_t = S_i|O,\lambda)\\ = & \frac{P(O|q_t = S_i,\lambda)P(q_t = S_i|\lambda)}{P((O|\lambda))}\\ & \vdots \\ = & \frac{\alpha_t(i)\beta_t(i)}{\sum_{j=1}^{N}\alpha_t(j)\beta_t(j)}\end{aligned}$$



- در هر گام (t) حالتي انتخاب مي‌شود كه بيشترين احتمال را داشته باشد.

$$q_t^* = \arg\max_i Y_t(i)$$

ریز محاسبات

$$\begin{aligned}
\gamma_t(i) \equiv & P(q_t = S_i|O,\lambda) \qquad \text{Bayse theore} \\
= & \frac{P(O|q_t = S_i,\lambda)P(q_t = S_i|\lambda)}{P(O|\lambda)} \\
= & \frac{P(o_1 \dots o_t|q_t = S_i,\lambda)P(o_{t+1} \dots o_T|q_t = S_i,\lambda)P(q_t = S_i|\lambda)}{\sum_{j=1}^{N} P(O|q_t = S_j,\lambda)P(q_t = S_j)} \\
= & \frac{\alpha_t(i)\beta_t(i)}{\sum_{j=1}^{N} \alpha_t(j)\beta_t(j)}
\end{aligned}$$

مقدمه

زنجیره‌ی مارکوف

مدل‌های مارکوف قابل مشاهده

مدل پنهان مارکوف

مسائل سه‌گانه

سایر منابع

Viterbi's Algorithm

$$\delta_t(i) = \max_{q_1 q_2 \dots q_{t-1}} P(q_1 q_2 \dots q_{t-1}, q_t = S_i, o_1 \dots o_t|\lambda) \quad (۳۵)$$

- مقداردهی اولیه[۴]

$$\delta_1(i) = \pi_i b_i(o_1) \quad (۳۶)$$

$$\psi_1(i) = 0 \quad (۳۷)$$

- بازگشتی[۵]

$$\delta_t(j) = \max_i [\delta_{t-1}(i) a_{ij}] \cdot b_j(o_t) \quad (۳۸)$$

$$\psi_t(i) = \arg\max_i [\delta_{t-1}(i) a_{ij}] \quad (۳۹)$$

- خاتمه[۶]

$$P^* = \max_i \delta_T(i) \quad (۴۰)$$

$$q_T^* = \arg\max_i \delta_T(i) \quad (۴۱)$$

- برگشت مسیر: (دنباله‌ي حالت‌ها) [۷]

$$q_T^* = \psi_{t+1}(q_{t+1}^*), \quad t = T-1, T-2, \dots, 1 \quad (۴۲)$$

[4] initialization
[5] induction
[6] termination
[7] Path backtrackin

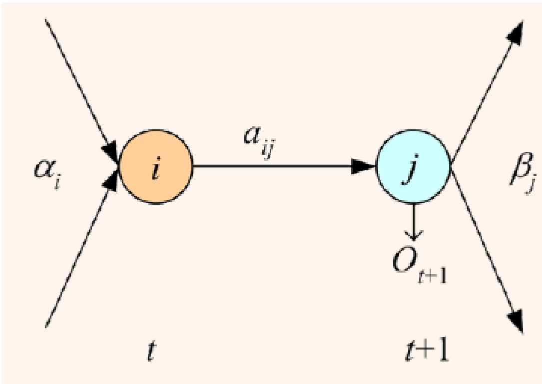
۳.۵ یادگیری (Learning)

- هدف این است که با در اختیار داشتن یك مجموعه‌ي آموزشي از مشاهدات $x = \{O^k\}_{k=1}^{K}$ پارامترهای مدل $\lambda^* = (A, B, \Pi)$ به گونه‌ای برآورد شوند که تابع درست نمایی $P(x|\lambda^*)$ بیشینه شود.

- راه حل تحليلي براي اين مساله وجود ندارد.

  – از یک فرآیند تکرار شونده استفاده می‌شود:

$$\text{Baum-Welch algorithm:} \qquad \xi_t(i,j) \equiv P(q_t = S_i, q_{t+1} = S_j | O, \lambda) \qquad (۴۳)$$



ریز محاسبات

$$
\begin{aligned}
\xi_t(i,j) \equiv & P(q_t = S_i, q_{t+1} = S_j | O, \lambda) \\
= & \frac{P(O|q_t = S_i, q_{t+1} = S_j, \lambda) P(q_t = S_i, q_{t+1} = S_j|\lambda)}{P(O|\lambda)} \\
= & \frac{P(O|q_t = S_i, q_{t+1} = S_j, \lambda) P(q_{t+1} = S_j|q_t = S_i, \lambda) P(q_t = S_i|\lambda)}{P(O|\lambda)} \\
= & \frac{1}{P(O|\lambda)} P(o_1 \dots o_t|q_t = S_i, \lambda) P(o_{t+1}|q_{t+1} = S_j, \lambda) \\
& P(o_{t+2} \dots o_T|q_{t+1} = S_j, \lambda) a_{ij} P(q_t = S_i|\lambda) \\
= & \frac{1}{P(O|\lambda)} P(o_1 \dots o_t|q_t = S_i, \lambda) P(o_{t+1}|q_{t+1} = S_j, \lambda) \\
& P(o_{t+2} \dots o_T|q_{t+1} = S_j, \lambda) a_{ij} \\
= & \frac{\alpha_t(i) b_j(o_{t+1}) \beta_{t+1}(j) a_{ij}}{\sum_k \sum_l \alpha_t(k) a_{kl} b_l(o_{t+1}) \beta_{t+1}(l)} \\
= & \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_k \sum_l \alpha_t(k) a_{kl} b_l(o_{t+1}) \beta_{t+1}(l)}
\end{aligned}
$$

مقدمه

زنجیره‌ی مارکوف

مدل‌های مارکوف قابل مشاهده

مدل پنهان مارکوف

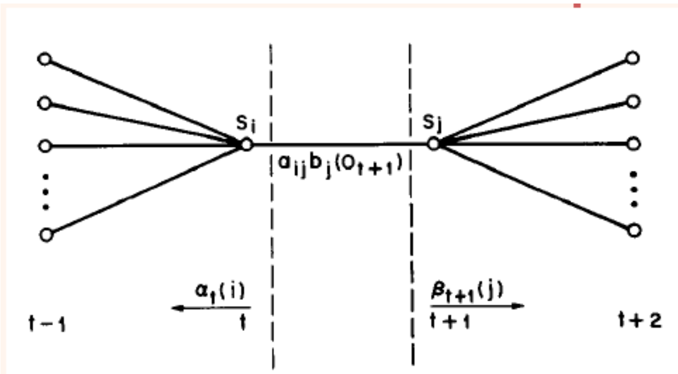
مسائل سه‌گانه

سایر منابع

Baum-Welch algorithm

$$\xi_t(i,j) \equiv P(q_t = S_i, q_{t+1} = S_j | O, \lambda) \quad (۴۴)$$

$$= \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_k \sum_l \alpha_t(k) a_{kl} b_l(o_{t+1}) \beta_{t+1}(l)} \quad (۴۵)$$



- مي‌توان احتمال حضور در يك حالت را محاسبه كرد:

$$\gamma_t(i) \equiv P(q_t = S_i | O, \lambda) \quad (۴۶)$$

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j) \quad (۴۷)$$

- در صورتي كه مدل ماركوف قابل مشاهده باشد، هر كدام از مقادير γ و ξ صفر و يک خواهند بود.

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

- $$\hat{\pi}_i = \frac{\#\{\text{ sequences starting with } S_i\}}{\text{sequences}} = \frac{\sum_k 1(q_1^k = S_i)}{K} \quad (۴۸)$$

- $$\hat{\pi}_i = \frac{\#\{\text{ transitions from } S_i to S_j\}}{\#\{\text{ transitions from } S_i\}} = \frac{\sum_k \sum_{t=1}^{T-1} 1(q_t^k = S_i \text{ and } q_{t+1}^k = S_j)}{\sum_k \sum_{t=1}^{T-1} 1(q_t^k = S_i)} \quad (۴۹)$$

- $$\gamma_t(i) \equiv P(q_t = S_i | O, \lambda) \quad (۵۰)$$

- $$\xi_t(i,j) \equiv P(q_t = S_i, q_{t+1} = S_j | O, \lambda) \quad (۵۱)$$

- $$z_i^t = \begin{cases} 1 & \text{if} \quad q_t = S_i \\ 0 & \text{otherwise} \end{cases} \qquad z_{ij}^t = \begin{cases} 1 & \text{if} \quad q_t = S_i \quad \text{and} \quad q_{t+1} = S_j \\ 0 & \text{otherwise} \end{cases} \quad (۵۲)$$

- $$E[z_i^t] = \gamma_t(i) \qquad E[z_{ij}^t] = \xi_t(i,j) \quad (۵۳)$$

نکته

- در گام ۱ (E-Step)، با پارامترهاي با مقدار فعلي پارامترهاي مدل مقادير γ و ξ تخمين زده می‌شوند.
- (M-Step)، بر اساس تخمين زده شده، پارامترهاي مدل به روز مي‌شوند.

$$\hat{\pi}_i = \frac{\sum_{k=1}^{K} \gamma_1^k(i)}{K} \qquad (۵۴)$$

$$\hat{a}_{ij} = \frac{\sum_{k=1}^{K} \sum_{t=1}^{T_{k-1}} \xi_t^k(i,j)}{\sum_{k=1}^{K} \sum_{t=1}^{T_{k-1}} \gamma_t^k(i)} \qquad (۵۵)$$

$$\hat{b}_j(m) = \frac{\sum_{k=1}^{K} \sum_{t=1}^{T_{k-1}} \gamma_t^k(j)(O_t^k = v_m)}{\sum_{k=1}^{K} \sum_{t=1}^{T_{k-1}} \gamma_t^k(i)} \qquad (۵۶)$$

نکته

- اين روند تا همگرايي ادامه خواهد يافت، ثابت شده است که $P(O|\lambda)$ نزولی خواهد شد.

نکته

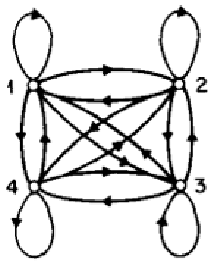
- اين الگوريتم، ماكزيمم محلي را مي‌يابد و در عمل رويه هدف (maximization surface) شکل پیچیده‌ای داردو دارای تعداد زیادی ماکزیمم محلی است.

نکته

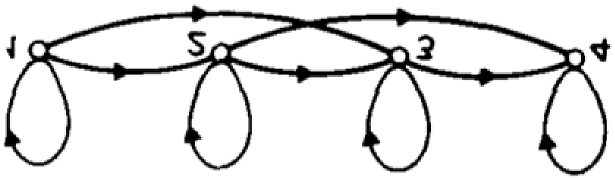
- نظر به اين كه كليت مسأله‌ي آموزش به نوعي يك مسأله‌ي بهينه‌سازي است و از تكنيك‌هاي نظير نزول گراديان براي حل اين مسأله مي‌توان بهره جست.

Model Selection in HMM

- در برخي کاربردها مانند تشخيص گفتار استفاده از مدل‌هاي خاصي توصيه ميشود.



Ergodic model



Bakis

$$a_{ij} = 0 \quad i < j$$
$$a_{ij} = 0 \quad i < j + \Delta$$

$$\pi_i = \begin{cases} 0, & \text{if} \quad i \neq 1 \\ 1, & \text{if} \quad i = 1 \end{cases}$$

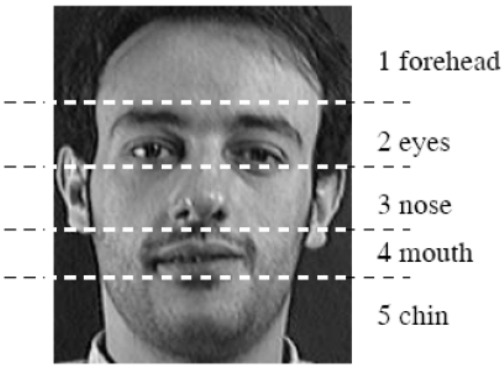
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

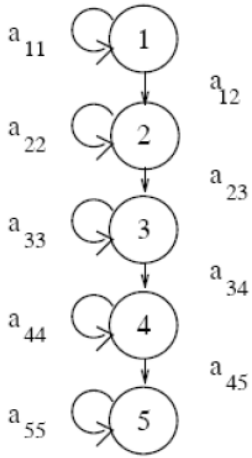
۴.۵ دسته‌بندی

- يك مجموعه از HMMها خواهيم داشت كه هر يك، دنباله‌هاي مربوط به يك دسته را مدل مي‌كنند.
  – مثلاً در بازشناخت كلمات ادا شده به ازاي هر كلمه، يك HMM جداگانه آموزش داده مي‌شود.
  – با ارائه‌ي يك كلمه‌ي جديد براي شناسايي، تمام مدل‌هاي موجود مورد ارزيابي قرار مي‌گيرند و مقدار محاسبه مي‌شود. سپس با استفاده از قانون بيز خواهيم داشت:

$$P(\lambda_i|O) = \frac{P(O|\lambda_i)P(\lambda_i)}{\sum_j P(O|\lambda_j)P(\lambda_j)}$$

  – مدلي كه دراي بيشترين احتمال $P(\lambda_i|O)$ باشد به عنوان دسته‌ی شناسایی شده معرفی می‌گردد.
- مثال_ شناسایی چهره





Ferdinando Silvestro Samaria, “Face Recognition Using Hidden Markov Model”, 1994

۵.۵ بخش پایانی

همملرن یک کتابخانه پایتون است که مدل های مارکوف پنهان را در پایتون پیاده سازی می کند همملرن سه مدل را ارائه می دهد _ یک مدل انتشار چند جمله ای، یک مدل انتشار گاوسی و یک مدل انتشار مخلوط گاوسی، این کتابخانه امکان اجرای مدل های سفارشی را فراهم می کند.

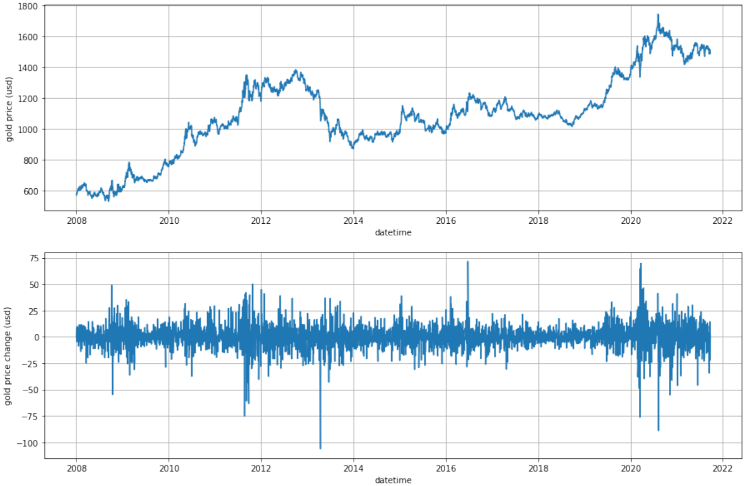
pip install hmmlearn==0.2.1

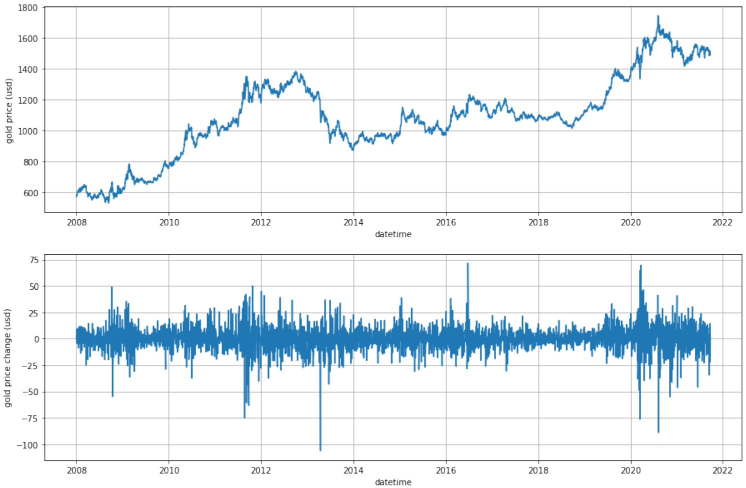
در ابتدا کتابخانه های لازم و همچنین داده ها را به پایتون وارد می کنیم و داده های با کمکن کتابخانه متپلات لیب مصور می کنیم. از آنجا که برخورد با تغییر قیمت به جای خود قیمت واقعی منجر به مدل سازی بهتر شرایط واقعی بازار می شود درنتیجه تغییرات روزانه قیمت طلا را محاسبه کرده و داده ها را از سال ۲۰۰۸ به بعد محدود کرده ایم.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from hmmlearn import hmm
base_dir = "https://github.com/natsunoyuki/Data_Science/blob/master
/gold/gold/gold_price_usd.csv?raw=True"
data = pd.read_csv(base_dir)
# Convert the datetime from str to datetime object.
data["datetime"] = pd.to_datetime(data["datetime"])
# Determine the daily change in gold price.
data["gold_price_change"] = data["gold_price_usd"].diff()
```

```
# Restrict the data to later than 2008 Jan 01.
data = data[data["datetime"] >= pd.to_datetime("2008-01-01")]
# Plot the daily gold prices as well as the daily change.
plt.figure(figsize = (15, 10))
plt.subplot(2,1,1)
plt.plot(data["datetime"], data["gold_price_usd"])
plt.xlabel("datetime")
plt.ylabel("gold price (usd)")
plt.grid(True)
plt.subplot(2,1,2)
plt.plot(data["datetime"], data["gold_price_change"])
plt.xlabel("datetime")
plt.ylabel("gold price change (usd)")
plt.grid(True)
plt.show()
data = pd.read_csv(base_dir)
# Convert the datetime from str to datetime object.
data["datetime"] = pd.to_datetime(data["datetime"])
# Determine the daily change in gold price.
data["gold_price_change"] = data["gold_price_usd"].diff()
# Restrict the data to later than 2008 Jan 01.
data = data[data["datetime"] >= pd.to_datetime("2008-01-01")]
```

```
plt.figure(figsize = (15, 10))
plt.subplot(2,1,1)
plt.plot(data["datetime"], data["gold_price_usd"])
plt.xlabel("datetime")
plt.ylabel("gold price (usd)")
plt.grid(True)
plt.subplot(2,1,2)
plt.plot(data["datetime"], data["gold_price_change"])
plt.xlabel("datetime")
plt.ylabel("gold price change (usd)")
plt.grid(True)
plt.show()
```





مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

به‌جای مدل‌سازی مستقیم قیمت طلا، تغییرات روزانه قیمت طلا را مدل‌سازی می‌کنیم این مسئله به ما امکان می‌دهد وضعیت بازار را بهتر درک کنیم . تغییرات روزانه قیمت طلا را با مدل گاوسی با حالت پنهان مطابقت دادیم. دلیل استفاده از ۳ حالت پنهان این است که ما حداقل ۳ روند مختلف را در تغییرات روزانه انتظار داریم (نوسان کم، متوسط و زیاد).

```
# Use the daily change in gold price as the observed measurements X.
X = data[["gold_price_change"]].values# Build the HMM model and fit
    to the gold price change data.
model = hmm.GaussianHMM(n_components = 3, covariance_type = "diag",
        n_iter = 50, random_state = 42)
model.fit(X)# Predict the hidden states corresponding to observed X.
Z = model.predict(X)
states = pd.unique(Z)
```

```
print("Unique states:")
print(states)

Unique states:
[0 1 2]
```

```
print("\nStart probabilities:")
print(model.startprob_)

Start probabilities:
[1.00000000e+00 4.28952054e-24 1.06227453e-46]
```

```
print("\nTransition matrix:")
print(model.transmat_)

Transition matrix:
[[8.56499275e-01 1.42858023e-01 6.42701428e-04]
 [2.43257082e-01 7.02528333e-01 5.42145847e-02]
 [1.33435298e-03 1.67318160e-01 8.31347487e-01]]
```

```
print("\nGaussian distribution means:")
print(model.means_)

Gaussian distribution means:
[[0.27988823]
 [0.2153654 ]
 [0.26501033]]
```

```
print("\nGaussian distribution covariances:")
print(model.covars_)

Gaussian distribution covariances:
[[[ 33.89296208]]
 [[142.59176749]]
 [[518.65294334]]]
```

```
plt.figure(figsize = (15, 10))
plt.subplot(2,1,1)
for i in states:
    want = (Z == i)
    x = data["datetime"].iloc[want]
    y = data["gold_price_usd"].iloc[want]
    plt.plot(x, y, '.')
plt.legend(states, fontsize=16)
plt.grid(True)
plt.xlabel("datetime", fontsize=16)
plt.ylabel("gold price (usd)", fontsize=16)
plt.subplot(2,1,2)
for i in states:
    want = (Z == i)
    x = data["datetime"].iloc[want]
    y = data["gold_price_change"].iloc[want]
    plt.plot(x, y, '.')
plt.legend(states, fontsize=16)
plt.grid(True)
plt.xlabel("datetime", fontsize=16)
plt.ylabel("gold price change (usd)", fontsize=16)
plt.show()
```

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

مقدمه
زنجیره‌ی مارکوف
مدل‌های مارکوف قابل مشاهده
مدل پنهان مارکوف
مسائل سه‌گانه
سایر منابع

۶ سایر منابع

مراجع

[1] abiner, L.R. (2011). ”A tutorial on hidden Markov models and selected applications in speech recognition.” Proceedings of the IEEE 77(2): 257-286.

مقدمه

زنجیره‌ی مارکوف

مدل‌های مارکوف قابل مشاهده

مدل پنهان مارکوف

مسائل سه‌گانه

سایر منابع

با تشکر
از توجه شما