

قرائینهای تعدادی: جلسه 21 و 25 اردیبهشت 1400

* مطالب این جلسه:

• حالت پایداری زنجیرههای مارکوف متناهی حالت

• چند مثال

• شروع بحث تئوری تقنین

• مقدمه

• اصل کفایت (The sufficiency principle)

آمار کافی (Sufficient Statistics)

• تقنین factorization

• برای زنجیره مارکوف متناهی حالت با تعداد حالت $M > 2$ کلیت اصول تحلیل شبیه مباحثی است که در جلسه قبل مطرح شد و جزئیات بیشتری دارد که به فصل 4 کتاب Gallager ارجاع می دهیم.

در مقایسه می توانیم قضیه زیر را بیان کنیم :

• قضیه: ماتریس گذار یک unichain متناهی حالت یک مقدار ویژه $\lambda = 1$

به همراه یک بردار ویژه هست $\underline{e}$ که $\underline{e} \gg 0$ و $\underline{e} = \sum \underline{e}_i$ و یک بردار

ویژه هست راست $\underline{e} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ دارد.

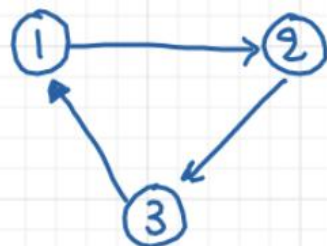
برای مقادیر ویژه دیگر λ_i داریم $|\lambda_i| < 1$. این خاصیتی است که است مگر اینکه

unichain متناهی باشد. اگر دوره تناوبی unichain d باشد، آنگاه

d مقدار ویژه با اندازه 1 به صورت متناوب حول دایره واحد توزیع شده است.

اگر unichain اگر یک باشد، آنگاه $\underline{p}^n$ به حالت پایدار $\underline{e} \cdot \underline{p}$ میل می کند.

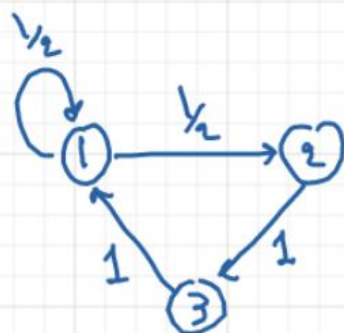
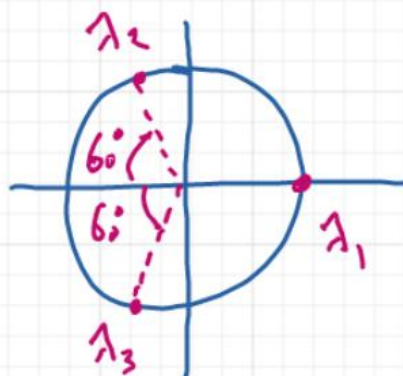
زنجیره مارکوف دلخواه می توان به کلاس های بازگشتی اشی تقسیم کرد و این قضیه
مابقی هم کلاس به طور جداگانه اعمال کرد.



$$\underline{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

مثال:

$$\Rightarrow \det(\underline{P} - \lambda \underline{I}) = 0 \Rightarrow \lambda_1 = 1, \lambda_2 = -\frac{1}{2} + \frac{\sqrt{3}}{2}j, \lambda_3 = -\frac{1}{2} - \frac{\sqrt{3}}{2}j$$



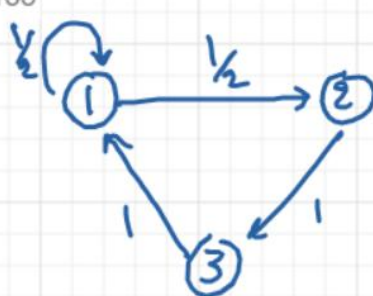
$$\underline{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow d=1 \Rightarrow \text{رتبه اول یک}$$

مثال:

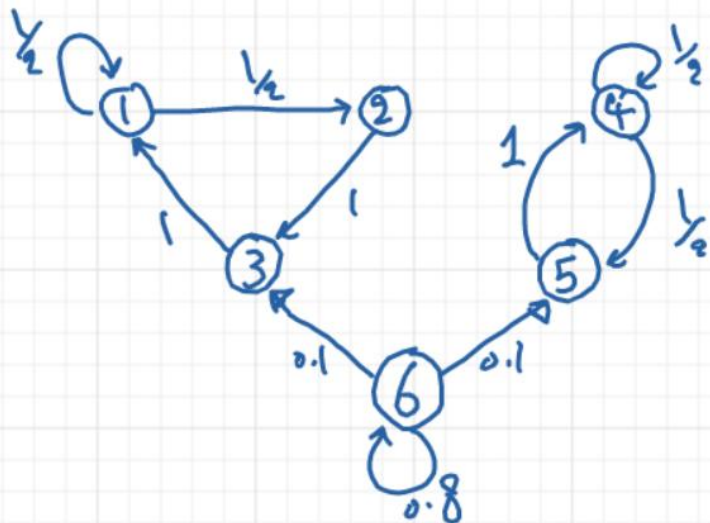
$$\lambda_1 = 1, \lambda_{2,3} = -\frac{1}{4} \pm 0.6614j$$

$$\underline{z} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$



مثال:

$$P = \begin{bmatrix} \boxed{\begin{matrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \boxed{\begin{matrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{matrix}} \end{bmatrix}$$



مثال:

• تئوری تعین:

فرض $X_1, \dots, X_n$: مشاهده مستقل از هم از یک آزمایشی متناهی هستند.

$$X_i \sim f(x|\theta)$$

$$= f(x; \theta)$$

↔ پای توزیع مشترک $X_1, \dots, X_n$:

$$f_{\underline{X}}(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f(x_i | \theta)$$

هدف از روی مشاهدات $X_1, \dots, X_n$ می‌توانیم حرفی در مورد θ بزنیم.
پارامتر نامعلوم

• دو دیدگاه برای تعین θ :

Bayesian point of view : فرضی این است که θ یک متغیر متناهی است

و ما از قبل یک توزیع پای θ فرض می‌کنیم.

Prior

• دیدگاه frequentist ها :

① یک پارامتر نامعلوم ولی غیرمتغیری.

• در هر کدام از این دیدگاه‌ها یک تنبلی برای ② بر حسب مشاهدات پرستی آید.

• تعریف آمار (Statistics) :

فرض کنید $x_1, \dots, x_n$ یک مجموعه مشاهدات باشد - تابع $T(x_1, \dots, x_n)$

$$T: \mathcal{X}^n \rightarrow \mathbb{R}^k \quad \text{به طوریکه:}$$

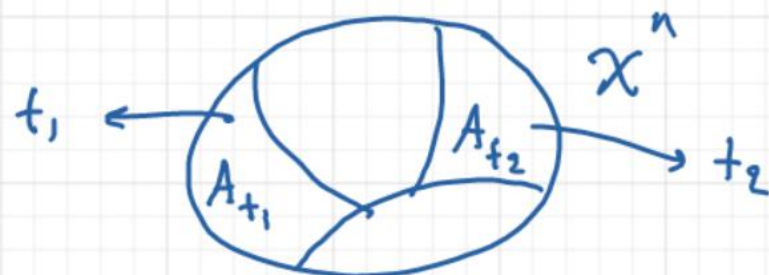
یک آمار نامیده می‌شود (دقت کنید T یک متغیر تصادفی است).

T : خلاصه کردن داده به data summarization.

• آمار T داده را به نحوی خلاصه می‌کند که برای هر $\underline{x}$ که $T(\underline{x}) = T(\underline{y})$ باشد

کسی که به جای مشاهده اصل داده با مقدار آمار T را مشاهده کند فرقی بین $\underline{x}$ و $\underline{y}$ قایل نمی‌شود

به طور دقیق تر:



$$A_t = \{ \underline{x} : T(\underline{x}) = t \}$$

• اصل کفایت (Sufficiency Principle):

اگر $T(\underline{X})$ یک آمار کافی برای θ باشد، هر نتیجه‌ای در مورد θ باید

مقتضای طریق $T(\underline{X})$ به دست آمده $\underline{x}$ باشد و رابطه داشته باشد.

برای دو داده $\underline{x}, \underline{y}$ که $T(\underline{x}) = T(\underline{y})$ نتیجه‌ای که درباره

θ در این دو حالت می‌گیریم باید یکسان باشد.

• آمار کافی (Sufficient Statistic):

آمار $T(X)$ یک آمار کافی برای θ است اگر که (برای متغیرهای گسسته):

$$P[X=x | T(X)=t; \theta] = P[X=x | T(X)=t]$$

مستقل از θ :

$$X \longleftrightarrow T(X) \longleftrightarrow \theta$$

ترتیب مارکف را داریم:

در حالت کلی ترتیب را داریم (حتی برای آمارهای که آمار کافی نیستند):

$$\theta \longleftrightarrow X \longleftrightarrow T(X)$$

data processing Ineq. : $I(\theta; X) \geq I(\theta; T(X))$

dpI

$$I(\theta; X) \leq I(\theta; T(X))$$

در نتیجه: اگر دیدگاه میزبان را استقایی که او باکیم:

$$I(\theta; X) = I(\theta; T(X))$$

برای یک آمار کافی داریم:

برای بررسی اینکه $T(X)$ آیا از θ آزاد است یا نه باید بررسی کنیم

به سبب دارد یا نه $P[X=x | T(X)=t; \theta]$:

توجه: اگر $T(X) \neq t \Leftarrow$ این احتمال صفر است.

پس باید احتمال زیر را بررسی کنیم:

مستقل از θ هست یا نه

$P[X=x | T(X)=T(x); \theta]$:

$$= \frac{P[X=x, T(X)=T(x) | \theta]}{P[T(X)=T(x) | \theta]}$$

توجه می کنیم: $\{X=x\} \subseteq \{T(X)=T(x)\}$

$$= \frac{\overbrace{P[X=x | \theta]}^{= P(x|\theta)}}{P[T(X)=T(x) | \theta]} = \frac{P(x|\theta)}{Q(T(x)|\theta)}$$

تابع چگالی احتمال مشترک
 بردار مشاهده X
 تابع چگالی احتمال $T(X)$

↓
 باید بررسی کنیم مستقل از θ هست یا نه.

معنی: اگر $P(x|0)$ تابع گسلی احتمال یا تابع حری احتمال مشرک x و

$Q(+1|0)$ هم تابع گسلی احتمال یا تابع حری احتمال $T(x)$ باشد، آنگاه

$$\frac{P(x|0)}{Q(T(x)|0)} \Leftrightarrow \begin{matrix} T(x) \text{ یک تابع کافی برای} \\ \text{است} \end{matrix} \begin{matrix} x \in x^n \\ \text{مستقل از } 0 \text{ باشد} \end{matrix}$$