

خرابیه های تصادفی : جلسه 6 ۴۵ ۱۰ خرداد ۱۴۰۱

* مطالب این جلسه

- مثال هایی برای روش گشتاورها
- تقنینه بیشینه likelihood
- تقنینه های بیز

• مسئله : مشاهده پیرامون مدل

$$f(\underline{x} | \theta_1, \dots, \theta_k) = \prod_{i=1}^n f(x_i | \theta_1, \dots, \theta_k)$$

• بر روی روش گشتاورها :

k گشتاور مدل را محاسبه کنیم :

$$\begin{cases} \mu_1 = \mathbb{E}[X] \text{ و } \dots \mu_k = \mathbb{E}[X^k] \\ m_1 = \frac{1}{n} \sum_{i=1}^n X_i \text{ و } \dots m_k = \frac{1}{n} \sum_{i=1}^n X_i^k \end{cases}$$

$$\begin{cases} \mu_1(\theta_1, \dots, \theta_k) = m_1 \\ \mu_k(\theta_1, \dots, \theta_k) = m_k \end{cases} \quad : \quad \text{درگاه کسار و جیو}$$

مثال: فرض کنید $X_1, \dots, X_n$ مشاهدات iid از توزیع گاوسی $\mathcal{N}(\mu, \sigma^2)$ باشند.

$$\underline{\theta} = (\theta_1, \theta_2) = (\mu, \sigma^2)$$

$$\begin{cases} m_1 = \bar{X} & , & m_2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \\ \mu_1 = \mu & , & \mu_2 = \sigma^2 + \mu^2 \end{cases}$$

$$\Rightarrow \begin{cases} \frac{1}{n} \sum X_i = \mu \\ \frac{1}{n} \sum X_i^2 = \mu^2 + \sigma^2 \end{cases} \Rightarrow \underline{\sigma^2 = \frac{1}{n} \sum (X_i - \bar{X})^2}$$

مثال: فرض کنید $X_1, \dots, X_n$ مشاهدات iid از توزیع دوجمله‌ای با پارامترهای

$$(p, k) \quad . \quad \text{که } k \text{ و } p \text{ نامعلوم}$$

مثالی از مقدار هم‌هاریکد روز و نرخ گذارشی p (P).

با استفاده از روشی گشتاورها:

$$\begin{cases} \bar{X} = P \cdot K \\ \frac{1}{n} \sum X_i^2 = K P (1-P) + K^2 P^2 \end{cases}$$

$$\begin{cases} \hat{K} = \frac{\bar{X}^2}{\bar{X} - (\frac{1}{n}) \sum (X_i - \bar{X})^2} \\ \hat{P} = \frac{\bar{X}}{\hat{K}} \end{cases}$$

امکان منفی شدن $\hat{K}$ و $\hat{P}$ وجود دارد

• تعریف بیشینه likelihood:

$$L(\underline{\theta} | \underline{x}) = \prod_{i=1}^n f(x_i | \theta_1, \dots, \theta_k)$$

تآیث: برای هر نقطه $\underline{x}$ عرض کنیم

$$\hat{\theta}_{ML}(\underline{x}) = \underset{\theta}{\operatorname{argmax}} L(\theta | \underline{x})$$

تذیکه ML : $\hat{\theta}_{ML}(\underline{x})$

• برای تذیکه ML یا در مسئله مواجه هستیم:

① میتوانیم بیشه سرکاری یک تابع likelihood را در مسئله خوردن صدای کنیم.

② حسیت عددی به تقییم ورودی .

• برای یافتن $\hat{\theta}_{ML}$ یا یک مسئله بهینه سازی حل کنیم .

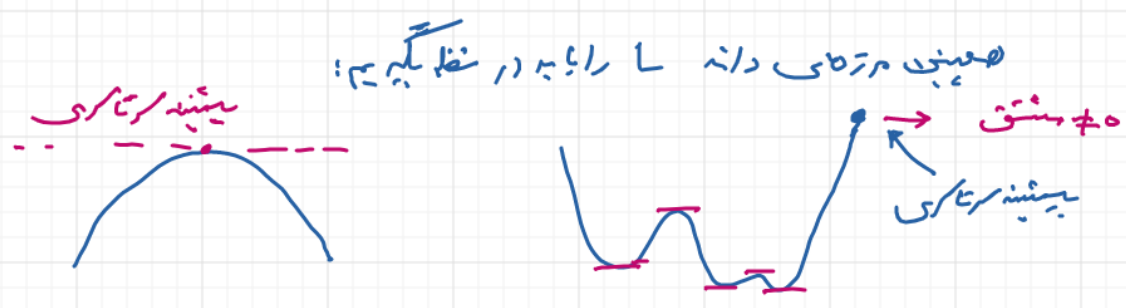
اگر تابع likelihood متبسه θ_i را مشتق پذیر؛ که $\Leftarrow$

که تریله ای برای نقاط بیشه:

$$\frac{\partial}{\partial \theta_i} L(\theta | \underline{x}) = 0$$

نقاط کید محلی ، بیشه محلی و نقاط محط $\rightarrow$

دامل دانه تايچ



مثال: فرض کړئ $x_1, \dots, x_n$ مشاهدات لږ از توږيچ $\mathcal{N}(\mu, 1)$ بياي تايچ likelihood:

$$L(\mu | \underline{x}) = \underbrace{(2\pi)^{-\frac{n}{2}}}_{=C} \exp\left(-\frac{\sum (x_i - \mu)^2}{2}\right)$$

$$\frac{dL}{d\mu} = C \cdot \exp\left(-\frac{\sum (x_i - \mu)^2}{2}\right) \left(\sum (x_i - \mu)\right) = 0$$

$\Rightarrow \hat{\mu}_{ML} = \frac{1}{n} \sum x_i$: $\hat{\mu}$ واقعا این بیشینه
سرتا می آید!

$\left. \frac{d^2 L}{d\mu^2} \right|_{\mu = \frac{1}{n} \sum x_i} < 0$ ✓

$\lim_{\mu \rightarrow \pm \infty} L = 0$ ✓ استقنای بیازده ها:

$\hat{\mu}_{ML} = \frac{1}{n} \sum x_i$ ← واقعا بیشینه کراری تابع L هست ← تنگه ML .

• نکته: می توانیم به جای L با $-L$ کار کنیم و همان نتیجه را بدست می آوریم.

• مثال: $x_1, \dots, x_n$ متغیرات iid از توزیع $N(\mu, 1)$ و می دانیم $\mu > 0$.

سوال: تنگه ML در این حالت چگونه می شود؟

↔ مسئله بهینه‌سازی مقبوض داریم.

$$L(\mu | \underline{x}) = f(\underline{x} | \mu) = (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum (x_i - \bar{x})^2\right) \exp\left(-\frac{n}{2} (\bar{x} - \mu)^2\right)$$

اگر $\bar{x} < 0$ بر حسب μ برای $\mu \geq 0$ نزدیک خواهد بود.

↔ مقدار بیشینه $L(\mu | \underline{x})$ در نقطه $\hat{\mu} = 0$ اتفاق افتاده.

$$\Rightarrow \hat{\mu}_{ML} = \begin{cases} \bar{x} & : \bar{x} \geq 0 \\ 0 & : \bar{x} < 0 \end{cases}$$

خواص تغییر ML (خواص سری):

① Consistency (مطابقت):

$$\hat{\theta}_{ML} \xrightarrow[n \rightarrow \infty]{P} \theta_{true}$$

② خاصیت تغییر ناپذیری تخمین ML:

فرض کنید یک مدل با پارامتر θ ← $\hat{\theta}_{ML}$
به همی اینکه به یافتن θ علاقه نداشته باشیم
علاقه به تخمین تابعی از θ داشته باشیم $\gamma(\theta)$
هستیم.

اگر $\hat{\theta}_{ML}$ تخمین ML برای پارامتر θ باشد، آنگاه $\gamma(\hat{\theta}_{ML})$ تخمین ML برای $\gamma(\theta)$ است.

③ تخمین ML برای $n \rightarrow \infty$ در کان Cramer-Rao صدق میکند.
(Efficiency)

④ تغییر μ_{ML} های $n \rightarrow \infty$ عزیز مثال پیدا کنی.

مثال: $x_1, \dots, x_n$ متغیرات iid از توزیع $\mathcal{N}(\mu, \sigma^2)$ که μ و σ^2 نامعلوم هستند.

$$\ln L(\mu, \sigma^2 | \underline{x}) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^2}$$

$$\begin{cases} \frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \\ \frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = 0 \end{cases}$$

$$\begin{cases} \hat{\mu}_{ML} = \bar{x} \\ \hat{\sigma}_{ML}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \end{cases}$$

در واقع $\hat{\mu}_{ML}$ یک کلمه بی‌بیم بیهوشگر است
هسته یا نه

برای پیدا کردن اینکه نقاط (θ_1, θ_2) مینه محلی اند

① $\frac{\partial L}{\partial \theta_1} \Big|_{\hat{\theta}_1, \hat{\theta}_2} = 0, \quad \frac{\partial L}{\partial \theta_2} \Big|_{\hat{\theta}_1, \hat{\theta}_2} = 0$

② در امتداد یکی از مشتقات مرتبه دوم منفی باشند

$\frac{\partial^2 L}{\partial \theta_1^2} \Big|_{(\hat{\theta}_1, \hat{\theta}_2)} < 0 \quad \text{یا} \quad \frac{\partial^2 L}{\partial \theta_2^2} \Big|_{(\hat{\theta}_1, \hat{\theta}_2)} < 0$

③ در مینیمم ماتریس هسیان تابع L مثبت باشد:

$$\begin{vmatrix} \frac{\partial^2 L}{\partial \theta_1^2} & \frac{\partial^2 L}{\partial \theta_1 \partial \theta_2} \\ \frac{\partial^2 L}{\partial \theta_1 \partial \theta_2} & \frac{\partial^2 L}{\partial \theta_2^2} \end{vmatrix} \Big|_{(\hat{\theta}_1, \hat{\theta}_2)} > 0$$

۱. داده‌ای به بیفت صابیت کردی تغییر ML :

مشاهده $x \leftarrow L(\theta | x)$

مشاهده $x + \epsilon \leftarrow L(\theta | x + \epsilon)$: انتظار داریم اگر ϵ کوچک باشد

تغییرات تابع L هر کم باشد. ولی لزوماً
ایکسان نیست.

مثال: برای مسئله مشاهده تعدادی جرم که از توزیع دو جمله‌ای (p, k) مشاهده شده اند

① مشاهدات: 16, 18, 22, 25, 27 $\Rightarrow \hat{k}_{ML} = 99$

② مشاهدات: 16, 18, 22, 25, 28 $\Rightarrow \hat{k}_{ML} = 190$

انتخاب این یعنی اسک L ;
 flat



there is no finite maximum