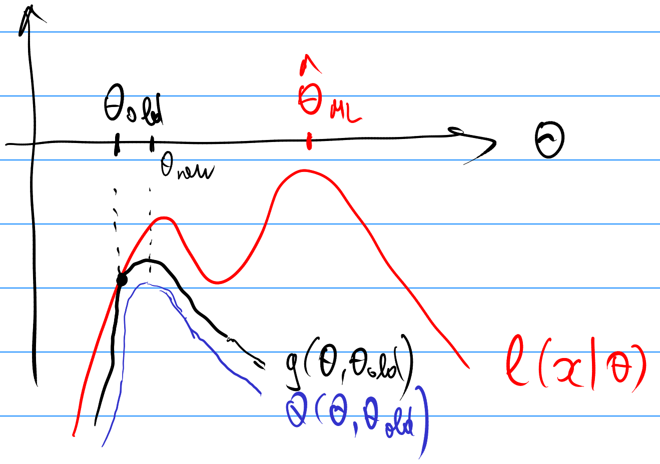
Привет !! :) 2021-10-08

→ EM

→ Bootstrap : как строить дов-ые интервалы, если забыл матстат !



EM

одна итерация EM-алгоритма.

вход: θ_{old}

E-шаг: находим ф-ции $\underline{g(\theta, \theta_{old})}$, $\underline{Q(\theta, \theta_{old})}$

M-шаг $\theta_{new} = \arg\max_{\theta} Q(\theta, \theta_{old})$

выход: θ_{new}

$$g(\theta, \theta_{old}) = E\left[\ln f(x, z|\theta) - \ln f(z|x, \theta_{old}) \,\middle|\, x, \theta_{old}\right]$$

$$Q(\theta, \theta_{old}) = E\left[\underline{\ln f(x, z|\theta)} \,\middle|\, x, \theta_{old}\right]$$

модель: $\nearrow (1)$ $\searrow (2)$ $z_i \in \{1, 2\}$ $N(\mu_1, 1)$ $N(\mu_2, 1)$ p_1 $p_2 = 1 - p_1$

$$\theta = \begin{pmatrix} p_1 \\ \mu_1 \\ \mu_2 \end{pmatrix}$$

модель → $$f(x_i, z_i | \theta) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i - \mu_1)^2\right) \cdot p_1 & \text{в } z_i = 1 \\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i - \mu_2)^2\right) \cdot p_2 & \text{в } z_i = 2 \end{cases}$$

x_i есть !

$h(z) = z^2$

z	0	1	2	3
h	0	1	4	9

$$\theta = \begin{pmatrix} p_1 \\ \mu_1 \\ \mu_2 \end{pmatrix}$$

6 7 8 9 10

$$P(z_i=1 \mid x_i, \theta_{old}) = \frac{f(x_i, z_i=1 \mid x_i, \theta_{old})}{f(x_i \mid \theta_{old})} =$$

$$= \frac{\frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}(x_i - \mu_1^{old})^2\right) \cdot p_1^{old}}{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x_i - \mu_1^{old})^2\right) \cdot p_1^{old} + \frac{1}{\sqrt{2\pi}} \exp(\underbrace{\ldots}_{\mu_2^{old}}) \cdot p_2^{old}}$$

$$E\left(\ln f(x_i, z_i \mid \theta) \mid x_i, \theta_{old}\right) =$$

$$= \underbrace{P(z_i=1 \mid x_i, \theta_{old}) \cdot \ln f(x_i, 1 \mid \theta) + P(z_i=2 \mid x_i, \theta_{old}) \cdot \ln f(x_i, 2 \mid \theta)}_{\text{зависит от: } x_i,\ \theta = \begin{pmatrix} p_1 \\ \mu_1 \\ \mu_2 \end{pmatrix} \ \theta_{old} = \begin{pmatrix} p_1^{old} \\ \mu_1^{old} \\ \vdots \\ \mu_2^{old} \end{pmatrix}}$$

$$\theta_{new} = \arg\max_{\theta} \sum_{i=1}^{n} E(\ldots\ldots)$$

Трудность! Думать надо 

E → выводится явно
E → сумма слаг. E( ) для дискретного z
E → сумма аппрокс. E( ) для непр z

E. $$Q(\theta, \theta_{old}) = E\left[\underbrace{\ln f(x, z \mid \theta)}_{\text{модель есть!}} \mid x, \theta_{old}\right]$$

$\boxed{\text{надо рассчитать } f(z \mid x, \theta_{old})}$

$\boxed{\text{Трюк!}}$ Вместо того, чтобы выводить руками $f(z \mid x, \theta_{old})$ мы запустим оптимизатор, который её найдёт!

EM — для ленивых.

$x, \theta^{old} \to$

E. $\max\limits_q$ [illegible] $\qquad q^* = f(z|x,\theta_{old})$

$$Q(\theta,\theta_{old}) = E\left[\ln f(x,z|\theta) \,|\, x, \theta_{old}\right]$$

M.

$$\theta_{new} = \arg\max_\theta Q(\theta,\theta_{old})$$

$\theta_{new} \leftarrow$

$\min\limits_q (q-5)^2$

? $\quad KL\left(f(z|x,\theta_{old}) \,||\, f(z|x,\theta_{old})\right) = 0$

наивность $\qquad \min\limits_q KL(q \,||\, f(z|x,\theta_{old})) \qquad q^* = f(z|x,\theta_{old})$

почему не работает? $\qquad$ чтобы его реализовать надо знать ответ!

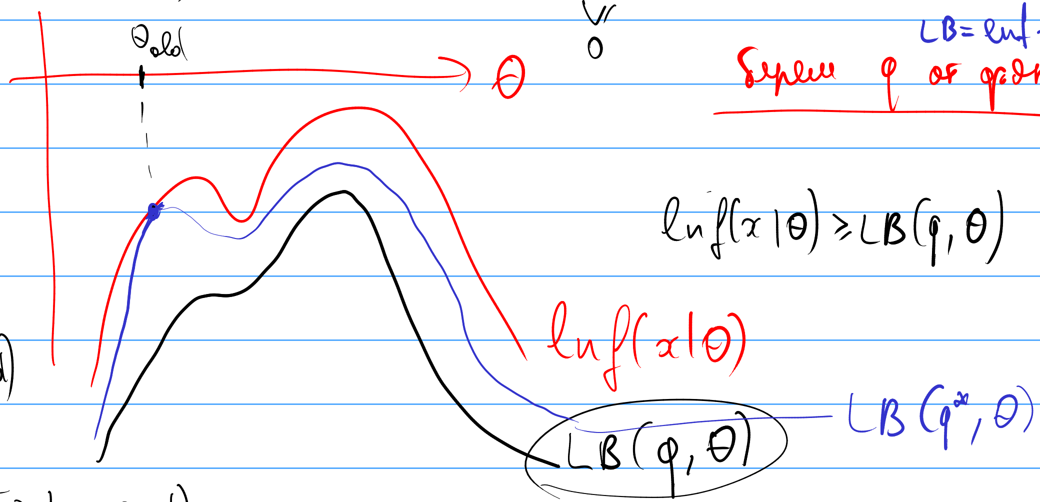
Спасение наивного метода!

есть такое разложение

$$\ln f(x|\theta) = \underbrace{KL(q \,||\, f(z|x,\theta))}_{\geq 0} + LB(q,\theta)$$

$LB = \ln f - KL$

берём q от фонаря



$$\ln f(x|\theta) \geq LB(q,\theta)$$

EM-итер.

Вход: θ_{old}

E $\quad q^* = \arg\max\limits_q LB(q,\theta_{old})$

$q^* = f(z|x,\theta_{old})$

$Q(\theta,\theta_{old}) = E(\ln f(x,z|\theta) \,|\, x,\theta_{old})$

M $\quad \theta_{new} = \arg\max\limits_\theta Q(\theta,\theta_{old})$

$$\arg\max_q LB(q,\theta) = \arg\min_q KL(q \,||\, f(z|x,\theta))$$

$n=100$

$z_i \in \{1,2\}$

$$LB(q,\theta) = \ln f(x|\theta) - KL\left(q \,\|\, f(z|x,\theta)\right) =$$

$$= \ln f(x|\theta) - CE\left(q \,\|\, f(z|x,\theta)\right) + H(q) =$$

$$= \ln f(x|\theta) + \int q(z)\cdot \ln f(z|x,\theta)\,dz - \int q(z) \ln q(z)\,dz =$$

$$= \int q(z) \ln f(x|\theta)\,dz + \text{—//—} =$$

$$= \int q(z)\cdot\left[\ln f(x|\theta) + \ln f(z|x,\theta) - \ln q(z)\right] dz =$$

$$LB(q,\theta) = \int q(z)\cdot \ln \frac{f(x,z|\theta)}{q(z)}\,dz$$

$q(z)$ ищем!

$f(x,z|\theta)$ – модель! (задали)

EM-2

x данные

θ_{old}

E-шаг

$$q^* = \arg\max_q LB(q, \theta_{old}) \quad \Big\{ \; q^* = f(z|x,\theta_{old})$$

$$Q(\theta, \theta_{old}) = E\left(\ln f(x,z|\theta) \,\middle|\, x, \theta_{old}\right)$$

$$\int q^*(z)\cdot \ln f(x,z|\theta)\cdot dz$$

просто сумма

M-шаг

$$\theta_{new} = \arg\max_\theta Q(\theta, \theta_{old})$$

(оценит) **Bootstrap.** Бутстрэп

μ

$[\hat{\mu} - 1.96 \cdot se(\hat{\mu}); \hat{\mu} + 1.96 \cdot se(\hat{\mu})]$

θ - неизв. пар.

n набл. $\to \xi \to \hat{\theta}$

$\theta = Med(f(y_i))$

$\hat{\theta} = sMed(y)$

Цель: CI для θ [95%]

conf interval

$y_1, \dots, y_n \to \hat{\theta}(y_1, \dots, y_n)$

бутстрэп выб №1: из исх-ой выборки размера n берём наугад рав-но n набл-ий с возм-стью повт-рного выбора.

$y_1^*, \dots, y_n^* \to \hat{\theta}_1^*$

$Law(\hat{\theta}_1^*) \approx Law(\hat{\theta})$

бут-выборка №2

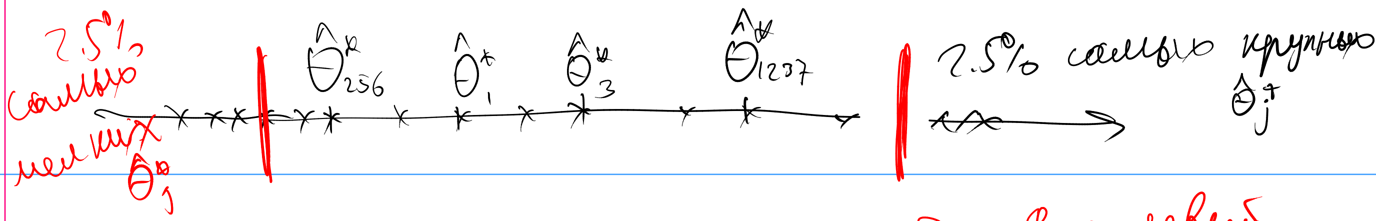
$y_1^*, \dots, y_n^* \to \hat{\theta}_2^* \quad \dots$

$y_1 \dots y_n \to \hat{\theta}$

$y_1^* \dots y_n^*$

$n_{boot} = 10000$

$\hat{\theta}_1^* \quad \hat{\theta}_2^* \quad \hat{\theta}_3^* \quad \dots \quad \hat{\theta}_{10000}^*$



CI для θ: $\left[q_L(\hat{\theta}^*) ; q_R(\hat{\theta}^*) \right]$

q_L – левый квантиль

q_R – правый квантиль.