Đinh Đức Tài

23122013

Homework 2

Toán cho AI - 23TNT1

1. Ước lượng hợp lý cực đại (MLE) cho phân phối chuẩn

Hàm phân phối xác suất của phân phối chuẩn:

$$p(x,\alpha) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]}, \quad \alpha = (\mu, \sigma)$$

a/ $\theta = \{x_1, x_2, \dots, x_n\}$. $|\theta| = n$

$$\Rightarrow L(\alpha | x_i) = p(x_i, \alpha) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]}$$

$$\forall i = \overline{1,n}$$

$$\Rightarrow L(\alpha|\theta) = \prod_{i=1}^{n} \cancel{p(x_i,} L(\alpha | x_i)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{\left[-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right]}$$

b/ $\Rightarrow \ell(\alpha|\theta) = \ln\left(L(\alpha|\theta)\right)$

$$= \frac{-n}{2} \cdot \ln(2\pi\sigma^2) + \left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right)$$

c/ $\Rightarrow \dfrac{\partial \ell}{\partial \mu} = \dfrac{2\sum_{i=1}^{n}(x_i-\mu)}{2\sigma^2} = \dfrac{1}{\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)$

$$\frac{\partial \ell}{\partial \mu} = 0 \Rightarrow \sum_{i=1}^{n}(x_i-\mu) = 0 \Rightarrow \mu = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x}$$

$$\frac{\partial l}{\partial \sigma} = \frac{-n}{2} \cdot \frac{(2\pi\sigma^2)'}{2\pi\sigma^2} - \sum_{i=1}^{n}(x_i-\mu)^2 \cdot \frac{1}{2} \cdot (\sigma^{-2})'$$

$$= \frac{-n}{2} \cdot \frac{2 . 2\pi . \sigma}{2\pi\sigma^2} + \sum_{i=1}^{n}(x_i-\mu)^2 \, \sigma^{-3}$$

$$= \frac{-n}{\sigma} + \sum_{i=1}^{n}(x_i-\mu)^2 . \sigma^{-3}$$

$$\frac{\partial l}{\partial \sigma} = 0 \Rightarrow \frac{n}{\sigma} = \cancel{n} \sum_{i=1}^{n}(x_i-\mu)^2 . \sigma^{-3}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$\Rightarrow \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2} = \sqrt{Var}$$

$$\Rightarrow \alpha_{MLE} = \arg\max l(\alpha|\theta) = (\bar{x}, \sqrt{Var})$$

2. Thuật toán phân cụm mean shift và gradient ascent

$$\theta = \{x_1, x_2, \ldots, x_n\} \subset R^2, \quad h = 1.$$

$$\hat{f}(x,\theta) = \frac{1}{n.h^2} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right),$$

$$K(u) = \frac{1}{2\pi} e^{\left(-\frac{1}{2}\|u\|^2\right)}$$

Thuật toán cập nhật $\theta^{(t)}$:

$$x_i^{(t+1)} = \frac{\sum_{J=1}^{n} x_J^{(t)} K\left(x_i^{(t)} - x_J^{(t)}\right)}{\sum_{J=1}^{n} K\left(x_i^{(t)} - x_J^{(t)}\right)}$$

a/ 2 bước đầu tiên của thuật toán mean shift:

$$\theta^{(0)} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\}$$

Bước 1:

i	$\sum_{J=1}^{n} x_J^{(0)} K\left(x_i^{(0)} - x_J^{(0)}\right)$	$\sum_{J=1}^{n} K\left(x_i^{(0)} - x_J^{(0)}\right)$	$x_i^{(1)}$
1	$(0{,}00852129; 0{,}0976047)^T$	0,258528	
2	$(0{,}00852129; 0{,}160923)^T$	0,258528	
3	$(0{,}0767062; 0{,}0976047)^T$	0,258528	
4	$(0{,}0767062; 0{,}160923)^T$	0,258528	

$$\Rightarrow \theta^{(1)} = \left\{ \begin{pmatrix} 0,0329608 \\ 0,377541 \end{pmatrix}, \begin{pmatrix} 0,0329608 \\ 0,622459 \end{pmatrix}, \begin{pmatrix} 2,96704 \\ 0,377541 \end{pmatrix}, \begin{pmatrix} 2,96704 \\ 0,622459 \end{pmatrix} \right\}$$

Bước 2:

i	$\sum_{J=1} x_J^{(1)} K(x_i^{(1)} - x_J^{(1)})$	$\sum_{J=1}^{n} K(x_i^{(1)} - x_J^{(1)})$	$x_i^{(2)}$
1	$(0,0229066; 0,158338)^T$	0,317844	
2	$(0,229066; 0,159506)^T$	0,317844	
3	$(0,930625; 0,158338)^T$	0,317844	
4	$(0,930625; 0,159506)^T$	0,317844	

$$\Rightarrow \theta^{(2)} = \left\{ \begin{pmatrix} 0,0720686 \\ 0,498164 \end{pmatrix}, \begin{pmatrix} 0,0720686 \\ 0,501836 \end{pmatrix}, \begin{pmatrix} 2,92793 \\ 0,498164 \end{pmatrix}, \begin{pmatrix} 2,92793 \\ 0,501836 \end{pmatrix} \right\}$$

Với $h = 1$, $\theta^{(t)} = \{x_1^{(t)}, x_2^{(t)}, \dots, x_n^{(t)}\}$

$$\Rightarrow \hat{f}(x, \theta^{(t)}) = \frac{1}{n} \sum_{k=1}^{n} K(x - x_k^{(t)})$$

$$\Rightarrow \nabla_x \hat{f}(x, \theta^{(t)}) = \nabla_x \left(\frac{1}{n} \sum_{k=1}^{n} K(x - x_k^{(t)}) \right)$$

$$= \frac{1}{n} \sum_{k=1}^{n} \nabla_x K(x - x_k^{(t)})$$

mà $K(u) = \frac{1}{2\pi} e^{\left(-\frac{1}{2}\|u\|^2\right)}$

$$\Rightarrow \nabla_u K(u) = \frac{1}{2\pi} e^{\left(-\frac{1}{2}\|u\|^2\right)} \cdot (-u) = -uK(u)$$

$$\Rightarrow \nabla_x K(x - x_k^{(t)}) = -(x - x_k^{(t)}) K(x - x_k^{(t)})$$

$$\Rightarrow \nabla_x \hat{f}(x, \theta^{(t)}) = \frac{1}{n} \sum_{k=1}^{n} \left[-(x - x_k^{(t)}) K(x - x_k^{(t)}) \right]$$

$$= \frac{1}{n} \left[\sum_{k=1}^{n} x_k^{(t)} K(x - x_k^{(t)}) - x \sum_{k=1}^{n} K(x - x_k^{(t)}) \right]$$

Với $x = x_i^{(t)}$

$$\Rightarrow \nabla_x \hat{f}(x, \theta^{(t)}) \Big|_{x = x_i^{(t)}}$$

$$= \frac{1}{n} \left[\sum_{k=1}^{n} x_k^{(t)} K(x_i^{(t)} - x_k^{(t)}) - x_i^{(t)} \sum_{k=1}^{n} K(x_i^{(t)} - x_k^{(t)}) \right]$$

Đặt $S_k = \sum_{k=1}^{n} K(x_i^{(t)} - x_k^{(t)})$

$$\Rightarrow \nabla_x \hat{f}(x \mid \theta^{(t)}) \Big|_{x = x_i^{(t)}}$$

$$= \frac{S_k}{n} \left[\frac{\sum_{k=1}^{n} x_k^{(t)} K(x_i^{(t)} - x_k^{(t)})}{S_k} - x_i^{(t)} \right]$$

Với công thức mean shift, ta có:

$$x_i^{(t+1)} = \frac{\sum_{k=1}^{n} x_k^{(t)} K(x_i^{(t)} - x_k^{(t)})}{\sum_{k=1}^{n} K(x_i^{(t)} - x_k^{(t)})}$$

$$= \frac{\sum_{k=1}^{n} x_k^{(t)} K(x_i^{(t)} - x_k^{(t)})}{S_k}$$

$$\Rightarrow \nabla_x \hat{f}(x \mid \theta^{(t)}) \Big|_{x = x_i^{(t)}} = \frac{S_k}{n} \left(x_i^{(t+1)} - x_i^{(t)} \right)$$

$$\Rightarrow x_i^{(t+1)} = x_i^{(t)} + \frac{n}{S_k} \nabla_x \hat{f}(x \mid \theta^{(t)}) \Big|_{x = x_i^{(t)}} \quad (*)$$

Dạng cập nhật Gradient Ascent để tối đa hoá hàm KDE: $\hat{f}(x \mid \theta^{(t)})$, đi từ $x_i^{(t)}$:

$$x_i^{(t+1)} = x_i^{(t)} + \alpha_i^{(t)} \nabla_x \hat{f}(x \mid \theta^{(t)}) \Big|_{x = x_i^{(t)}} \quad (**)$$

(đi theo chiều gradient hàm KDE)

Từ (*), (***)

$$\Rightarrow x_i^{(t+1)} = x_i'^{(t+1)} \text{ với } \alpha_i^{(t)} = \frac{n}{S_k}$$

Hay mỗi bước cập nhật mean shift tương ứng với một bước của thuật toán gradien ascent cho hàm KDE tại thời điểm t.

Tốc độ học tương ứng là:

$$\alpha_i^{(t)} = \frac{n}{S_k} = \frac{n}{\sum_{k=1}^{n} K\left(x_i^{(t)} - x_k^{(t)}\right)}$$

$\alpha_i^{(t)}$ phù hợp với nhận xét tốc độ học thay đổi theo từng bước t.