سوال 8

a)

چون X و Y Independent هستند ← $f_{X,Y}=f_X f_Y \Leftarrow$

از طرفی μ را نیز از آن دو مستقل درنظر میگیریم

$$f_{\mu XY} = f_{XY} f_\mu = 1 \times f_{XY} = f_X f_Y$$

$$f_{\mu,x,y}(t,x,y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\mu)^2}$$

$$\Rightarrow \text{joint} = f_{\mu,X,Y}(t,x,y) = \boxed{\frac{1}{2\pi} e^{-\frac{1}{2}((x-\mu)^2+(y-\mu)^2)}}$$

b find MAP:

$f((X,Y))$ دیتاها مشترک $(X,Y)_i = (X_i,Y_i)$

به تعداد N و به صورت جفت فرض کرده ایم. (دوتایی مرتب)

به خاطر توزیع یونیفرم ظاهر شده است.

$$f(\text{Data}) = \prod f_\mu(X,Y)\, f(\mu) = \prod f_{XY}(x,y)\, [u(\mu) - u(\mu-1)]$$

u تابع پله است.

$$\Rightarrow L(\mu) = \sum \lg\left[\frac{1}{2\pi} e^{-\frac{1}{2}((x-\mu)^2+(y-\mu)^2)}\right) u(\mu)-u(\mu-1)]$$

$$\Rightarrow L(\mu) = \sum_{(x,y)} \lg\left[\frac{e^{-\frac{1}{2}((x-\mu)^2+(y-\mu)^2)}}{2\pi}(u(\mu)-u(\mu-1))\right.$$

$$\hat{\mu} = \underset{\mu}{\text{Argmax}}\, L \Rightarrow \frac{\partial L}{\partial \mu} = 0$$

$$\Rightarrow L(\mu) = N \lg\left(\frac{1}{2\pi}\right) + N \lg\left(u(\mu - u(\mu-1))\right) + \sum_{(x,y)}^{N} \left(-\frac{1}{2}\right)\left((x-\mu)^2+(y-\mu)^2\right)$$

$$\frac{\partial L}{\partial \mu} = N \frac{\partial}{\partial \mu} \lg\left(\frac{1}{2\pi}\right) + N \frac{\delta(\mu)-\delta(\mu-1)}{u(\mu)-u(\mu-1)} + \sum_{(x,y)}^{N} [(x_i-\mu)+(y_i-\mu)] = 0$$

$$\Rightarrow \sum_{(x,y)}^{N} 2\mu = \sum_{(x,y)}^{N} (x_i+y_i) + \frac{\delta(\mu)-\delta(\mu-1)}{u(\mu)-u(\mu-1)} \times N$$

$$\Rightarrow 2N\mu = \sum_N (x_i+y_i) + \frac{\delta(\mu)-\delta(\mu-1)}{u(\mu)-u(\mu-1)} \times N$$

$$\Rightarrow \hat{\mu} = \frac{\sum_N (x_i+y_i)}{2N} + \frac{1}{2} \times \underbrace{\frac{\delta(\mu)-\delta(\mu-1)}{u(\mu)-u(\mu-1)}}_{0 < \hat{\mu} \leq 1 \;=\; 0} \Rightarrow \begin{cases} \frac{\sum_N x_i+y_i}{2N} & 0 < \hat{\mu} \leq 1 \\ 0 & \hat{\mu} \leq 0 \\ 1 & \hat{\mu} > 1 \end{cases}$$

Scanned by CamScanner

ادامه 8

بنابراین داریم

$$\hat{\mu} = \frac{\sum_{N} x_i + y_i}{2N}$$

$$\Rightarrow \hat{\mu}_{MaP} = \begin{cases} \hat{\mu} & 1 > \hat{\mu} > 0 \\ 1 & \hat{\mu} > 1 \\ 0 & \hat{\mu} < 0 \end{cases}$$

دقت کنید که جزء $\frac{1}{2}\frac{\delta(\mu)-\delta(\mu-1)}{u(\mu)-u(\mu-1)}$ در تساوی بالا حذف شده اند زیرا

این جزء اجازه خروج $\hat{\mu}$ از محدوده $1 > \hat{\mu} > 0$ را نمی داد

زیرا در $\hat{\mu} > 1$ یا $\hat{\mu} < 0$ یک عبارت $\frac{0}{0}$ می شود با حتما

می شود $\Leftarrow$ Postrior اثر خود را گذاشت و $\hat{\mu}$ بین ۱ و ۰ ماند که

ادامه قسمت ط

شاید عبارت $u(t)$ در عبارت max likelihood که گیج کننده باشد.

⟸ μ را به صورت زیر تعریف می‌کنیم

$$\mu_{MAP} = \begin{cases} 1 & \mu_{MLE} > 1 \\ \mu_{MLE} & 0 \le \mu_{MLE} \le 1 \\ 0 & \mu_{MLE} < 0 \end{cases}$$

دقت کنید توزیع μ یک توزیع یکنواخت است بنابراین

$$\mu_{MAP} = \operatorname{Argmax} f(\mu|x) = \operatorname{Argmax} f(x|\mu)\, g(\mu)$$

چون $g(\mu)=1 \Longrightarrow \;\to\; \mu_{MAP} = \operatorname{Argmax} f(x|\mu) \;\to\; \hat{\mu} = \operatorname{Argmax} f(x|\mu)$

if $0 \le \hat{\mu} \le 1$ $\quad \hat{\mu} = \mu_{MLE}$

$$\boxed{\hat{\mu} = \mu_{MLE}}$$

$$\mu_{MAP} = \begin{cases} 1 & \hat{\mu} > 1 \\ \hat{\mu} & 0 \le \hat{\mu} \le 1 \\ 0 & \hat{\mu} < 0 \end{cases}$$

دقت کنید ← چون مسأله را با مشتق‌گیری حل کرده‌ایم فرض کرده‌ایم تابع پیوسته است و دوم آنکه چون در یک نقطه مشتق صفر شده است. بنابراین نقاط نزدیک‌تر به نقطه بدست آمده پاسخ بهتری‌اند (چون بازه باید بین صفر و یک باشد)

$$\hat{\mu} \to \operatorname{Argmax} \prod \frac{1}{2\pi} e^{-\frac{1}{2}\left((x-\mu)^2+(y-\mu)^2\right)}$$

$$\to L(\mu|x,y) = N \log\left(\frac{1}{2\pi}\right) + \sum_{(x,y)} -\frac{1}{2}\left[(x_i-\mu)^2+(y_i-\mu)^2\right] =$$

$$\to \frac{\partial L}{\partial \mu} = 0 = N \cancel{\frac{\partial}{\partial \mu}\log\frac{1}{2\pi}} + \sum_{(x,y)} (x_i+y_i) - 2N\mu = 0 \Longrightarrow$$

$$\boxed{\hat{\mu} = \frac{\sum_{(x,y)}^{N} x_i+y_i}{2N}} \longrightarrow \mu_{MAP} = \begin{cases} 1 & \hat{\mu} > 1 \\ \hat{\mu} & 0 \le \hat{\mu} \le 1 \\ 0 & \hat{\mu} < 0 \end{cases}$$