ص ①

هوش مصنوعی پیشرفته

تکلیف شماره‌ی ۲   فصل ۱۴ : استدلال احتمالاتی

۱)

الف) (ii) بطور واضح نادرست است ← بیان می‌کند که با داشتن اندازه‌گیری‌های M_1 و M_2 تعداد ستاره‌ها مستقل از کانون عدسی است.

(iii) ساختار علّی را به درستی بازنمایی می‌کند : هر اندازه‌گیری با تعداد واقعی ستاره‌ها و کانون عدسی‌ها تحت تأثیر قرار می‌گیرد و دو تلسکوپ مستقل از هم هستند.

(iii) یک شبکه‌ی درست، اما پیچیده‌تر را نشان می‌دهد.

ب) شبکه (ii) بهتر است، زیرا تعداد پارامتر کمتری لازم دارد (نسبت به iii)

ج) برای محاسبه $P(M_1|N)$ لازم است بر روی F_1 شرط گذاری شود (هر دو حالت F_1 درنظر گرفته می‌شود و با احتمال آنها وزن دهی می‌شود) :

$$P(M_1|N) = P(M_1|N,F_1)\,P(F_1|N) + P(M_1|N,\neg F_1)\,P(\neg F_1|N)$$

$$= P(M_1|N,F_1)\,P(F_1) + P(M_1|N,\neg F_1)\,P(\neg F_1)$$

فرض کنید که f احتمال این باشد که تلسکوپ خارج از کانون باشد (کانون تلسکوپ خراب باشد). این تمرین بیان می‌کند که این موجب «زیر شمارش سه یا بیشتر ستاره» می‌شود، اما اگر $N = 3$ یا کمتر باشد، شمارش صفر خواهد بود اگر تلسکوپ خارج از کانون باشد. اگر تلسکوپ در کانون باشد (کانون تلسکوپ سالم باشد)، در این صورت فرض می‌کنیم یک احتمال e برای شمارش یکی دوتا کمتر و احتمال e برای شمارش یک خیلی بیشتر وجود دارد. در مابقی زمانها $(1-2e)$ شمارش دقیق خواهد بود. در این صورت جدول احتمال شرطی به صورت زیر خواهد بود :

	$N=1$	$N=2$	$N=3$
$M_1=0$	$f+e(1-f)$	f	f
$M_1=1$	$(1-2e)(1-f)$	$e(1-f)$	0
$M_1=2$	$(1-f)e$	$(1-2e)(1-f)$	$e(1-f)$
$M_1=3$	0	$e(1-f)$	$(1-2e)(1-f)$
$M_1=4$	0	0	$e(1-f)$

جمع هر ستون برابر با 1 است.

مقادیر منطقی برای e و f می‌تواند برابر باشد با : $e = 0.05$ $f = 0.002$

د) امکان محاسبه‌ی محتمل‌ترین تعداد ستاره‌ها بدون دانستن توزیع پیشین $P(N)$ وجود ندارد.

احتمالات پیشین را p_2 ، p_4 و $p_{\geq 6}$ در نظر می‌گیریم.

احتمال پسین برای $N = 2$ می‌شود $p_2 e^2(1-f)^2$

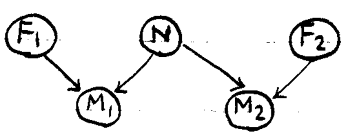
" " برای $N = 4$ حداکثر می‌شود $p_4 e f$

(زیرا $N = 4$ حالت out-of-focus تلسکوپ است که به جای 1، 0 اندازه می‌گیرد)

" " برای $N \geq 6$ حداکثر می‌شود $p_{\geq 6} f^2$

اگر فرض کنیم احتمالات پیشین تقریباً مساوی هستند، آنگاه $N = 2$ محتمل‌ترین است، زیرا f بسیار کوچکتر از e فرض شده است.

۲)



$M_1, M_2 \in \{0, 1, 2, 3, 4\}$

$N \in \{1, 2, 3\}$

$P(N \mid M_1 = 2, M_2 = 2)$

$$P(N \mid M_1, M_2) = P(N, M_1, M_2) / P(M_1, M_2)$$

$$P(N \mid M_1 = 2, M_2 = 2) = P(N, M_1 = 2, M_2 = 2) / P(M_1 = 2, M_2 = 2)$$

$$= \alpha P(N, M_1 = 2, M_2 = 2)$$

$$= \alpha \sum_{F_1} \sum_{F_2} P(F_1, F_2, N, M_1 = 2, M_2 = 2)$$

با توجه به معناشناسی سراسری شبکه بیزی:

$$= \alpha \sum_{F_1} \sum_{F_2} P(F_1) P(F_2) P(N) P(M_1 \mid F_1, N) P(M_2 \mid F_2, N)$$

$$= \alpha P(N) \sum_{F_1} P(F_1) P(M_1 \mid F_1, N) \sum_{F_2} P(F_2) P(M_2 \mid F_2, N)$$

$$= \alpha P(N) \left(P(F_1) P(M_1 \mid F_1, N) + P(\neg F_1) P(M_1 \mid \neg F_1, N) \right)$$

$$\times \left(P(F_2) P(M_2 \mid F_2, N) + P(\neg F_2) P(M_2 \mid \neg F_2, N) \right)$$

س۳ (۳

$$\begin{tikzcd} X_1 \arrow[d] \\ X_2 \end{tikzcd}$$

$P(X_1)$ گاوسی

$P(X_2|X_1)$ گاوسی خطی

$$P(x)=\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}\,e^{-\frac{1}{2}\left((x-\underbrace{\mu}_{\text{بردار میانگین}})^T\underbrace{\Sigma^{-1}}_{\text{ماتریس کوواریانس}}(x-\mu)\right)}$$

$$x=\begin{bmatrix}x_1\\x_2\end{bmatrix}\qquad \mu=\begin{bmatrix}m_1\\m_2\end{bmatrix}$$

فرض: $\Sigma^{-1}=\begin{bmatrix}c & d\\ d & e\end{bmatrix}$

$$-\frac{1}{2}\left((x-\mu)^T\Sigma^{-1}(x-\mu)\right)=$$
$$-\frac{1}{2}\left(c(x_1-m_1)^2+2d(x_1-m_1)(x_2-m_2)+e(x_2-m_2)^2\right)\qquad (*)$$

$$P(x_1)=\frac{1}{\sigma_1\sqrt{2\pi}}\,e^{-(x_1-\mu_1)^2/(2\sigma_1^2)}$$
$$P(x_2|x_1)=\frac{1}{\sigma_2\sqrt{2\pi}}\,e^{-(x_2-(ax_1+b))^2/(2\sigma_2^2)}$$

$\Downarrow$

$$P(x_1,x_2)=P(x_2|x_1)\,P(x_1)$$
$$=\frac{1}{\sigma_1\sigma_2(2\pi)}\,e^{-(\sigma_2^2(x_2-(ax_1+b)^2+\sigma_1^2(x_2-(ax_1+b))^2)/(2\sigma_1^2\sigma_2^2)}$$

$\Downarrow\ (*)$

$$c=(\sigma_2^2+a^2\sigma_1^2)/\sigma_1^2\sigma_2^2$$
$$2d=-2a/\sigma_2^2$$
$$e=1/\sigma_2^2$$

$$\frac{1}{\sigma_1\sigma_2(2\pi)}=\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}=\frac{1}{(2\pi)\sqrt{1/|\Sigma^{-1}|}}=\frac{1}{(2\pi)\sqrt{\underbrace{(ce-d^2)}_{1/\sigma_1^2\sigma_2^2}}}$$

به طور مشابه m_1 و m_2 هم به دست می‌آید. با قرار دادن در فرمول $P(x_1,x_2)$ می‌بینیم که واقعاً گاوسی چند متغیره است.

ماتریس کوواریانس:
$$\Sigma=\begin{bmatrix}c & d\\ d & e\end{bmatrix}^{-1}=\frac{1}{ce-d^2}\begin{bmatrix}e & -d\\ -d & c\end{bmatrix}=\begin{bmatrix}\sigma_1^2 & a\sigma_1^2\\ a\sigma_1^2 & \sigma_2^2+a^2\sigma_1^2\end{bmatrix}$$

۴) الف)

$$P(B \mid j, m)$$
$$= \alpha P(B) \sum_e P(e) \sum_a P(a \mid b, e) P(j \mid a) P(m \mid a)$$
$$= \alpha P(B) \sum_e P(e) \left[0.9 \times 0.7 \times \begin{pmatrix} 0.95 & 0.29 \\ 0.94 & 0.001 \end{pmatrix} + 0.05 \times 0.01 \times \begin{pmatrix} 0.05 & 0.71 \\ 0.06 & 0.999 \end{pmatrix} \right]$$
$$= \alpha P(B) \sum_e P(e) \begin{pmatrix} 0.598525 & 0.183055 \\ 0.59223 & 0.011295 \end{pmatrix}$$
$$= \alpha P(B) \left[0.002 \times \begin{pmatrix} 0.598525 \\ 0.183055 \end{pmatrix} + 0.998 \times \begin{pmatrix} 0.59223 \\ 0.0011295 \end{pmatrix} \right]$$
$$= \alpha \begin{pmatrix} 0.001 \\ 0.999 \end{pmatrix} \times \begin{pmatrix} 0.59224259 \\ 0.01493351 \end{pmatrix} = \alpha \begin{pmatrix} 0.0059224259 \\ 0.0149185776 \end{pmatrix}$$
$$\approx \langle 0.284, 0.716 \rangle$$

ب) با وارد کردن گام نرمال‌سازی، ۷ جمع، ۱۶ ضرب، ۲ تقسیم داریم.
الگوریتم برشماری ۲ ضرب اضافی نیاز دارد.

ج) برای محاسبه‌ی $P(X_1 \mid X_n = true)$ با استفاده از برشماری باید دو درخت دودویی کامل (هر کدام برای یک مقدار X_1) هر یک با عمق $n-2$ را ارزیابی کنیم. پس کل کار می‌شود $O(2^n)$.
با استفاده از حذف متغیر، فاکتورها خارج از دو متغیر رشد نمی‌کند. برای مثال مرحله‌ی اول می‌شود:

$$P(X_1 \mid X_n = true)$$
$$= \alpha P(X_1) \cdots \sum_{x_{n-2}} P(x_{n-2} \mid x_{n-3}) \sum_{x_{n-1}} P(x_{n-1} \mid x_{n-2}) P(X_n = true \mid x_{n-1})$$
$$= \alpha P(X_1) \cdots \sum_{x_{n-2}} P(x_{n-2} \mid x_{n-3}) \sum_{x_{n-1}} f_{X_{n-1}}(x_{n-1}, x_{n-2}) f_{X_n}(x_{n-1})$$
$$= \alpha P(X_1) \cdots \sum_{x_{n-2}} P(x_{n-2} \mid x_{n-3}) f_{\overline{X_{n-1}} X_n}(x_{n-2})$$

خط آخر متناظر با یک مسئله با $n-1$ متغیر (به جای n متغیر) است. کار انجام شده در گام اول یک ثابت مستقل از n است، پس با استقرا روی n می‌توان گفت کار کل می‌شود $O(n)$.

د) از استقرا روی تعداد گره‌ها در polytree (pt) استفاده می‌کنیم. پایه بدیهی است. فرض استقرا: فرض می‌کنیم هر pt با n گره بتواند در زمانی متناسب با اندازه‌ی pt (یعنی مجموع اندازه‌ی CPT‌ها) محاسبه شود. حال یک pt با $n+1$ گره درنظر می‌گیریم. هر ترتیب از گره‌ها که با توپولوژی سازگار باشد، ابتدا برخی گره‌های برگ را از این pt حذف می‌کند. برای حذف همه‌ی برگ‌ها باید کاری متناسب با اندازه‌ی CPT آن انجام شود. در این صورت چون شبکه یک pt است، برای هر والد یک زیرمسئله‌ی مستقل باقی می‌ماند. کاری که هر زیرمسئله می‌برد متناسب با مجموع اندازه‌های CPT آن است. پس مجموع کار برای $n+1$ گره متناسب با مجموع اندازه‌های CPT آن است. (⇐ خطی است)

polytree : حداکثر یک مسیر بدون جهت بین هر دو گره در شبکه وجود دارد