درس هوش مصنوعی - تمرین تئوری سری 3

مهران مظاهری - 98102346

سوال 1

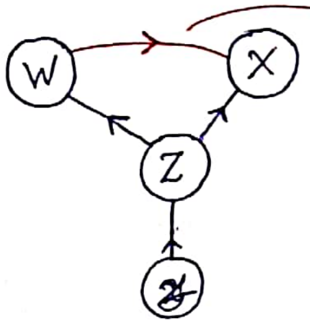
می دانیم اگر در گراف جهت داری ، $\binom{n}{2} + 1$ یال داشته باشیم ، وجود دارد دو راسی که بین آن دو ، دو یال رفت و برگشتی است چراکه گراف کامل بدون جهت $\binom{n}{2}$ یال دارد و این 1 یال اضافه طبق اصل لانه کبوتری حتما بین دو راسی قرار دارد که یال داشته اند؛ پس این گراف جهت دار دور خواهد داشت. پس $\binom{n}{2}$ کران بالایی برای تعداد یال های یک شبکه بیزی است.

برای اینکه نشان دهیم وجود دارد شبکه بیزی با $\binom{n}{2}$ یال، مثال می زنیم:

- شبکه بیزی با رئوس $x_1, \dots, x_n$ داریم که بین رئوس x_i و x_j که $j > i$ باشد یال از x_i به x_j وجود دارد؛

از آنجایی که این گراف دور ندارد (چرا که هیچ مسیری وجود ندارد که از راس با اندیس بزرگتر به راس با اندیس کوچکتر برسیم) پس متعلق به یک شبکه بیزی خواهد بود و شبکه بیزی با $\binom{n}{2}$ یال داریم. (تعداد یال $\binom{n}{2}$ شد چراکه بین هر دو راس دقیقا 1 یال وجود داشت)

سوال 2



برای اینکه شرط استقلال بین W, X نباشد

$y \perp\!\!\!\perp x \mid z$ ✓

$y \perp\!\!\!\perp w \mid z$ ✓

causal chain

سوال 3

$$P(\cap x_i) = \prod_i \left(P(x_i \mid parents(x_i))\right) = P(A)\,P(B|A)\,P(C|A)\,P(D|A)\,P(E|A,D)\,P(F|D)\,P(G|B,C,E,F)$$

سوال 4

$P(B|G,E)$

$$P(B|G,E) = \frac{P(B,G,E)}{P(G,E)}$$

آ) ابتدا CPT ها را تشکیل می دهیم

$P(A)$ $P(B|A)$ $P(C|A,B)$ $\cancel{P(F|A,D)}$ $P(D|A,C)$

$P(E|A,D)$ $P(F|A,E)$ $P(G|B,C)$

باید به ترتیب متغیرهای F ، D ، C ، A را حذف کنیم

F:

$P(F|A,E) \xrightarrow{\sum_F}$ نیازی به این نداریم

D

$P(E|D,A)\,P(D|A,C) \xrightarrow{\sum_D} P(E|A,C)$

C:

$P(E|A,C) \times P(C|A,B) \times P(G|C,B) \xrightarrow[\times]{\cancel{B}} P(E,C,G|A,B) \xrightarrow{\Sigma_C} P(E,G|A,B)$

B:

$P(E,G|A,B) \times P(B|A) \times P(G|B) \xrightarrow{\times} P(E,G,B|A)$

A:

$P(E,G,B|A) \times P(A) \xrightarrow{\times} P(A,B,E,G) \xrightarrow{\Sigma_A} P(E,G,B) \xrightarrow[\text{over } B_{G,E} \,/P(B)_{G,E}]{\text{normalize}} P(E,G|B)$

A:

ب) $\cancel{P(E|A,E)} \times P(A) \times P(F|A,E) \times P(E|A,D) \times P(D|A,C) \times P(C|A,B) \times P(B|A)$

$\xrightarrow{\times} P(A,B,C,D,E,F) \xrightarrow{\Sigma_A} P(B,C,D,E,F)$

F

~~E~~

B

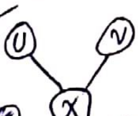
$P(G|E,B) \times P(B,C,D,E,F) \xrightarrow{\times} P(B,C,D,E,F,G) \xrightarrow{\Sigma_C \Sigma_D \Sigma_F} P(B,E,G)$

$P(B|E,G) \leftarrow \text{normalize over } P(B)_{G,E}$

ج) ترتیب قسمت الف بهینه‌تر است چرا که در ب نیاز است که جدول $f(B,C,D,E,F)_{/G}$ ساخته شود که نیاز به محاسبات بیشتر دارد و تعداد بیشتری سطر به جدول احتمال ها اضافه خواهد کرد که بهینه نیست

5 آ) $R \perp\!\!\!\perp Z | U$ ✗ causal chain ~~$U \perp\!\!\!\perp V | S, X$~~

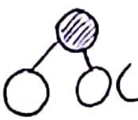
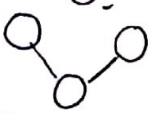
بدلیل ساختار V شکل بین R و S مشروط به J: $R \not\perp\!\!\!\perp S | J$ و بدلیل زنجیره علی از S به V به X به Z ⇐ $S \not\perp\!\!\!\perp Z \Rightarrow R \not\perp\!\!\!\perp Z | J$

ب) $U \perp\!\!\!\perp V | S, X$ ← نادرست بدلیل وجود ساختار V بین  که مسیر فعال است

ج) $W \perp\!\!\!\perp Y | S$ ← درست

مسیر گذرنده از U به دلیل ساختار  غیرفعال است

مسیر گذرنده از X هم به دلیل ساختار  غیرفعال است

د) $W \perp\!\!\!\perp Y | U, V$ ← درست؛ مسیرهایی که از S می‌گذرند به دلیل causal chain بودن مستقل اند و مسیرهایی که از X می‌گذرند بدلیل  و  مستقل اند

ه) $W \perp\!\!\!\perp T \mid U, X, Z$ درست :- مسیر $\underline{WUSVT}$: causal chain ← غیرفعال

- مسیر $\underline{WUXVT}$: common cause ← غیراکتیو - $\underline{WZXVT}$: causal chain ← غیرفعال

همه مسیرها غیرفعال بودند

ⓑ سوال 6

rejection sampling:

$P_1 = 0.5 \,,\, P_{2|1} = 0.8$

P_1	+	−	−	+	+	−	−	+
P_2	+	−	+	−	−	+	+	+
P_3	+	reject	+	reject	reject	−	تردید	−
	✓					✓		✓

$$\Rightarrow P(P_1 \mid P_2, \neg P_3) = \frac{1}{3}$$

likelyhood ~~sampling~~ weighting:

P_1	+	−	−	−	−	−	+	+	+	−	+	−	−	+	−	−	−	−	+	−
W	0.8	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{4}{5}$	$\frac{4}{5}$	$\frac{4}{5}$	$\frac{1}{2}$	$\frac{4}{5}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{4}{5}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{4}{5}$	$\frac{1}{2}$

$$\Rightarrow P(P_1 \mid P_2, \sim P_3) = \frac{7 \times 0.8}{7 \times 0.8 + 13 \times \frac{1}{2}} = \frac{5.6}{12.1} = 0.46$$

Gibbs sampling:

$$P(P_1 \mid P_2, \sim P_3, P_4) = \frac{0.4 \times 0.8}{0.6 \times 0.5 + 0.4 \times 0.8} = 0.516$$

$$P(P_4 \mid P_2, \sim P_3, P_1) = \frac{0.8}{1} = 0.8$$

P_1	1	0	0	1	1	1	0	0	1	1
P_4	1	0	1	1	0	0	1	1	0	1

مقداردهی اولیه

$$\longrightarrow P(P_1 \mid P_2, \sim P_3) = 0.6$$

در rejection sampling ما نیاز به نمونه‌های زیادی برای رسیدن به احتمال درست خواهیم داشت و هرچقدر احتمال شرط کمتر باشد ~~سرعت~~ تعداد نمونه‌های لازم بیشتری شود همچنین در این نمونه‌برداری نیاز به محاسبه احتمال جدیدی هم وجود ندارد در روش likelyhood weighting کم بودن احتمال مشاهدات دیگر تاثیر منفی نخواهد گذاشت و همچنین می‌توانیم در هر بار در هر بار استفاده، نمونه معتبر استخراج کنیم، همچنان این روش تطابق خوبی با داده‌ها و احتمالات دارد؛ در روش gibbs sampling می‌توانیم همه مشاهدات را برای نمونه‌گیری هریک از متغیرها در نظر بگیریم و روش کاملتری محسوب می‌شود اما همانطور که مشاهده می‌شود پیچیدگی محاسباتی دارد به نحوی که در روند محاسبه ما مجبور شدیم یکبار $P(P_1 \mid P_2, \sim P_3, P_4)$ را محاسبه کنیم که معادل محاسبه $P(P_1 \mid P_2, \sim P_3)$ است