به نام خدا

یادگیری تقویتی

نیم‌سال بهار ۰۲-۱۴۰۱

اساتید: دکتر رهبان، آقای حسنی

دانشکده‌ی مهندسی کامپیوتر

تمرین سری صفرم | دوره مفاهیم پیش نیاز | زمان تحویل: ۲۸ بهمن

لطفا نکات زیر را رعایت کنید:

- سوالات خود را از طریق پست مربوط به تمرین در Quera مطرح کنید.
- پاسخ ارسالی واضح و خوانا باشد.
- در هر کدام از سوالات، اگر از منابع خاصی استفاده کرده‌اید، آن را ذکر کنید.
- اگر با افرادی همفکری کرده‌اید، نام ایشان را ذکر کنید.
- پاسخ ارسالی باید توسط خود شما نوشته شده باشد. به اسکرین‌شات از منابع یا پاسخ افراد دیگر نمره‌ای تعلق نمی‌گیرد.
- تمام پاسخ‌های خود را در یک فایل با فرمت RL_HW#_[SID]_[Fullname].zip روی کوئرا قرار دهید.
- برای ارسال هر تمرین تا ساعت ۲۳:۵۹ روز ددلاین فرصت دارید. علاوه بر آن، در هر تمرین می توانید تا سقف سه روز از تأخیر مجاز باقیمانده‌ی خود استفاده کنید.

سوال ۱: واریانس تخمینگر (۳۰ نمره)

برای محاسبه امید ریاضی یک تابع از متغیر تصادفی X، می توان به جای انجام محاسبات ریاضی یا حل با روش های عددی از تخمین Monte-Carlo استفاده کرد. این روش با تولید یک دنباله از متغیر های iid از توزیع X و محاسبه میانگین نمونه ها امید ریاضی مورد نظر را تخمین می زند.
دو تا از خواص مورد توجه تخمینگر های نقطه ای، بایاس و واریانس است. بایاس یک تخمینگر W برای پارامتر θ را تفاوت امید ریاضی W و مقدار واقعی θ تعریف میکنیم. واریانس یک تخمینگر نیز میزان پراکندگی تخمین ها را از میانگین تخمین ها می سنجد. یعنی:

$$Bias_\theta(W) = E_\theta[W] - \theta$$
$$Var_\theta(W) = E_\theta[(W - E_\theta[W])^2]$$

(آ) یک معیار برای ارزیابی تخمینگر ها می تواند میانگین مربعات خطا یعنی $E_\theta[(W-\theta)^2]$ باشد. نشان دهید:

$$MSE(W,\theta) = E_\theta[(W-\theta)^2] = Var_\theta(W) + (Bias_\theta(W))^2$$

یعنی یک تخمینگر برای اینکه خطای کمی داشته باشد باید هم Accuracy بالا (بایاس کم) و هم Precision مناسب (واریانس کم) داشته باشد.

(ب) نشان دهید تخمینگر Monte Carlo یعنی:

$$\hat{\theta}_n = \frac{1}{n}\sum f(X_i) \qquad \text{where} \qquad X_i \overset{\text{iid}}{\sim} P(X)$$

یک تخمینگر بدون بایاس برای $\theta = E[f(X)]$ است

(ج) برای دو تابع $f_1(x) = 1 + (x/2)^2$ و $f_2(x) = (x/4)^{10}$ امید ریاضی را با فرض توزیع یکنواخت روی بازه ی [۰،۴] محاسبه کنید. سپس یک برنامه کوتاه بنویسید که با ۱۰۰ نمونه $E_P[f_1(x)]$ و $E_P[f_2(x)]$ را محاسبه کند. بایاس و واریانس تخمین ها را مقایسه کرده و در صورت تفاوت علت را توضیح دهید. (نیازی به تحویل کد نیست)

(د) حال فرض کنید تخمینگر Monte Carlo برای محاسبه $E_P[f(x)]$ به جای تولید نمونه های iid از توزیع p از توزیع دیگری مانند q نمونه تولید کند. نشان دهید $E_q[\frac{p(x)}{q(x)} * f(x)]$ یک تخمینگر بدون بایاس است. تفاوت واریانس تخمینگر جدید با تخمینگر اولیه را بدست آورید.

(ه) به ازای چه توزیع q(x) ای واریانس تخمینگر جدید کمینه خواهد بود ؟ آیا این توزیع به ازای همه ی توابع f یک توزیع احتمالاتی معتبر است ؟ نشان دهید اگر توزیع q(x) "متناسب" با $p(x)|f(x)|$ باشد تخمینگر جدید کمترین واریانس را به ازای توزیع های q(x) ممکن دارد.

(و) (امتیازی) یک کاربرد مهم تخمینگر جدید می تواند کاهش واریانس تخمین پارامتر هایی از توزیع باشد که در نقاط کم احتمال توزیع اصلی هستند و در نتیجه در تخمینگر Monte Carlo ساده مشاهدات کمی از آن نواحی بررسی می شود.

متغیر تصادفی X با توزیع نرمال استاندارد P را در نظر بگیرید. فرض کنید می خواهیم پارامتر $p = P(X > 5)$ را تخمین بزنیم. مشکل اصلی این است که در تخمینگر مونته کارلو تعداد بسیار زیادی باید نمونه گرفته شود تا داده هایی با $P(X > 5)$ تولید شود. واریانس تخمینگر مونته کارلو $I_{X>5}$ با توزیع $P(x)$ را به صورت تقریبی با واریانس تخمینگری با توزیع ثانویه ی Q(x) مقایسه کنید.

$$Q(x) = \begin{cases} e^{-(x-5)} & x \geq 5 \\ 0 & \text{else} \end{cases} \quad (۱)$$

سوال ۲: زنجیره ی مارکوف (۳۰ نمره)

یک زنجیره مارکوف یک فرآیند تصادفی متشکل از یک مجموعه حالات S و تعداد شمارایی متغیر تصادفی $X_0, X_1, X_2, ...$ است که مقادیر مجموعه S را به خود می گیرند و خاصیت زیر را دارند:

$$P(X_n = s_n | X_{n-1} = s_{n-1}, ..., X_0 = s_0) = P(X_n = s_n | X_{n-1} = s_{n-1})$$

به P ماتریس گذار فرآیند نیز گفته می شود و $P^{(n)}$ (ماتریس P به توان n) نیز ماتریس گذار n مرحله ای فرآیند است.

(آ) نشان دهید توزیع حالات در هر زمان n به طور یکتا توسط توزیع اولیه حالات و ماتریس گذار $P^{(n)}$ مشخص می شود.

(ب) نشان دهید به طور کلی تر توزیع متغیر تصادفی مارکوف در هر زمان با توجه به مشاهدات در دسترس قبلی به آخرین مشاهده بستگی دارد یعنی:

$$P(X_{t1+t2} | X_{F_1}, X_{F_2}, ..., X_{F_n}) = P(X_{t1+t2} | X_{max(F_i)}) \quad \text{where} \quad F \subset \{1, 2, ..., t1\}$$

(ج) نشان دهید اگر مجوعه حالات متناهی باشد داریم:

$$P^{(n+m)} = P^{(n)} P^{(m)}$$

(د) یکی از خواص مورد توجه زنجیره های مارکوف، توزیع استیت ها در بی نهایت است. اگر فرآیند مارکوف شرایط خاصی داشته باشد توزیع استیت ها به یک توزیع حالت پایا (π_{ss}) همگرا خواهد شد یعنی:

$$\lim_{n \to \infty} P^{(0)} P^{(n)} = \pi_{ss}$$

و طبیعتا توزیع حالت پایا در معادله زیر صدق خواهد کرد:

$$\pi_{ss} P = \pi_{ss}$$

n زنجیره مارکوف با ماتریس های گذار P_n و...و P_1 و توزیع های حالت پایای π_n ,..., π_1 را در نظر بگیرید. اگر زنجیره ی مارکوف جدیدی با ماتریس گذار زیر بسازیم، توزیع حالت پایای آن چه توزیع (هایی) است ؟ اگر توزیع حالت پایا یکتا نباشد رفتار بلند مدت زنجیره مارکوف توسط چه عاملی تعیین می شود؟

$$\begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & P_n \end{pmatrix}$$

(ه) (امتیازی) توزیع π_l روی فضای حالات توزیع حدی فرآیند نام دارد اگر:

$$\lim_{n \to \infty} P_{i,j}^{(n)} = \pi_l(j) \quad \forall i, j$$

برای اینکه توزیع حدی وجود داشته باشد باید به ازای تمامی i,j ها حد بالا وجود داشته باشد.

نشان دهید اگر یک توزیع π توزیع حدی فرآیند باشد توزیع حالت پایا نیز می باشد. همچنین نشان دهید اگر توزیع حدی وجود داشته باشد توزیع اولیه حالات تاثیری بر رفتار بلند مدت زنجیره مارکوف ندارد (راهنمایی: از معادله قسمت (ج) استفاده کنید).

سوال ۳: نظریه اطلاعات (۲۷ نمره)

برای یک توزیع احتمالاتی متغیری به نام انتروپی تعریف می شود که بسیاری از خواص آن از لحاظ شهودی با مفهوم اطلاعات همخوانی دارد. همچنین برای بررسی میزان عدم قطعیت یک توزیع نیز می توان انتروپی استفاده کرد. به طور مثال می توان گفت هر چه مشاهدات ما از یک متغیر تصادفی اطلاعات کمتری به ما بدهد، عدم قطعیت این متغیر بیشتر است. انتروپی متغیر های تصادفی گسسته (این مفاهیم قابل تعمیم به متغیر های پیوسته نیز می باشند اما برای سادگی از متغیر گسسته استفاده می کنیم) به شکل زیر تعریف می شود:

$$H(X) = -\sum_x P(x)\log P(x)$$

$$H(X1,...,X_n) = -\sum_{(x_1,...,x_n)} P(x_1,...,x_n)\log P(x_1,...,x_n)$$

(آ) نشان دهید $H(X) \geq 0$

(ب) به طور شهودی از بین توزیع های ممکن روی یک مجموعه X توزیعی که بیشترین عدم قطعیت را دارد کدام است؟ انتروپی آن را محاسبه کنید.

(ج) انتروپی شرطی نیز به شکل زیر تعریف می شود:

$$H(X|Y) = -\sum_{x,y} p(x,y)\log p(x|y)$$

نشان دهید:

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$$

(د) یک ابزار برای مقایسه میزان اطلاعات بین دو توزیع انتروپی نسبی یا فاصله KL است که برای دو توزیع P و Q به شکل زیر تعریف می شود:

$$D_{KL}(P||Q) = \sum_x P(x)\log(\frac{P(x)}{Q(x)})$$

ابتدا به کمک نامساوی ینسن نشان دهید که انتروپی نسبی دو توزیع کمیتی نامنفی است و تنها زمانی صفر می شود که دو توزیع یکسان باشند.

(ه) سپس به کمک نتایج قسمت قبل و توزیعی که در قسمت (ب) حدس زده اید یک حد بالا برای انتروپی یک توزیع روی مجموعه گسسته X بدست آورده و نشان دهید تنها یک توزیع به این انتروپی می رسد.

(و) برای دو متغیر تصادفی یک نوع انتروپی نسبی خاص از اهمیت ویژه ای برخوردار است. اطلاعات متقابل دو متغیر تصادفی X و Y را به شکل زیر تعریف میکنیم.

$$I(X;Y) = D_{KL}(P(X,Y) \;||\; P(X)P(Y))$$

$$I(X;Y|Z) = D_{KL}(P(X,Y|Z) \;||\; P(X|Z)P(Y|Z))$$

(ز) به کمک معیار اطلاعات متقابل نشان دهید که شرطی کردن یک متغیر تصادفی بر اساس متغیر تصادفی دیگری (اطلاعات جدید) انتروپی آن را افزایش نمی دهد.

(ح) نامساوی های زیر را ثابت کنید و شرط تساوی را نیز بیان کنید:

۱. $I(X,Y;Z) \geq I(X;Z)$

۲. $H(X,Y|Z) \geq H(X|Z)$

(ط) (امتیازی) بر خلاف انتروپی، شرطی کردن کردن بر اساس یک متغیر دیگر (اطلاعات جدید) می تواند اطلاعات متقابل را کاهش یا افزایش دهد. اگر X،YوZ متغیر های تصادفی باینری باشند توزیع های مشترک P(X،Y،Z) مثال بزنید که تحت آنها:

۱. $I(X;Y) = 0$ و $I(X;Y|Z) = 1$

۲. $I(X;Y) = 1$ و $I(X;Y|Z) = 0$

سوال ۴: مدل های احتمالاتی و متغیر های نهان (۳۳ نمره)

(آ) نشان دهید کمینه کردن فاصله توزیع مدل KL با توزیع تجربی داده ها معادل تخمین بیشینه درست نمایی است.

(ب) مدل احتمالاتی با متغیر های قابل مشاهده X و متغیر های نهان Z را در نظر بگیرید. فرض کنید X از یک توزیع نامعلوم با پارامتر های θ بدست می آید.
با استفاده از نامساوی ینسن نشان دهید برای هر توزیع فرضی دلخواه Q(Z) باند پایین زیر برای تابع لایکلیهود برقرار است:

$$l(\theta; X) = \log P_\theta(X) \geq \mathcal{L}[\theta, Q] = E_{Q(Z)}[\log P_\theta(X, Z)] + H[Q]$$

(ج) نشان دهید با انتخاب $Q(Z) = P_\theta(Z|X)$ فاصله تابع لایکلیهود و باند پایین آن کمینه خواهد شد

(د) روش Expectation-Maximization با توجه به نتیجه بالا در هر مرحله ابتدا با پارامتر های θ^t توزیع $P_\theta(Z|X)$ را محاسبه می کند و سپس با توجه به توزیع بدست آمده با بیشینه کردن باند پایین پارامتر ها را به θ^{t+1} آپدیت میکند. نقطه تعادل این الگوریتم همان بیشینه تابع لایکلیهود است.
با استفاده از قانون بیز توضیح دهید که چرا ممکن است توزیع $P_\theta(Z|X)$ قابل محاسبه نباشد.

(ه) برای حل این مشکل می توان توزیع $P_\theta(Z|X)$ را نیز با توزیع $Q_\phi(Z|X)$ تخمین زد. ابتدا نشان دهید که باند پایین قسمت (ج) را می توان به صورت زیر بازنویسی کرد:

$$L[\theta, \phi] = E_{Q_\phi(Z|X)}[\log P_\theta(X|Z)] - D_{KL}(Q_\phi(Z|X) \ || \ P_\theta(Z))$$

(و) این باند پایین اساس آموزش VAE ها است. به طور خلاصه مراحل آموزش VAE را توضیح دهید. بهینه سازی عبارت اول باعث افزایش قدرت بازسازی تصاویر ورودی از روی بازنمایی پنهان آنها می شود. اثر بیشینه کردن عبارت دوم از نظر مفهومی چیست ؟

(ز) یکبار دیگر پاند پایین را به صورت زیر بازنویسی کنید:

$$L[\theta, \phi] = E_{Q_\phi(Z|X)}[\log P_\theta(X|Z)] - D_{KL}(Q_\phi(Z) \ || \ P_\theta(Z)) - I_{Q_\phi}(X; Z)$$

(ح) (امتیازی) تعدادی از مشکلات رایج در آموزش VAE ها را مطالعه و به صورت خلاصه ذکر کنید. سپس با بررسی مفهومی ترم دوم و سوم عبارت بالا و مقایسه این دو با یکدیگر سعی کنید مشکلات ذکر شده را توجیه کنید.