به نام خدا

یادگیری ماشین — نیمسال اول ۹۸-۹۷

تمرین سری اول

موعد تحویل: ۴ مهر — نمره: ۱۲۰

سوال ۱: Probability and Statistics (35 points)

۱.۱. (۵ نمره) متغیرهای تصادفی X_1 و X_2 را با چگالی احتمال توامان [1] زیر در نظر بگیرید که در آن $k > 0$:

$$f_{X_1,X_2}(x_1,x_2) = kx_1^2 x_2 \mathbb{1}(x_1 > 0 \wedge x_2 > 0 \wedge x_1 + x_2 \leq 1)$$

(آ) (۲ نمره) آیا X_1 و X_2 مستقل هستند؟ برای پاسخ خود دلیل بیاورید.

(ب) (۳ نمره) مقدار k را محاسبه نمایید.

۲.۱. (۲ نمره) فرض کنید ۵ درصد از مردان و ۰٫۲۵ درصد از زنان مشکل کوررنگی دارند. فردی به تصادف انتخاب می‌شود و مشخص می‌شود این فرد کوررنگی دارد. این فرد با چه احتمالی مرد است؟ (فرض کنید تعداد زنان و مردان برابر است.)

۳.۱. (۳ نمره) x و y متغیرهای تصادفی یکنواخت در بازه $(0, \alpha)$ و مستقل از هم هستند. تابع چگالی احتمال [2] متغیر تصادفی $z = |x-y|$ را بیابید.

۴.۱. (۷ نمره) x و y متغیرهای تصادفی یکنواخت در بازه $(0, 1)$ و مستقل از هم هستند. متغیرهای تصادفی w و z را به این صورت تعریف می‌کنیم:

$$w = max(x,y)$$
$$z = min(x,y)$$

تابع چگالی احتمال را برای موارد زیر به دست آورید:

(آ) (۳نمره) $r = w - z$

(ب) (۴ نمره) $s = w + z$

۵.۱. (۱۱ نمره) برای هر دو متغیر تصادفی دلخواه X و Y ثابت کنید:

(آ) (۴ نمره) $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$

(ب) (۴ نمره) $var(X) = \mathbb{E}[var(X|Y)] + var(\mathbb{E}[X|Y])$

(ج) (۳ نمره) $Cov(aX+b, cY+d) = acCov(X,Y)$

۶.۱. (۷ نمره) X و Y دو متغیر تصادفی هستند.

(آ) (۳ نمره) ثابت کنید اگر X و Y مستقل باشند آنگاه ناهمبسته [3] نیز هستند.

(ب) (۴ نمره) مثالی بزنید که در آن X و Y ناهمبسته باشند ولی مستقل نباشند.

---

[1] Joint probability density function
[2] Probability density function
[3] Uncorrelated

سوال ۲: Estimation (45 points)

۱،۲. (۵ نمره) فرض کنید $X_1, X_2, ..., X_N$ نمونه‌های تصادفی مستقل از یک توزیع با چگالی $f_{\theta,j}$ باشند که $\theta > 0$ و $j = 1, 2$ است. اگر $f_{\theta,1}(x) = \mathcal{N}(x|0, \theta)$ و $f_{\theta,2}(x) = (2\theta)^{-1} e^{-\frac{|x|}{\theta}}$ باشد، MLE را برای پارامترهای (θ, j) پیدا کنید.

۲،۲. (۴نمره) فرض کنید $X_1, X_2, ..., X_N$ نمونه‌های تصادفی مستقل هستند از یک توزیع با چگالی زیر:

$$f(x|\theta) = \theta x^{-2} \quad , 0 < \theta \leq x < \infty$$

MLE را برای پارامتر θ پیدا کنید.

۳،۲. (۵نمره) فرض کنید $X_1, X_2, ..., X_N$ نمونه‌های تصادفی مستقل هستند از یک توزیع با چگالی زیر:

$$f(x|\theta) = \frac{1}{2} e^{-|x-\theta|} \quad , -\infty < x < \infty, -\infty < \theta < \infty$$

MLE را برای پارامتر θ پیدا کنید.

۴،۲. (۴نمره) فرض کنید $X_1, X_2, ..., X_N$ نمونه‌های تصادفی مستقل از توزیع یکنواخت در بازه $(\theta, \theta + 1)$ هستند که در آن مقدار پارامتر θ عددی نا مشخص $(-\infty < \theta < \infty)$ است. آیا MLE برای پارامتر θ یکتاست؟ برای پاسخ خود دلیل بیاورید.

۵،۲. (۵ نمره) عمر مفید یک لامپ (بازه زمانی شروع به کار لامپ تا قبل از سوختن آن)، یک متغیر تصادفی با تابع چگالی احتمال $ce^{-cx}U(x)$ است. در آزمایشی ۸۰ لامپ را مورد بررسی قرار دادیم و دیدیم ۲۰۰ ساعت پس از شروع آزمایش ۶۲ لامپ همچنان روشن بودند. تخمین ML را برای پارامتر c بیابید.

۶،۲. (۸ نمره) تابع چگالی احتمال توامان زیر را در نظر بگیرید:

$$f_{X,Z}(x,z) = \begin{cases} e^{-z}, & 0 \leq x \leq z \\ 0, & \text{otherwise} \end{cases}$$

(آ) (۱ نمره) x و y ای را بیابید که این تابع چگالی را بیشینه کند.

(ب) (۱ نمره) تابع چگالی حاشیه [۴] $f_Z(z)$ و مقدار z که این تابع را بیشینه می‌کند را بیابید.

(ج) (۳ نمره) متغیر جدیدی به این تابع اضافه می‌کنیم و تابع چگالی احتمال توامان جدید را به این صورت تعریف می‌کنیم:

$$f_{X,Z,Y}(x,z,y) = \begin{cases} y^2 e^{-yz}, & 0 \leq x \leq z, \ 1 \leq y \leq 2 \\ 0, & \text{otherwise} \end{cases}$$

به شرط داشتن مشاهده $Y = y$، تخمین توامان MAP را برای X و Z بیابید.

(د) (۳ نمره) با توجه به قسمت قبل، تابع $f_{Z|Y}(z|y)$، چگالی حاشیه‌ای Z به شرط $Y = y$، را بیابید. سپس تخمین MAP را برای Z به شرط $Y = y$ پیدا کنید.

۷،۲. (۱۴ نمره) تاسی داریم که در یک آزمایش آن را N بار پرتاب کرده‌ایم. در این N پرتاب N_1 بار عدد ۱ ، N_2 بار عدد ۲ و به همین ترتیب N_6 بار عدد ۶ ظاهر شده است. همچنین احتمال آمدن اعداد ۱ تا ۶ را نیز به ترتیب با θ_1 و ... و θ_6 نمایش می‌دهیم.

(آ) (۳ نمره) فرض کنید متغیری که عدد تاس را نشان می‌دهد، از توزیع $Multinomial$ پیروی می‌کند. MLE را برای پارامترهای θ_1 تا θ_6 به دست آورید.

(ب) (۳ نمره) نشان دهید توزیع $Dirichlet$ یک $ConjugatePrior$ برای توزیع $Multinomial$ است.

(ج) (۳ نمره) با در نظر گرفتن توزیع پیشین $Dirichlet$ با پارامترهای α_1 تا α_6 بر روی پارامترهای θ_1 تا θ_6، پارامترهای θ را این بار با استفاده از MAP تخمین بزنید.

(د) (۵ نمره ــ اختیاری) با در نظر گرفتن قسمت قبل، با استفاده از رویکرد بیزین مقدار $P(x|D)$ را به دست بیاورید.

[۴] marginal density function

سوال ۳: Linear Algebra and Matrix Calculus (30 points)

۱٫۳. (۱۲ نمره) اگر a و x بردارهای ستونی و A ماتریس مربعی باشد موارد زیر را ثابت کنید:

(آ) $\frac{da^\top x}{dx} = \frac{dx^\top a}{dx} = a^\top$

(ب) $\frac{dx^\top x}{dx} = ۲x^\top$

(ج) $\frac{d(x^\top a)^۲}{dx} = ۲x^\top aa^\top$

(د) $\frac{dAx}{dx} = A$

(ه) $\frac{dx^\top A}{dx} = A^\top$

(و) $\frac{dx^\top Ax}{dx} = x^\top(A + A^\top)$

۲٫۳. (۴ نمره) اگر α را به صورت زیر تعریف کنیم:

$$\alpha = y^\top x$$

که در آن x و y دو بردار ستونی $n \times ۱$ و هر دو تابعی از بردار z هستند، آنگاه ثابت کنید:

$$\frac{\partial \alpha}{\partial z} = x^\top \frac{\partial y}{\partial z} + y^\top \frac{\partial x}{\partial z}$$

۳٫۳. (۴ نمره) ماتریس H را در نظر بگیرید. مقادیر ویژه و بردارهای ویژه را برای ماتریس $HH^\top$ به دست آورید.

$$H = \begin{bmatrix} ۱ & ۲ & ۰ \\ ۰ & ۱ & ۱ \end{bmatrix}$$

۴٫۳. (۵ نمره) اگر A یک ماتریس مربعی و λ یک eigenvalue برای ماتریس A باشد، ثابت کنید λ یک eigenvalue برای ماتریس $A^\top$ است.

۵٫۳. (۵ نمره) فرض کنید S مجموعه‌ای متعامد [۵] از بردارهای ناصفر باشد. ثابت کنید S مستقل خطی است.
(تعریف مجموعه متعامد از بردارها: فرض کنید $S = \{u_۱, u_۲, u_۳, ..., u_n\}$ مجموعه ای از بردارهای $\mathbb{C}^m$ باشد. S متعامد است اگر هر دو بردار غیر یکسان از S بر هم عمود باشند، یعنی $i \neq j$ ، $< u_i, u_j >= ۰$)

سوال ۴: Lagrange Multiplier (10 points)

۱٫۴. با استفاده از ضریب لاگرانژ مقدار بیشینه و کمینه سراسری تابع $f(x, y) = x^۲ + ۲y^۲ - ۴y$ را با توجه به محدودیت $x^۲ + y^۲ = ۹$ به دست آورید.

[۵] orthogonal set