模式识别
作业一

181220076, 周韧哲, 本科，人工智能学院，人工智能学院选课

2021 年 6 月 26 日

Problem 1

(a) 需要令 $\frac{8a-1}{3} \geq 0$，即 $a \geq \frac{1}{8}$。

(b) 等于 1。

(c) 当 $a = \frac{1}{2}$，上式为 1。我认为上式恒等于 1。

(d) 返回值为 $1.2182 + 0.1260i$。

(e) 原因是使用 ˆ(1/3) 命令时会调用 pow(A,B)，当 B 的绝对值（这里是 $\frac{1}{3}$）小于 1 时会返回复数根。使用 nthroot(A,B) 命令可以得到正确结果为 1。多个 a 的取值支持了我的想法。

(f) 证明：先令 $x = \sqrt{\frac{8a-1}{3}}$，则 $a = \frac{3x^2+1}{8}$，代入原式，得到

$$\begin{aligned}
&\sqrt[3]{\frac{3x^2+1}{8} + \frac{\frac{3x^2+1}{8}+1}{3}x} + \sqrt[3]{\frac{3x^2+1}{8} - \frac{\frac{3x^2+1}{8}+1}{3}x} \\
&= \sqrt[3]{\frac{1}{8}(x^3+3x^2+3x+1)} + \sqrt[3]{\frac{1}{8}(-x^3+3x^2-3x+1)} \\
&= \sqrt[3]{\frac{1}{8}(x+1)^3} + \sqrt[3]{\frac{1}{8}(1-x)^3} \\
&= 1
\end{aligned}$$

(g) 当 $a = 2$ 时，此表达式刚好符合上述定理，所以等于 1。

(h) 由 Cardano's formula 知当 $4p^3 + 27q^2 > 0$ 时，此式有唯一实数根

$$\sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

当 $p = 2a-1, q = -2a$ 时，方程变为 $z^3 + (2a-1)z - 2a = 0$，容易看出其有唯一实数根 1，且将 p, q 代入上面的实数根表达式后恰好得到本题关于 a 的表达式，所以本题中的表达式等于 1。

Problem 2

(a) $x_\perp = \frac{x^Ty}{y^Ty}y = \frac{2\sqrt{3}}{4}\cdot(1,\sqrt{3})^T = (\frac{\sqrt{3}}{2},\frac{3}{2})^T$。

(b) $x - x_\perp = (\frac{\sqrt{3}}{2}, -\frac{1}{2})^T$，所以 $(x-x_\perp)^Ty = \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} = 0$，所以 $y \perp (x - x_\perp)$。

(c) 如下图所示。

(d) 由提示得 $\|x - x_\perp\|^2 \le \|x - \lambda y\|^2$，所以 $\|x-x_\perp\| \le \|x-\lambda y\|$，等号取得条件为 $\|x_\perp - \lambda y\| = 0$，即 $\lambda = \frac{\sqrt{3}}{2}$。

Problem 3

(a) 实对称矩阵正定的充要条件是特征值都是正数，所以 X 为正定矩阵的充要条件为 $x > 0$。

(b) $det(X)$ 等于 X 的特征值乘积，所以 $12x = 72$，$x = 6$。

Problem 4

(a) $\mathbb{E}[X] = \int_{-\infty}^{\infty} p(x)xdx = \int_0^\infty \beta e^{-\beta x}xdx = -\int_0^\infty xde^{-\beta x} = \int_0^\infty e^{-\beta x}dx - xe^{-\beta x}|_0^\infty = \frac{1}{\beta}$，
$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} p(x)x^2dx = -\int_0^\infty x^2de^{-\beta x} = 2\int_0^\infty xe^{-\beta x}dx - x^2e^{-\beta x}|_0^\infty = \frac{2}{\beta^2}$，所以
$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{1}{\beta^2}$

(b) 易知 $x < 0$ 时，$F(x) = 0$；当 $x \ge 0$ 时，$F(x) = \int_0^x p(t)dt = \int_0^x \beta e^{-\beta t}dt = 1 - e^{-\beta x}$

(c) 贝叶斯公式：$Pr(X \ge a+b|X \ge a) = \frac{Pr(X\ge a+b)}{Pr(X \ge a)} = \frac{1-F(a+b)}{1-F(a)} = e^{-\beta b} = 1 - F(b) = Pr(X \ge b)$

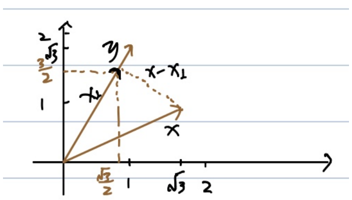
(d) 易得期望寿命为 $\mathbb{E}[X] = \frac{1}{10^{-3}} = 10^3$。由于指数分布的无记忆性，2000 小时后的能再工作 b 小时的概率 $Pr(X \ge b+2000|X \ge 2000) = Pr(X \ge b)$，故剩余寿命的期望不变，仍为 10^3。

Problem 5

(a) 因为 $f''(x) = a^2e^{ax} \ge 0$，所以 $f(x)$ 为凸函数。

(b) 因为 $f''(x) = -\frac{1}{x^2} < 0$，所以 $f(x)$ 为凹函数。

(c) $x > 0$ 时，$f''(x) = \frac{1}{x} > 0$，为凸函数。扩充定义域到 $[0,\infty)$ 后，对于 $0 \le \lambda \le 1$，当 $x = y = 0$ 有 $f(\lambda x + (1-\lambda)y) = 0 = \lambda f(x) + (1-\lambda)f(y)$ 成立；当其中一个不为 0，不妨令 $x = 0$，则 $f(\lambda x + (1-\lambda)y) = (1-\lambda)y\ln((1-\lambda)y) \le (1-\lambda)y\ln y = \lambda f(x) + (1-\lambda)f(y)$，所以 $f(x)$ 在 $[0,\infty)$ 为凸函数。



(d) 约束为 $g(p_1,\cdots,p_n)=\sum_{i=1}^{n}p_i=1$，令 $f(p_1,\cdots,p_n)=-H=\sum_{i=1}^{n}p_i\log p_i$，则问题成为

$$\begin{aligned}&\min f(p_1,\cdots,p_n),\\&\text{s.t. } g(p_1,\cdots,p_n)=1\end{aligned}$$

使用拉格朗日乘子法, 拉格朗日函数 $L(p_1,\cdots,p_n,\lambda)=f(p_1,\cdots,p_n)+\lambda(g(p_1,\cdots,p_n)-1)$, 对各变量求导得

$$\begin{aligned}\frac{\partial L}{\partial p_i}&=\log p_i+1+\lambda=0,\quad i=1,\cdots,n\\\frac{\partial L}{\partial \lambda}&=\sum_{i=1}^{n}p_i-1=0\end{aligned}$$

解得 $\lambda=\log n-1,\quad p_i=\frac{1}{n},i=1,\cdots,n$。所以最大化熵的 p_i 都相等且为 $\frac{1}{n}$。