Homework 0

6501-003 Reinforcement Learning (Spring 2025)

This problem set is for self-assessment only—no submission required. It tests your foundation in high-school math, calculus, probability, and linear algebra.

Please try to work through these problems independently first. While you may look up concepts and formulas when stuck, avoid searching for complete solutions. If you struggle with more than 1 problem even after reviewing the concepts, please either consult with me or strengthen the prerequisites before taking the course.

P.S. These are not random questions, but the exact arguments we will use in the course.

1. Prove that for any $n \in \mathbb{N}$, $2\sqrt{n+1} - 2 < 1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots \frac{1}{\sqrt{n}} < 2\sqrt{n}$.

   Relating the series sum with the area under the curve $y = \frac{1}{\sqrt{x}}$ allows us to bound

   $$\int_1^{n+1} \frac{1}{\sqrt{x}} \mathrm{d}x < \sum_{i=1}^{n} \frac{1}{\sqrt{i}} < \int_0^n \frac{1}{\sqrt{x}} \mathrm{d}x.$$

   Evaluating the two integrals (recall that $\int \frac{1}{\sqrt{x}} \mathrm{d}x = 2\sqrt{x} + C$) leads to the desired inequality.

2. Find $a, b > 0$ that minimizes $\frac{1}{a^2} + \frac{a}{b} + b$. Justify your answer.

   By the AM-GM inequality (`https://artofproblemsolving.com/wiki/index.php/AM-GM_Inequality`),

   $$\frac{1}{5}\left(\frac{1}{a^2} + \frac{a}{b} + b\right) = \frac{1}{5}\left(\frac{1}{a^2} + \frac{a}{2b} + \frac{a}{2b} + \frac{b}{2} + \frac{b}{2}\right) \geq \sqrt[5]{\frac{1}{a^2} \cdot \frac{1}{2b} \cdot \cdot \frac{1}{2b} \cdot \frac{b}{2} \cdot \frac{b}{2}} = 2^{-\frac{4}{5}}.$$

   The inequality holds when $\frac{1}{a^2} = \frac{a}{2b} = \frac{b}{2} \Rightarrow b = 2^{\frac{1}{5}}$ and $a = 2^{\frac{2}{5}}$.

3. Let $x_1, \ldots, x_n \in \mathbb{R}^d$ and $y_1, \ldots, y_n \in \mathbb{R}$. Let $\lambda > 0$ and define

   $$F(\theta) = \sum_{i=1}^{n} (x_i^\top \theta - y_i)^2 + \lambda \|\theta\|_2^2.$$

   Prove that

   $$\theta = \left(\lambda I + \sum_{i=1}^{n} x_i x_i^\top\right)^{-1} \left(\sum_{i=1}^{n} x_i y_i\right)$$

   minimizes $F(\theta)$.

   $$\nabla_\theta F(\theta) = \sum_{i=1}^{n} 2x_i(x_i^\top \theta - y_i) + 2\lambda\theta = 2\left(\lambda I + \sum_{i=1}^{n} x_i x_i^\top\right)\theta - 2\sum_{i=1}^{n} x_i y_i.$$

   Solving for θ under $\nabla_\theta F(\theta) = 0$ gives the desired θ.

4. Let $z_1, \dots, z_n$ be independent variables following the standard normal distribution $\mathcal{N}(0,1)$, and let $x_1, \dots, x_n \in \mathbb{R}^d$ be fixed vectors. Define $X = \sum_{i=1}^n z_i x_i$. Prove that the covariance matrix of X, i.e., $\mathbb{E}[XX^\top]$, is equal to $\sum_{i=1}^n x_i x_i^\top$.

$$
\begin{aligned}
\mathbb{E}\left[XX^\top\right] &= \mathbb{E}\left[\left(\sum_{i=1}^n z_i x_i\right)\left(\sum_{i=1}^n z_i x_i^\top\right)\right] \\
&= \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^n z_i z_j x_i x_j^\top\right] \\
&= \sum_{i=1}^n \mathbb{E}[z_i^2] x_i x_i^\top && (\mathbb{E}[z_i z_j] = 0 \text{ for } i \neq j) \\
&= \sum_{i=1}^n x_i x_i^\top. && (\mathbb{E}[z_i^2] = 1)
\end{aligned}
$$

5. Let $A, B \in \mathbb{R}^{d\times d}$ be symmetric positive definite matrices such that $A \succ B$. Prove that $B^{-1} \succ A^{-1}$.
(Recall: $A \succ B$ means that for any $x \in \mathbb{R}^d \backslash \{0\}$, $x^\top (A - B)x > 0$.)

By the condition $A \succ B$, we have for any $x \neq 0$, $x^\top (B^{-\frac{1}{2}} AB^{-\frac{1}{2}} - I)x = x^\top B^{-\frac{1}{2}} (A - B)B^{-\frac{1}{2}} x > 0$. This implies that all eigenvalues of $B^{-\frac{1}{2}} AB^{-\frac{1}{2}}$ are larger than 1.

Notice that $B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}$ is the inverse of $B^{-\frac{1}{2}} AB^{-\frac{1}{2}}$. Therefore, all eigenvalues of $B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}$ are smaller than 1.

Therefore, for any $x \neq 0$, $x^\top (B^{-1} - A^{-1})x = x^\top B^{-\frac{1}{2}} \left(I - B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}\right) B^{-\frac{1}{2}} x > 0$.

6. Let $y \in \Delta_d$ and $\theta \in \mathbb{R}^d$, where Δ_d is the d-dimensional probability space defined as $\{x \in \mathbb{R}^d : \sum_{i=1}^d x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i\}$. Let $F(x) = x^\top \theta + \mathrm{KL}(x, y)$. Prove that the minimizer of $F(x)$ over $x \in \Delta_d$ satisfies

$$
x_i = \frac{y_i e^{-\theta_i}}{\sum_{j=1}^d y_j e^{-\theta_j}}.
$$

Recall: the KL divergence between two distributions $x, y \in \Delta_d$ is defined as $\mathrm{KL}(x, y) = \sum_{i=1}^d x_i \log \frac{x_i}{y_i}$.

We solve the minimization of $F(x)$ over $x \in \Delta_d$ by finding a solution that satisfies the KKT condition. The Lagrangian function of the problem can be written as

$$
\mathcal{L}(x, \lambda, \mu) = x^\top \theta + \mathrm{KL}(x, y) + \lambda \left(\sum_{i=1}^d x_i - 1\right) + \sum_{i=1}^d \mu_i x_i
$$

with $\lambda \in \mathbb{R}$ and $\mu_i \leq 0$. The first-order stationarity condition requires

$$
\nabla_{x_i} \mathcal{L}(x, \lambda, \mu) = \theta_i + \log \frac{x_i}{y_i} + 1 + \lambda + \mu_i = 0
$$

for all i, and the complementary slackness condition requires

$$
\mu_i x_i = 0
$$

for all i.

The first condition gives $x_i = y_i \exp\left(-\theta_i - 1 - \lambda - \mu_i\right)$. Combining this with the condition $\mu_i x_i = 0$ we get $\mu_i = 0$ for all i. Thus, $x_i = y_i \exp(-\theta_i - 1 - \lambda)$ for some λ. To make the solution feasible, we need $\sum_i x_i = 1$. This implies that the $\exp(-1 - \lambda)$ factor is simply a normalization factor, and we have

$$
x_i = \frac{y_i \exp(-\theta_i)}{\sum_{j=1}^d y_j \exp(-\theta_j)}.
$$

7. Let $\pi_\theta(x)$ be a function of x parameterized by θ. Prove that

$$\nabla^2_\theta \log \pi_\theta(x) = \frac{\nabla^2_\theta \pi_\theta(x)}{\pi_\theta(x)} - (\nabla_\theta \log \pi_\theta(x))(\nabla_\theta \log \pi_\theta(x))^\top.$$

We first compute the first-order derivative.

$$\nabla_\theta \log \pi_\theta(x) = \frac{\nabla_\theta \pi_\theta(x)}{\pi_\theta(x)} \qquad \text{(by } \tfrac{\mathrm{d}}{\mathrm{d}y}(\log y) = \tfrac{1}{y} \text{ and chain rule)}$$

Then the second derivative:

$$\nabla^2_\theta \log \pi_\theta(x) = \nabla_\theta \left(\frac{\nabla_\theta \pi_\theta(x)}{\pi_\theta(x)} \right) = \frac{\pi_\theta(x) \nabla^2_\theta \pi_\theta(x) - (\nabla \pi_\theta(x))(\nabla \pi_\theta(x))^\top}{(\pi_\theta(x))^2}.$$

The last expression can be further simplified as

$$\frac{\nabla^2_\theta \pi_\theta(x)}{\pi_\theta(x)} - (\nabla_\theta \log \pi_\theta(x))(\nabla_\theta \log \pi_\theta(x))^\top$$

using the fact that $\nabla_\theta \log \pi_\theta(x) = \frac{\nabla_\theta \pi_\theta(x)}{\pi_\theta(x)}$.