Quiz 3 Practice Problems

10-606

October 4, 2025

1 Matrix Calculus and Optimization

1. Let $f(x) = (Ax + b)^\top (Ax + b)$ where $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$.
   (a) Compute $\nabla_x f(x)$.
   (b) Compute the Hessian $H(f) = \nabla_x^2 f(x)$.
2. Let $X \in \mathbb{R}^{n \times m}$ and $Y \in \mathbb{R}^{m \times n}$.
   (a) The trace of a square matrix A, denote $\mathrm{Tr}(A)$, is the sum of its diagonal elements. Show that $\mathrm{Tr}(XY) = \sum_{i=1}^n \sum_{j=1}^n x_{ij} y_{ji}$.
   (b) Show that $\nabla_X \mathrm{Tr}(XY) = Y^\top$.
3. For $\theta, y \in \mathbb{R}^K$ such that $\sum_{i=1}^K y_i = 1$, let

$$f(\theta) = -\sum_{i=1}^{K} y_i \log \left(\frac{e^{\theta_i}}{\sum_{j=1}^K e^{\theta_j}} \right).$$

   What is $\nabla_\theta f(\theta)$?
4. Let $h(w) = \log\big(1 + \exp(-y\, w^\top x)\big)$ with $x \in \mathbb{R}^d$ and $y \in \{-1, 1\}$. Compute $\nabla_w h(w)$.
5. For a square matrix $A \in \mathbb{R}^{n \times n}$ and vector $x \in \mathbb{R}^n$, let $f(x) = \frac{x^\top A x}{x^\top x}$. Compute $\nabla_x f(x)$.
6. Let $f_1, \dots, f_n$ be convex functions from $\mathbb{R}^n$ to $\mathbb{R}$.
   (a) Show that αf_1 is convex for any scalar $\alpha \geq 0$.
   (b) Show that $\sum_i \alpha_i f_i$ is convex for nonnegative real numbers α_i.
   (c) Is $\sum_i \alpha_i f_i$ for arbitrary real numbers α_i?

7. Let $f(x) = \frac{1}{2}x^\top Qx - c^\top x$ with

$$Q = \begin{bmatrix} k & 1 \\ 1 & k \end{bmatrix}, \qquad c = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

for some $k \in \mathbb{R}$.

   (a) Compute $\nabla f(x)$.
   (b) From $x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, take one gradient descent step with step size $\eta = 0.2$. What is x_1?

2 Probability

1. You roll two fair dice.
   (a) $P(\text{sum} = 7)$
   (b) $P(\text{at least one six})$
   (c) Are the events in (a) and (b) independent?
2. From a standard 52-card deck, draw two cards without replacement.
   (a) What is the probability both are hearts?
   (b) What is $P(\text{second is a heart} \mid \text{first is red})$?
3. Let $X \sim \text{Bernoulli}(p)$ (i.e., X is 1 with probability p, 0 with probability $1-p$) and define $Y = 1 - X$.
   (a) Write the joint pmf $P(X, Y)$.
   (b) Are X and Y independent?
4. Flip a fair coin three times. Let X be the number of heads, and $Y = \mathbf{1}\{\text{first flip is H}\}$.
   (a) Write the joint pmf $P(X, Y)$.
   (b) Compute $P(X = 3 \mid Y = 1)$.
5. Show that

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

if $P(B) \neq 0$. This is known as Bayes' theorem.

6. An urn contains 5 red, 3 blue, and 2 green balls. Two balls are drawn *without replacement*. What is:

   (a) The probability both are red.

   (b) The probability the two balls are the same color.

   (c) The probability the second ball is blue, given the first was green.

7. This is a classic phenomenon known as the "birthday paradox." Suppose there are 23 people in a room. Let's assume their birthdays are uniformly distributed across the days of the year. We're going to show that the probability that two people share a birthday is more than 50%.

   (a) Let A be the event that some two people in the room share a birthday. Express this probability in terms of the chances that no two people share a birthday.

   (b) Label the 23 people from 1 to 23. Let E_k be the event that person k does not share a birthday with person 1 through $k-1$. What is $P(E_k | E_1 \cap E_2 \cap \cdots \cap E_{k-1})$?

   (c) Express $P(\cap_{i=1}^{23} E_i)$ as a product of terms that look like $P(E_k | \cap_{j<k} E_j)$. (Recall that $\cap_{j<k} E_j$ is just a concise way of writing $E_1 \cap E_2 \cap \cdots \cap E_{k-1}$.)

   (d) What is $P(A)$?