## Problem 1

Let $A$ be an $m \times n$ real matrix (where $m \geq n$). Let us prove the existence of a singular value decomposition of $A$:

$$U^\top AV = \Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_n \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}, \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

in accordance with the following procedure. In this equation, $^\top$ denotes the matrix transpose, $\mathbf{0}$ a zero vector with suitable dimension, $\Sigma$ an $m \times n$ real matrix, $U$ an $m \times m$ orthogonal matrix, and $V$ an $n \times n$ orthogonal matrix. The diagonal elements of $\Sigma$, $\sigma_i$ $(i = 1, \ldots, n)$, are called singular values.

(1) Define a function $\sigma(A)$ of a real matrix $A$ by

$$\sigma(A) = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\|,$$

where $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$ is the Euclid norm of $\boldsymbol{x}$. Assume the existence of $\boldsymbol{x}$ that attains the maximum in the above equation.

Show that this function satisfies the following properties of a matrix norm. That is, for any $m \times n$ real matrices $A, B$ and any real number $\alpha$,

1. $\sigma(A) \geq 0$, and $\sigma(A) = 0 \Leftrightarrow A = 0$,
2. $\sigma(\alpha A) = |\alpha|\sigma(A)$,
3. $\sigma(A + B) \leq \sigma(A) + \sigma(B)$.

(2) Define $\sigma_1 = \sigma(A)$. From the above assumption there exists a pair of unit vectors $\boldsymbol{x}_1$ and $\boldsymbol{y}_1$ that satisfy $\sigma_1 \boldsymbol{y}_1 = A\boldsymbol{x}_1$. We can construct a pair of orthonormal bases $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ and $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_m\}$. Let us define an $m \times m$ orthogonal matrix $U_1 = (\boldsymbol{y}_1\ \boldsymbol{y}_2\ \cdots\ \boldsymbol{y}_m)$ and an $n \times n$ orthogonal matrix $V_1 = (\boldsymbol{x}_1\ \boldsymbol{x}_2\ \ldots\ \boldsymbol{x}_n)$. Then we have

$$A_1 = U_1^\top AV_1 = \begin{pmatrix} \alpha & \boldsymbol{w}^\top \\ \boldsymbol{z} & A_2 \end{pmatrix}. \qquad (*)$$

Show $\alpha = \sigma_1$ and $\boldsymbol{z} = \mathbf{0}$.

(3) For $A_1$ in equation $(*)$ (where $\alpha = \sigma_1, \boldsymbol{z} = \mathbf{0}$), show

$$\left\| A_1 \begin{pmatrix} \sigma_1 \\ \boldsymbol{w} \end{pmatrix} \right\| \geq \sigma_1^2 + \boldsymbol{w}^\top \boldsymbol{w}.$$

(4) Show that $\sigma_1 = \sigma(A_1)$ is satisfied.

(5) Using (3) and (4), prove $w = 0$.

(6) In equation (*), assume $\alpha = \sigma_1, w = 0$, and $z = 0$. Show $\sigma(A_1) \geq \sigma(A_2)$ provided $\sigma_1 = \sigma(A_1)$.

(7) By using induction, show the existence of a singular value decomposition of any $m \times n$ matrix $A$ where $m \geq n$.

## Problem 2

Let $f(x)$ be a real function which is four times differentiable and satisfies $f'(x) > 0$ in an interval containing $r$, where $f(r) = 0$. Answer the following questions concerning the function

$$F(x) = \frac{f(x)}{\sqrt{f'(x)}}.$$

(1) Express $F'(x)$ in terms of $f(x)$ and its derivatives.

(2) Express $F''(x)$ in terms of $f(x)$ and its derivatives.

(3) The function $G(x)$ is defined as

$$G(x) = x - \frac{F(x)}{F'(x)}.$$

Calculate the following limit,

$$\lim_{x \to r} \frac{G(x) - r}{(x - r)^2}.$$

# Problem 3

Consider a method to "evenly" partition given $n$ positive integers

$$a_1, a_2, \cdots, a_n$$

into $m$ groups. We calculate the sum of the integers in each group. We say a partition is "even" if the largest of these sums is small. Formally, we define an *$m$-partition* as a list of sets $(G_1, \cdots, G_m)$ such that

$$G_1 \cup \cdots \cup G_m = \{1, 2, \cdots, n\},$$
$$i \neq j \Rightarrow G_i \cap G_j = \emptyset,$$

where $\emptyset$ denotes the empty set. For $X \subset \{1, \cdots, n\}$, we define $\sigma(X)$ by

$$\sigma(X) = \sum_{k \in X} a_k.$$

That is, $\sigma(X)$ is the sum of $a_k$'s whose indices are in $X$. Given a partition $\Delta = (G_1, \cdots, G_m)$, we define $P(\Delta)$ by

$$P(\Delta) = \max_{i \in \{1, \cdots, m\}} \sigma(G_i).$$

The goal is, given positive integers $(a_1, \cdots, a_n)$, to find $\Delta$ for which $P(\Delta)$ is as small as possible. Let us analyze the following method, which prepares $m$ sets and adds elements $a_1, \cdots, a_n$ in this order, one at a time, to the set whose sum is minimum at that time. That is,

(a) Let $G_i^0 = \emptyset$ $(i = 1, \cdots, m)$.

(b) Given $G_i^{l-1}$ $(i = 1, \cdots, m)$, $G_i^l$ $(i = 1, \cdots, m)$ are determined as follows. Choose $i \in \{1, \cdots, m\}$ that minimizes $\sigma(G_i^{l-1})$ and let such $i$ be $x$ (if there are multiple such $i$'s, choose an arbitrary one), and

$$\begin{cases} G_x^l = G_x^{l-1} \cup \{l\}, & \\ G_i^l = G_i^{l-1} & (1 \leq i \leq m, i \neq x). \end{cases}$$

(c) Repeat (b) to compute $G_i^1, G_i^2, \cdots$ in sequence and finally obtain $G_i^n$. The partition $(G_1, \cdots, G_m) = (G_1^n, \cdots, G_m^n)$ is the solution.

For example, the above procedure works as the following table for $(a_1, a_2, a_3, a_4, a_5) = (10, 20, 5, 6, 50)$ and $m = 2$.

| $l$ | $G_1^l$ | $\sigma(G_1^l)$ | $G_2^l$ | $\sigma(G_2^l)$ |
|---|---|---|---|---|
| 1 | { 1 } | 10 | $\emptyset$ | 0 |
| 2 | { 1 } | 10 | { 2 } | 20 |
| 3 | { 1, 3 } | 10 + 5 | { 2 } | 20 |
| 4 | { 1, 3, 4 } | 10 + 5 + 6 | { 2 } | 20 |
| 5 | { 1, 3, 4 } | 10 + 5 + 6 | { 2, 5 } | 20 + 50 |

As a result, we obtain $\Delta = (\{1,3,4\},\{2,5\})$ where

$$P(\Delta) = \max\{21, 70\} = 70.$$

An optimal partition (i.e., a partition that minimizes $P(\Delta)$) in this case is, clearly, $\Delta_{\text{opt}} = (\{1,2,3,4\},\{5\})$, where we have

$$P(\Delta_{\text{opt}}) = \max\{41, 50\} = 50.$$

In the following, let $(a_1, \cdots, a_n)$ and $m$ be given, $\Delta = (G_1, \cdots, G_m)$ an $m$-partition obtained by applying the above method to $(a_1, \cdots, a_n)$, and $\Delta_{\text{opt}}$ an optimal $m$-partition for $(a_1, \cdots, a_n)$. Answer the following questions.

(1) Prove that the following inequalities hold for all $i, j \in \{1, \cdots, m\}$.

$$\sigma(G_i) - \sigma(G_j) \le \max\{\, a_1, \cdots, a_n \,\} \le P(\Delta_{\text{opt}})$$

(2) Prove that the following inequality holds.

$$P(\Delta) \le \frac{1}{m}\sum_{k=1}^{n} a_k + \left(1 - \frac{1}{m}\right)\max\{\, a_1, \cdots, a_n \,\}$$

(3) Prove that the following inequality holds (Hint: use the result of either (1) or (2)).

$$P(\Delta) \le \left(2 - \frac{1}{m}\right) P(\Delta_{\text{opt}}).$$

(4) Prove that the above inequality is tight. That is, for any $m \ge 1$, there exists $(a_1, \cdots, a_n)$ for which the equality sign holds.

## Problem 4

Let $A$ be an $n \times n$ real symmetric matrix, $\psi : \mathbf{R}^n \to \mathbf{R}$ a function defined by $\boldsymbol{x}^\top A\boldsymbol{x}$, and $\boldsymbol{b}$ an $n$-dimensional real valued vector. Symbol $^\top$ denotes the matrix transpose. Consider the problem for finding the extremum (minimal value or maximal value) of $\psi(\boldsymbol{x})$ under the conditions that $||\boldsymbol{x}||^2 - 1 = 0$ and $\boldsymbol{b}^\top \boldsymbol{x} = 0$. Let $L(\boldsymbol{x}, \lambda, \mu)$ be the Lagrangean function defined by $L(\boldsymbol{x}, \lambda, \mu) = \psi(\boldsymbol{x}) + \lambda(1 - ||\boldsymbol{x}||^2) - 2\mu\boldsymbol{b}^\top\boldsymbol{x}$, and $\Omega$ the set defined by

$$\Omega = \left\{ (\boldsymbol{x}, \lambda, \mu) \;\middle|\; \begin{array}{lcl} \partial L/\partial x_i & = & 0 \quad (i = 1, 2, \ldots, n), \\ \partial L/\partial \lambda & = & 0, \\ \partial L/\partial \mu & = & 0 \end{array} \right\}.$$

Answer the following questions.

(1) Give explicit formulae for $\partial L/\partial \boldsymbol{x} = (\partial L/\partial x_1, \ldots, \partial L/\partial x_n)$, $\partial L/\partial \lambda$, and $\partial L/\partial \mu$.

(2) Show that if $\boldsymbol{b} = \boldsymbol{0}$, then every solution $(\overline{\boldsymbol{x}}, \overline{\lambda}, \overline{\mu}) \in \Omega$ satisfies that $\overline{\boldsymbol{x}}$ is an eigenvector of the matrix $A$.

(3) Let $A$ be a diagonal matrix satisfying $a_{11} < a_{22} < \cdots < a_{nn}$, and every element of $\boldsymbol{b}$ is non-zero. Suppose that $(\overline{\boldsymbol{x}}, \overline{\lambda}, \overline{\mu}) \in \Omega$. Answer the following questions.

(3-1) Give an explicit formula for $\partial L/\partial x_i$.

(3-2) Show that ($\overline{\mu} = 0$ and $\exists i \in \{1, 2, \ldots, n\}$, $\overline{\lambda} = a_{ii}$) does not hold.

(3-3) Show that ($\overline{\mu} = 0$ and $\forall i \in \{1, 2, \ldots, n\}$, $\overline{\lambda} \neq a_{ii}$) does not hold.

(3-4) Show that ($\overline{\mu} \neq 0$ and $\exists i \in \{1, 2, \ldots, n\}$, $\overline{\lambda} = a_{ii}$) does not hold.

(4) Let $A$ be a $3 \times 3$ diagonal matrix with $(a_{11}, a_{22}, a_{33}) = (2, 4, 6)$, and $\boldsymbol{b}^\top = (\sqrt{3}, \sqrt{2}, \sqrt{3})$. Based on the properties in (3), find all the points in $\Omega$.

(5) Let $A$ be a $4 \times 4$, non-zero diagonal matrix, and $\boldsymbol{b}^\top = (0, 1, 1, 1)$. Show an example of matrix $A$ such that the solution $\overline{\boldsymbol{x}}^\top = (1, 0, 0, 0)$ does not give an extremum. Prove that $\overline{\boldsymbol{x}}$ does not give an extremum in your example.

## Problem 5

Let $\boldsymbol{u}$ be an arbitrary vector, and $\boldsymbol{e}$ an arbitrary unit vector, both in a 3-dimensional Euclidean space $\mathbf{R}^3$. A vector $\boldsymbol{v} \in \mathbf{R}^3$ is obtained by rotating $\boldsymbol{u}$ around $\boldsymbol{e}$ for an angle $\theta$. Assume that the coordinate system of $\mathbf{R}^3$ is orthogonal and right-handed. The sign of angle $\theta$ is defined as follows; when the positive $z$-axis is aligned with the direction of $\boldsymbol{e}$, the rotation direction from the positive $x$-axis toward the positive $y$-axis (the rotation direction of a right-hand screw propelling in the direction of $\boldsymbol{e}$) is positive.

(1) Assume that $\boldsymbol{u}$ is orthogonal to $\boldsymbol{e}$. Derive a representation of $\boldsymbol{v}$ using $\boldsymbol{u}$, $\boldsymbol{e}$ and $\theta$.

(2) Assume that $\boldsymbol{u}$ has an arbitrary direction. Derive a representation of $\boldsymbol{v}$ using $\boldsymbol{u}$, $\boldsymbol{e}$ and $\theta$.

(3) Given $\boldsymbol{e} = (l, m, n)^\top$, $l^2 + m^2 + n^2 = 1$, assume that the above defined rotation of an arbitrary $\boldsymbol{u}$ is obtained by $\boldsymbol{v} = T\boldsymbol{u}$. Represent the transformation matrix $T$ using $l, m, n, \theta$. Note that $^\top$ on the right shoulder of a vector denotes transposition.

## Problem 6

Suppose we distribute $N$ equivalent balls to $S$ distinguishable boxes.

(1) Calculate the total number of distributions.

(2) Calculate the number of distributions under the condition that each box contains at least one ball.

(3) Assume that all the distributions of question (1) occur with the same probability. For a given box, calculate the probability $P(n, N, S)$ that the number of balls in the box is $n$. Let $r$ be a rational number. Show the following limit in terms of $r$ and $n$:

$$\lim_{N,S\to\infty,\, N/S=r} P(n, N, S).$$