数理情報学専攻

修士課程入学試験問題

専門科目 数理情報学

平成２６年８月１９日（火）　１０：００～１３：００

５問出題，３問解答

This booklet is an informal English translation of the original examination booklet. Answer in Japanese or English.

Answer three out of the five problems.

Please note:

(1) Do not open this booklet until the starting signal is given.

(2) Notify the supervisor if there are missing or incorrect pages in your booklet.

(3) Three answer sheets will be given. Use one sheet per problem. If necessary, you may use the back of the sheet.

(4) Fill in the examinee number and the problem number in the designated place of each answer sheet. Do not put your name.

(5) Do not separate a draft sheet from the booklet.

(6) Any answer sheet with marks or symbols unrelated to the answer will be invalid.

(7) Leave the answer sheets and this booklet in the examination room.

Examinee number	No.

Fill in your examinee number.

Problem numbers			

Fill in numbers of the three answered problems.

Problem 1

For an arbitrary vector $\boldsymbol{x} \in \mathbb{R}^n$, its ith component is denoted by x_i $(i = 1, \ldots, n)$ and the norm $\|\boldsymbol{x}\|_\infty$ is defined by $\|\boldsymbol{x}\|_\infty := \max_{i=1,\ldots,n} |x_i|$. We also denote the transpose of $\boldsymbol{x}$ by $\boldsymbol{x}^\top$. Answer the following questions.

(1) For a vector $\boldsymbol{a} \in \mathbb{R}^n$ $(\boldsymbol{a} \neq \boldsymbol{0})$, consider the following linear programming problem:

$$
\text{(P1)} \quad \left| \begin{array}{ll} \underset{\boldsymbol{u},\boldsymbol{v}}{\text{maximize}} & \boldsymbol{a}^\top(\boldsymbol{u} - \boldsymbol{v}) \\ \text{subject to} & \boldsymbol{e}^\top \boldsymbol{u} + \boldsymbol{e}^\top \boldsymbol{v} \leq 1, \\ & \boldsymbol{u} \geq \boldsymbol{0}, \ \boldsymbol{v} \geq \boldsymbol{0} \end{array} \right.
$$

with variables $\boldsymbol{u} \in \mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$, where $\boldsymbol{e}$ and $\boldsymbol{0}$ denote the vectors whose components are all ones and zeros, respectively.

(1-1) Let $(\boldsymbol{u}^*, \boldsymbol{v}^*)$ be an optimal solution of (P1). Show that $u_i^* \cdot v_i^* = 0$ holds for each i, where u_i^* and v_i^* are the ith components of $\boldsymbol{u}^*$ and $\boldsymbol{v}^*$, respectively.

(1-2) Describe the dual problem of (P1).

(1-3) Show that the optimal value of (P1) is equal to $\|\boldsymbol{a}\|_\infty$.

(2) Let $Q = (q_{ij}) \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix. Consider the following optimization problem:

$$
\text{(P2)} \quad \left| \begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \boldsymbol{x}^\top Q^{-1} \boldsymbol{x} \\ \text{subject to} & \|\boldsymbol{x}\|_\infty = 1, \end{array} \right.
$$

where Q^{-1} is the inverse of Q.

(2-1) Obtain an optimal solution of (P2).

(2-2) Obtain the minimum constant C (independent of $\boldsymbol{x}$) such that

$$
\|\boldsymbol{x}\|_\infty \leq C \sqrt{\boldsymbol{x}^\top Q^{-1} \boldsymbol{x}}
$$

holds for every $\boldsymbol{x} \in \mathbb{R}^n$.

Problem 2

Consider a multiple regression equation

$$y = a + bx + cz$$

with response variable y and explanatory variables x, z. In addition, consider a simple regression equation

$$y = \alpha + \beta x$$

with x as the only explanatory variable, and a simple regression equation

$$y = \delta + \theta z$$

with z as the only explanatory variable. Given real-valued data x_i $(i = 1, \ldots, n)$, y_i $(i = 1, \ldots, n)$ and z_i $(i = 1, \ldots, n)$ whose sample variances are normalized to be one, we obtain the coefficients $a, b, c, \alpha, \beta, \delta, \theta$ by the least squares method. Here we assume that $n \geq 4$.

(1) Let r_{xy}, r_{yz}, and r_{xz} be the sample correlation coefficients between x and y, y and z, and x and z, respectively. Describe the conditions satisfied by the triplet (r_{xy}, r_{xz}, r_{yz}).

(2) Show that the regression coefficients b, c, β, θ are expressed as

$$b = \frac{r_{xy} - r_{xz}r_{yz}}{1 - {r_{xz}}^2}, \quad c = \frac{r_{yz} - r_{xz}r_{xy}}{1 - {r_{xz}}^2}, \quad \beta = r_{xy}, \quad \theta = r_{yz},$$

where we assume that each sample correlation coefficient is not equal to ± 1.

(3) Show that $b \cdot \beta < 0$ and $c \cdot \theta < 0$ do not hold at the same time.

(4) Give a numerical instance of values r_{xy}, r_{xz}, r_{yz} that lead to $b \cdot \beta < 0$ and $c \cdot \theta > 0$. Check that the values satisfy the conditions shown in (1).

(5) Explain the case of $b \cdot \beta < 0$ by giving a practical example.

Problem 3

Consider the following ordinary differential equations with respect to real-valued functions $x(t)$ and $y(t)$ of time t:

$$(*)\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t}(t) = x(t) - y(t) - x(t)(x(t)^2 + y(t)^2) + \dfrac{x(t)y(t)}{\sqrt{x(t)^2 + y(t)^2}}\,, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t}(t) = x(t) + y(t) - y(t)(x(t)^2 + y(t)^2) - \dfrac{x(t)^2}{\sqrt{x(t)^2 + y(t)^2}}\,. \end{cases}$$

(1) Change the ordinary differential equations $(*)$ to those in the polar coordinates (r, θ) through the relations

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta.$$

(2) Find the solutions $r(t)$ and $\theta(t)$ of the ordinary differential equations derived in (1), where the initial values are given by $r(0) = r_0$ with $r_0 > 0$ and $\theta(0) = \theta_0$ with $0 \leq \theta_0 < 2\pi$.

(3) Sketch the solutions $r(t)$ and $\theta(t)$ obtained in (2) as functions of t on the plane with the abscissa t and the ordinate $r(t)$ and on that with the abscissa t and the ordinate $\theta(t)$, respectively.

(4) Sketch the phase portrait of the solutions of $(*)$ on the phase space with the abscissa $x(t)$ and the ordinate $y(t)$, paying attention to the behavior with $t \to +\infty$.

Problem 4

Let $f : I \to I$ be a continuous map on the closed interval $I = [0, 1]$ defined by

$$f(x) = \begin{cases} x + \frac{1}{2} & (x \in I_1), \\ 2 - 2x & (x \in I_2), \end{cases}$$

where $I_1 = [0, \frac{1}{2}]$ and $I_2 = [\frac{1}{2}, 1]$. For natural numbers n and k such that $1 \leq k \leq 2^n$, define

$$I_k^n = \left[\frac{k-1}{2^n}, \frac{k}{2^n} \right].$$

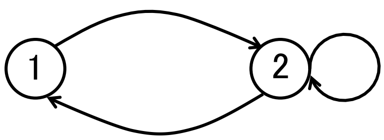
For a natural number p, define the composite function f^p by

$$f^p(x) = f(f^{p-1}(x)),$$

where $f^0(x) = x$, and call $x \in I$ satisfying $f^p(x) = x$ a periodic point of period p.

(1) Show that for a continuous map $g : I \to I$ and a closed interval $J \subseteq I$, if $g(J) \supseteq J$, then J contains a fixed point (a point $x \in J$ such that $g(x) = x$).

(2) Consider a directed path $P = (v_0, v_1, \ldots, v_p)$ of length p on the following directed graph.



Here, $v_i \in \{1, 2\}$ for $i = 0, 1, \ldots, p$. Let m be the number of times that the directed path P visits node 2 except for the end node v_p ($m := \#\{i \mid v_i = 2,\ 0 \leq i < p\}$).

(2-1) Show that there uniquely exists a natural number k satisfying $1 \leq k \leq 2^{m+1}$ such that

$$f^i(I_k^{m+1}) \subseteq I_j \quad (j = v_i)$$

for $i = 0, 1, \ldots, p$, and in particular

$$f^p(I_k^{m+1}) = I_j \quad (j = v_p).$$

We denote such an interval I_k^{m+1} by J_P.

(2-2) Show that f^p is an affine function on J_P for the directed path P.

(2-3) Show that there uniquely exists a periodic point of period p in J_P if the initial node of the directed path P coincides with the end node ($v_0 = v_p$).

(3) Show that the number $N(p)$ of the periodic points of period p is given by $N(p) = \operatorname{tr} A^p$ for the adjacency matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

of the directed graph in (2), where $\operatorname{tr} A^p$ designates the sum of the diagonal components of A^p.

(4) Calculate $N(p)$ and evaluate its growth rate $\displaystyle\lim_{p \to \infty} \frac{\log N(p)}{p}$.

Problem 5

Let S be an array of length n. Each entry $S[i]$ ($i = 1, 2, \ldots, n$) of the array is an element of a set $\mathcal{A} = \{1, 2, \ldots, \sigma\}$. Let k be the number of distinct elements that appear in S. We define sets D_i ($i = 1, 2, \ldots, n$) by

$$D_i = \{j \in \mathbb{N} \mid 1 \leq j < i, S[j] = S[i]\},$$

where $\mathbb{N}$ is the set of natural numbers. We also define $d(i)$ ($i = 1, 2, \ldots, n$) by

$$d(i) = \begin{cases} i - \max D_i & (D_i \neq \emptyset), \\ i & (D_i = \emptyset), \end{cases}$$

where $\max D_i$ indicates the maximum value in D_i. Consider algorithms for computing all $d(1), d(2), \ldots, d(n)$.

We use the RAM model for analyzing the algorithms. We assume that each entry of the array S is read in constant time but cannot be modified. The space complexity of each algorithm does not include the space for storing S. Once the value of each $d(i)$ is output, it is not necessary to be stored in the memory.

(1) Give an algorithm for computing all $d(1), d(2), \ldots, d(n)$ in $O(n + \sigma)$ time using $O(\sigma)$ working space.

(2) Describe an outline of an algorithm for computing all $d(1), d(2), \ldots, d(n)$ in $O(n \log k)$ time using $O(k)$ working space.

(3) Consider the following algorithm which computes all $d(1), d(2), \ldots, d(n)$ using $O(1)$ working space.

> For each i, we examine $S[i-1], S[i-2], \ldots$ in order. The first time we find j such that $S[j] = S[i]$, let $d(i) = i - j$. If we do not find such j after examining the entries until $S[1]$, let $d(i) = i$.

Show that the time complexity of this algorithm is $O(nk)$.

(4) Show that $\displaystyle\sum_{i=1}^{n} \log d(i) = O(n \log k)$.