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２２年８月２４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 the draft paper from this 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

Let A be a real symmetric matrix of order n, and consider the eigenvalue problem

$$A\boldsymbol{x} = \lambda \boldsymbol{x}.$$

We assume that all the eigenvalues are distinct, and are arranged in the ascending order as

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n. \qquad \cdots \qquad (*1)$$

The corresponding eigenvectors are denoted by $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$.

(1) Show that the set of $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ can be chosen so that it forms an orthonormal basis.

In what follows, suppose that each component of the matrix A varies depending on a parameter θ, and we consider the rate of change of the k-th eigenvalue λ_k, as well as that of the corresponding eigenvector $\boldsymbol{x}_k$, with respect to θ. We assume that, for any θ under consideration, A is a real symmetric matrix satisfying the assumption $(*1)$. It is known that if each component $A_{ij}(\theta)$ of $A = (A_{ij}(\theta) \mid i, j = 1, \ldots, n)$ is differentiable with respect to θ, then each eigenvalue is differentiable with respect to θ, and the set of eigenvectors forming an orthonormal basis can be chosen so that each component of the eigenvector is differentiable with respect to θ. Therefore, we denote the derivatives of λ_k and $\boldsymbol{x}_k$ as

$$\mu_k = \frac{\mathrm{d}\lambda_k}{\mathrm{d}\theta}, \qquad \boldsymbol{y}_k = \frac{\mathrm{d}\boldsymbol{x}_k}{\mathrm{d}\theta},$$

where $\dfrac{\mathrm{d}\boldsymbol{x}_k}{\mathrm{d}\theta}$ denotes the vector consisting of the derivatives of the components of $\boldsymbol{x}_k$ with respect to θ.

(2) Show that the system of linear equations

$$\begin{bmatrix} A - \lambda_k I & -\boldsymbol{x}_k \\ -\boldsymbol{x}_k^\top & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_k \\ \mu_k \end{bmatrix} = \begin{bmatrix} -\dfrac{\mathrm{d}A}{\mathrm{d}\theta}\boldsymbol{x}_k \\ 0 \end{bmatrix} \qquad \cdots \qquad (*2)$$

is satisfied. Here, $\dfrac{\mathrm{d}A}{\mathrm{d}\theta}$ is the matrix of order n, the (i, j)-component of which is given by $\dfrac{\mathrm{d}A_{ij}}{\mathrm{d}\theta}$, and $\boldsymbol{x}_k^\top$ denotes the transpose of the vector $\boldsymbol{x}_k$.

(3) Show that the system of linear equations $(*2)$ has a unique solution.

(4) Show that $\mu_k = \boldsymbol{x}_k^\top \dfrac{\mathrm{d}A}{\mathrm{d}\theta}\boldsymbol{x}_k$.

Problem 2

Let $z_1, z_2, \ldots, z_T$ be T random variables taking values 1 or 2. Consider a Markov chain specified by the initial probabilities $\pi_1 = \mathrm{P}(z_1 = 1)$, $\pi_2 = \mathrm{P}(z_1 = 2)$ and the transition probabilities $p_{ij} = \mathrm{P}(z_{t+1} = j \mid z_t = i)$ $(i, j = 1, 2;\ t = 1, \ldots, T-1)$, such that the joint probability of these random variables (i.e. the probability of the path $(i_1, i_2, \ldots, i_T)$) is written as

$$\mathrm{P}(z_1 = i_1, z_2 = i_2, \ldots, z_T = i_T) = \pi_{i_1} p_{i_1 i_2} p_{i_2 i_3} \ldots p_{i_{T-1} i_T}.$$

Here the parameters satisfy $\pi_1 + \pi_2 = 1$ and $p_{i1} + p_{i2} = 1$ $(i = 1, 2)$. We call π, p, and q the parameters of the probability distribution, where $\pi = \pi_1 = 1 - \pi_2$ is the initial probability, and $p = p_{12} = 1 - p_{11}$ and $q = p_{21} = 1 - p_{22}$ are transition probabilities.

We observe n independent paths from this Markov chain. Let $(i_1^{(k)}, i_2^{(k)}, \ldots, i_T^{(k)})$ denote the k-th observed path $(k = 1, \ldots, n)$. In these n observed paths, let

$$x_i = (\text{the number of paths } k \text{ such that } i_1^{(k)} = i), \quad i = 1, 2,$$

denote the initial frequency of $i = 1, 2$ and let

$$y_{ij} = \Big(\text{the number of pairs } (k, t) \text{ such that } (i_t^{(k)}, i_{t+1}^{(k)}) = (i, j)\ (1 \le k \le n;\ 1 \le t \le T-1)\Big), \quad i, j = 1, 2,$$

denote the total number of transitions from i to j.

(1) Express the joint probability of n paths in terms of $x_1, x_2, y_{11}, y_{12}, y_{21}, y_{22}$ and the parameters π, p, q.

(2) Obtain the values of π, p, q (maximum likelihood estimates) which maximize the probability $\prod_{k=1}^{n} \mathrm{P}(z_1 = i_1^{(k)}, z_2 = i_2^{(k)}, \ldots, z_T = i_T^{(k)}; \pi, p, q)$ of the observed n paths.

(3) The stationary distribution (π_1, π_2) of the Markov chain is defined by

$$(\pi_1, \pi_2) \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} = (\pi_1, \pi_2).$$

Express the stationary distribution in terms of p and q, where p and q are positive.

(4) Obtain the expected values of $x_1, x_2, y_{11}, y_{12}, y_{21}, y_{22}$, when the initial distribution of the Markov chain is the stationary distribution obtained in (3).

(5) Assume that the initial distribution (π_1, π_2) is the stationary distribution obtained in (3) and π_1 and π_2 are expressed as functions of p, q as $\pi_1 = \pi_1(p, q)$ and $\pi_2 = \pi_2(p, q)$. Under this assumption, let $L(p, q; x_1, x_2, y_{11}, y_{12}, y_{21}, y_{22})$ denote the probability of n paths in terms of $x_1, x_2, y_{11}, y_{12}, y_{21}, y_{22}$ and p, q. Evaluate the limit as $T \to \infty$ (with n fixed) of

$$\begin{pmatrix} \lim_{T\to\infty} \frac{1}{T}\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial p^2}\right) & \lim_{T\to\infty} \frac{1}{T}\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial p \partial q}\right) \\ \lim_{T\to\infty} \frac{1}{T}\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial p \partial q}\right) & \lim_{T\to\infty} \frac{1}{T}\mathrm{E}\left(-\frac{\partial^2 \log L}{\partial q^2}\right) \end{pmatrix}$$

(i.e. the limit of the Fisher information matrix of (p, q)), where E( ) denotes the expected value.

Problem 3

Consider the following initial value problem of partial differential equation:

$$\frac{\partial \phi}{\partial t} = -a\frac{\partial \phi}{\partial x} \quad (-\infty < x < +\infty,\ t > 0), \qquad \cdots \qquad (*1)$$
$$\phi(x,0) = f(x) \quad (-\infty < x < +\infty),$$

where x and t denote space and time respectively, $\phi = \phi(x,t)$ is an unknown real-valued function, a is a positive constant, and $f(x)$ is a periodic function with period 1. In this problem, a numerical method to obtain a solution with period 1 (i.e. function ϕ which satisfies $\phi(x+1,t) = \phi(x,t)$ for all x and $t \geq 0$) is considered.

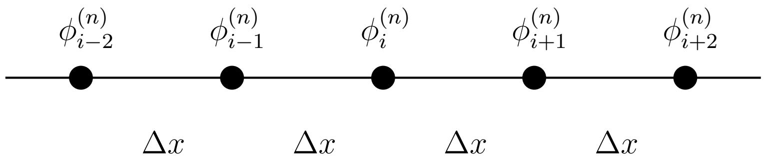

Fig. 1

The interval $[0,1)$ is discretized by equally-spaced N grid points, where the position of the grid point is given by $x_i = i\Delta x \ \ (i = 0, 1, \ldots, N-1;\ \Delta x = 1/N)$, as shown in Fig. 1. The following finite-difference scheme is considered:

$$\frac{\phi_i^{(n+1)} - \phi_i^{(n)}}{\Delta t} = -a\left\{\mu\ \frac{\phi_{i+1}^{(n)} - \phi_i^{(n)}}{\Delta x} + (1-\mu)\ \frac{\phi_i^{(n)} - \phi_{i-1}^{(n)}}{\Delta x}\right\}$$
$$(i = 0, 1, \ldots, N-1), \qquad \cdots \qquad (*2)$$

where Δt is a time step, and $\phi_N^{(n)} = \phi_0^{(n)},\ \phi_{-1}^{(n)} = \phi_{N-1}^{(n)}$. Here, $\phi_i^{(n)}$ denotes an approximate value of ϕ on grid point $x_i = i\Delta x$ at time $t_n = n\Delta t$, and μ is a real-valued parameter. The approximate value $\phi_i^{(n+1)}$ of ϕ at time $t_{n+1} = (n+1)\Delta t$ can be computed from $\phi_{i-1}^{(n)},\ \phi_i^{(n)},\ \phi_{i+1}^{(n)}$ at time $t_n = n\Delta t$ using the finite-difference scheme $(*2)$.

(1) To evaluate accuracy of the finite-difference scheme $(*2)$, assume that

$$\phi_{i-1}^{(n)} = \phi(x_{i-1}, t_n), \qquad \phi_i^{(n)} = \phi(x_i, t_n), \qquad \phi_{i+1}^{(n)} = \phi(x_{i+1}, t_n)$$

for a function $\phi(x,t)$ that satisfies equation $(*1)$. Obtain the order of $\dfrac{\phi_i^{(n+1)} - \phi(x_i, t_{n+1})}{\Delta t}$ in terms of Δx and Δt for sufficiently small Δx and Δt.

(2) To evaluate stability of the finite-difference scheme $(*2)$, assume that $\phi_i^{(n)}$ is given by

$$\phi_i^{(n)} = V^{(n)} \exp(i\theta\sqrt{-1}),$$

where $V^{(n)}$ is a complex number and θ is a real number. Substitution of $\phi_i^{(n)} = V^{(n)} \exp(i\theta\sqrt{-1})$ into $(*2)$ yields a relation of the form

$$V^{(n+1)} = G(\theta)V^{(n)},$$

where $G(\theta)$ is a complex number depending on θ, called the amplification factor. The finite-difference scheme $(*2)$ is considered stable when $|G(\theta)| \leq 1$ for all real θ, according to the von Neumann stability analysis. Determine the stability of the finite-difference scheme $(*2)$ in the following two cases:

(a) $\mu = 0$,

(b) $\mu = 1/2$.

Problem 4

Many engineering systems have both continuous and discrete aspects. Let us consider the following ordinary differential equation† having both the aspects with respect to time t:

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}(t) + \omega^2 x(t) = a\, x(\lfloor t \rfloor) \quad \text{for} \quad t \geq 0, \qquad \cdots \quad (*)$$

where $\lfloor t \rfloor$ denotes the largest integer which is not larger than t, and a and ω are positive constants. Let the initial conditions be

$$x(0) = x_0, \qquad \frac{\mathrm{d}x}{\mathrm{d}t}(0) = y_0.$$

† More specifically, we consider a real-valued function $x(t)$ such that $x(t)$ satisfies Equation $(*)$ for t which is not an integer, and both $x(t)$ and $\frac{\mathrm{d}x}{\mathrm{d}t}(t)$ are continuous for all $t \geq 0$.

(1) For an integer n (≥ 0), let x_n and y_n be the values of $x(t)$ and $\frac{\mathrm{d}x}{\mathrm{d}t}(t)$ at $t = n$, respectively. Express the solution $x(t)$ for t in the interval $n \leq t \leq n+1$ using x_n and y_n.

(2) Find a square matrix A of order 2 which is independent of n and satisfies

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = A \begin{bmatrix} x_n \\ y_n \end{bmatrix} \qquad (n = 0, 1, 2, \ldots).$$

(3) Let $\omega = \pi/3$. Find the condition which a must satisfy so that, for any initial values $(x_0,\, y_0)$, $x(t)$ remains bounded for $t \geq 0$.

Problem 5

Let $G = (V, E)$ be an undirected graph with the vertex set $V = \{1, 2, \ldots, n\}$ and the edge set $E = \{e_i = \{k_i, l_i\} \mid i = 1, 2, \ldots, m\}$. Given a nonnegative vertex cost c_j for each $j \in V$, consider the problem of computing a vertex cover W with the minimum $\sum_{j \in W} c_j$ (called the minimum cost vertex cover problem). Here a set W of vertices is called a vertex cover if $W \cap \{k, l\} \neq \emptyset$ holds for all edges $\{k, l\} \in E$.

Let us formulate the minimum cost vertex cover problem as an integer programming problem. Let A denote the transpose of the incidence matrix of G, i.e., A denotes an $m \times n$ matrix such that the rows and columns of A respectively correspond to edges and vertices in G, and the element (i, j) of A is 1, if edge e_i is incident to vertex j, and 0 otherwise. Let $\mathbf{1}$ denote the m-dimensional column vector with all 1's, and c denote the n-dimensional column vector whose j-th element is c_j. Then we have the following integer programming problem that represents the minimum cost vertex cover problem:

$$\begin{array}{lrl} \text{(Problem IP)} & \min & c^\top x \\ & \text{subject to} & Ax \geq \mathbf{1} \\ & & x = (x_1, \ldots, x_n)^\top \in \{0, 1\}^n, \end{array}$$

where $^\top$ denotes the transposition. Furthermore, we consider the following linear programming problem as a relaxation of Problem IP:

$$\begin{array}{lrl} \text{(Problem P)} & \min & c^\top x \\ & \text{subject to} & Ax \geq \mathbf{1} \\ & & x = (x_1, \ldots, x_n)^\top \geq 0. \end{array}$$

(1) Construct an instance of Problem IP whose optimal value (i.e. minimum of the objective function) is different from the optimal value of Problem P.

(2) We consider the following algorithm that computes an approximate solution x^* of IP from an optimal solution x of P. Given an optimal solution x of P, set x_j^* to 1 if $x_j \geq 1/2$, and 0 otherwise. Show that x^* is a feasible solution of Problem IP whose objective value $c^\top x^*$ is at most twice of the optimal value of IP.

(3) The dual of Problem P is

$$\begin{array}{lrl} \text{(Problem D)} & \max & \mathbf{1}^\top y \\ & \text{subject to} & A^\top y \leq c \\ & & y = (y_1, \ldots, y_m)^\top \geq 0. \end{array}$$

Describe and prove the weak duality theorem for Problems P and D.

(4) We next consider the following algorithm that computes an approximate solution x of Problem IP without solving Problem P.

Algorithm Primal-Dual

for $j = 1, \ldots, n$ **do** $x_j := 0$;

for $i = 1, \ldots, m$ **do** $y_i := 0$;

while x is infeasible to Problem IP **do**

Choose arbitrarily an edge $e_h = \{k_h, l_h\}$ $(1 \le h \le m)$ with $x_{k_h} + x_{l_h} = 0$;

$y_h := \min\{c_{k_h} - \sum_{i:i\neq h,\{k_i,l_i\}\ni k_h} y_i, \quad c_{l_h} - \sum_{i:i\neq h,\{k_i,l_i\}\ni l_h} y_i\}$;

if $y_h = c_{k_h} - \sum_{i:i\neq h,\{k_i,l_i\}\ni k_h} y_i$ **then** $x_{k_h} := 1$

else $x_{l_h} := 1$;

end{while}

Show that x obtained by Algorithm Primal-Dual is a feasible solution of Problem IP whose objective value $c^\top x$ is at most twice of the optimal value of IP.

(Hint) Show first that y obtained by Algorithm Primal-Dual is a feasible solution of Problem D, and $c_j = \sum_{i:\{k_i,l_i\}\ni j} y_i$ holds for any j with $x_j = 1$.

(Note) In the description of the algorithm above, $\sum_{i:i\neq h,\{k_i,l_i\}\ni k_h} y_i$ means the sum of all y_i's that satisfy $i \neq h$ and $\{k_i, l_i\} \ni k_h$. Similarly, $\sum_{i:i\neq h,\{k_i,l_i\}\ni l_h} y_i$ means the sum of all y_i's that satisfy $i \neq h$ and $\{k_i, l_i\} \ni l_h$.