2023年度10月期入学 / 2024年度4月期入学
京都大学 大学院情報学研究科
修士課程 知能情報学コース 入学者選抜試験問題
(情報学基礎)

2023年8月7日 9:00～11:00

【注意】

1. 問題冊子はこの表紙を含めて13枚ある。
2. 試験開始の合図があるまで中を見てはいけない。
3. 試験開始後、枚数を確認し、落丁または印刷の不鮮明なものがあれば直ちに申し出ること。
4. 問題は日本語と英語の両方で出題されている。すべて解答しなさい。
   - F1-1, F1-2　線形代数、微分積分 ……………………… 1-4ページ
   - F2-1, F2-2　アルゴリズムとデータ構造 ……………… 5-12ページ
5. 特に指定のない限り、日本語または英語で解答すること。
6. 解答用紙に記載されている注意事項についても留意すること。

October 2023 Admissions / April 2024 Admissions
Entrance Examination for Master's Program
Intelligence Science and Technology Course
Graduate School of Informatics, Kyoto University
(Fundamentals of Informatics)

August 7, 2023
9:00 - 11:00

NOTES

1. This is the Question Booklet in 13 pages including this front cover.
2. Do not open the booklet until you are instructed to start.
3. After the examination has started, check the number of pages and notify proctors (professors) immediately if you find missing pages or unclear printings.
4. Questions are written in Japanese and English. **Answer all the questions.**
   - F1-1, F1-2　Linear Algebra, Calculus ………………… Pages 1 to 4
   - F2-1, F2-2　Algorithms and Data Structures ………… Pages 5 to 12
5. Write your answer in Japanese or English, unless otherwise specified.
6. Read carefully the notes on the Answer Sheets as well.

F1–1、F1–2、 F2–1、 F2–2 それぞれ別の解答用紙を用いて解答すること。

以下の設問において i は虚数単位、$\mathbb{C}$ は複素数全体の集合を表す。また、行列 $\mathbf{A}$ に対して、$\mathbf{A}^{\mathsf{H}}$ は $\mathbf{A}$ の共役転置（エルミート転置）を、$\mathbf{A}^{-1}$ は $\mathbf{A}$ の逆行列をそれぞれ表す。

設問 1　行列 $\mathbf{D} \in \mathbb{C}^{4\times 4}$ を次で定義する。

$$\mathbf{D} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

このとき、以下の問いに答えよ。

(1) $\mathbf{D}$ がユニタリ行列であることを示せ。

(2) 行列 $\mathbf{G} \in \mathbb{C}^{4\times 4}$ を

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & i \end{pmatrix}$$

で定義するとき、$\mathbf{D}^{\mathsf{H}}\mathbf{G}\mathbf{D}$ の逆行列を導出せよ。

設問 2　n 次元ベクトル空間 V が 2 つの部分空間 W_1, W_2 の直和であるとする。ベクトル $\mathbf{x} \in V$ が

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \quad \mathbf{x}_1 \in W_1,\ \mathbf{x}_2 \in W_2$$

と分解されるとき、$\mathbf{x}$ を $\mathbf{x}_1$ に写す一次変換を考える。この一次変換を表す V のある基底に関する行列を $\mathbf{B}$ とする。このとき、以下の問いに答えよ。

(1) $\mathbf{x}$ を $\mathbf{x}_2$ に写す一次変換を表す上述の基底に関する行列を導出せよ。

(2) $\mathbf{B}^2 = \mathbf{B}$ が成り立つことを示せ。

(3) 適当な n 次正則行列 $\mathbf{P}$ をとれば

$$\mathbf{P}^{-1}\mathbf{B}\mathbf{P} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & \text{\Large 0} & \\ & & 1 & & & \\ & & & 0 & & \\ & \text{\Large 0} & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

となることを示せ。

Use one answer sheet for each of F1–1, F1–2, F2–1, and F2–2.

In the questions below, i denotes the imaginary unit and $\mathbb{C}$ is the set of all complex numbers. Also, $\mathbf{A}^{\mathsf{H}}$ stands for the conjugate transpose (Hermitian transpose) of a matrix $\mathbf{A}$, and $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$.

Q.1 Let $\mathbf{D} \in \mathbb{C}^{4\times 4}$ be a matrix defined as

$$\mathbf{D} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}.$$

Answer the following questions.

(1) Show that $\mathbf{D}$ is a unitary matrix.

(2) Let $\mathbf{G} \in \mathbb{C}^{4\times 4}$ be a matrix defined as

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & i \end{pmatrix}.$$

Derive the inverse of $\mathbf{D}^{\mathsf{H}}\mathbf{G}\mathbf{D}$.

Q.2 Suppose that an n-dimensional vector space V is the direct sum of two subspaces W_1 and W_2. When a vector $\mathbf{x} \in V$ is decomposed as

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2, \quad \mathbf{x}_1 \in W_1,\ \mathbf{x}_2 \in W_2,$$

we consider the linear transformation that maps $\mathbf{x}$ to $\mathbf{x}_1$. Let $\mathbf{B}$ denote the matrix of the linear transformation with respect to some basis of V. Answer the following questions.

(1) Derive the matrix of the linear transformation that maps $\mathbf{x}$ to $\mathbf{x}_2$ with respect to the basis mentioned above.

(2) Show that $\mathbf{B}^2 = \mathbf{B}$ holds.

(3) Show that

$$\mathbf{P}^{-1}\mathbf{B}\mathbf{P} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & 1 & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

holds for some non-singular matrix $\mathbf{P}$ of size $n \times n$.

F1–1、F1–2、 F2–1、 F2–2 それぞれ別の解答用紙を用いて解答すること。

設問 1　以下の積分を求めよ。計算過程も明示すること。

(1) $\displaystyle\int_0^\infty \frac{1}{(x^2+1)^2}\mathrm{d}x$

(2) $D = \left\{(x, y) \mid x^2 + \dfrac{y^2}{4} \leq 1\right\}$ としたときに、

$$\iint_D x^2y^2\mathrm{d}x\mathrm{d}y$$

設問 2　以下の問いに答えよ。計算過程も明示すること。

(1) $\log_e(1.02)$ の小数第 7 位を四捨五入し、小数第 6 位まで求めよ。

(2) $x > 0$ に対して、次の不等式が成り立つことを示せ。

$$x - \frac{x^2}{2} < \log_e(1+x) < 1 - \frac{x^2}{2} + \frac{x^3}{3}$$

設問 3　$3x^2 + 2y^2 + z^2 = 1$ の条件の下で、xyz の最大値と最小値を求めよ。

Use one answer sheet for each of F1–1, F1–2, F2–1, and F2–2.

Q.1 Compute the following integrals. Derivations must be clearly shown.

(1) $\displaystyle\int_0^{\infty} \frac{1}{(x^2+1)^2} \mathrm{d}x$

(2) $\displaystyle\iint_D x^2 y^2 \mathrm{d}x\mathrm{d}y, \quad$ where $D = \left\{(x, y) \mid x^2 + \frac{y^2}{4} \leq 1\right\}$

Q.2 Answer the following questions. Derivations must be clearly shown.

(1) Compute $\log_e(1.02)$ up to the sixth decimal place by rounding the seventh decimal place.

(2) For $x > 0$, show that the following inequality holds.

$$x - \frac{x^2}{2} < \log_e(1+x) < 1 - \frac{x^2}{2} + \frac{x^3}{3}$$

Q.3 Derive the maximum and minimum values of xyz, when $3x^2 + 2y^2 + z^2 = 1$.

F1–1、F1–2、F2–1、F2–2 それぞれ別の解答用紙を用いて解答すること。

設問 max ヒープは根以外のノードの値が必ず親ノードの値の大きさ以下になっている、おおよそ完全二分木である。max ヒープはインデックスが 1 から始まる配列として表現でき、インデックス i のノードの親ノード、左の子ノード、右の子ノードのインデックスはそれぞれ、Parent(i)=$\lfloor \frac{i}{2} \rfloor$、Left(i)=2i、Right(i)=2i+1 として求められる。ただし、$\lfloor \cdot \rfloor$ は床関数である。

(1) インデックス i のノードを根とする部分木が max ヒープとなるように変換する再帰的関数 MaxHeapify を表す次の疑似コードをそれぞれの四角（(1-a)、(1-b)、(1-c)、(1-d)）を埋めることにより完成させよ。この疑似コードでは、インデックス i の左の子ノードと右の子ノードを根とする部分木はともに max ヒープであると仮定する。

Algorithm 1 MaxHeapify(A, i)

```
lft ← Left(i)
rgt ← Right(i)
largest ← i
if lft ≤ A.heapsize and [(1-a)] then
   largest ← lft
end if
if rgt ≤ A.heapsize and [(1-b)] then
   largest ← rgt
end if
if [(1-c)] then
   tmp ← A[i]; A[i] ← A[largest]; A[largest] ← tmp
   [(1-d)]
end if
```

ただし、A.heapsize は配列 A に格納されているヒープを構成する要素の数を表す。(3) で用いる A.length は配列 A そのものの大きさを表し、A.heapsize ≤ A.length である。

(2) (1) の MaxHeapify に長さ n の整数列を与えたときの最悪実行時間の漸近的上界をビッグオー記法を用いて答えよ。ただし、オーダーとして最も低いものを答えること。

（次のページに続く）

(3) (1) の `MaxHeapify` を用いて任意の整数列を昇順にソートするアルゴリズム `HeapSort` を表す次の疑似コードをそれぞれの四角（`(3-a)`、`(3-b)`、`(3-c)`）を埋めることにより完成させよ。ただし、`(3-a)` には考えうる最も小さい値を入れること。

Algorithm 2 `HeapSort(A)`

```
A.heapsize ← A.length
for i ← [ (3-a) ],...,1 do
    [ (3-b) ]
end for
for i ← A.length,...,2 do
    tmp ← A[1]; A[1] ← A[i]; A[i] ← tmp
    A.heapsize ← A.heapsize - 1
    [ (3-c) ]
end for
```

(4) (3) のアルゴリズムに長さ n の整列されていない整数列を与えたときの最悪実行時間の漸近的上界をビッグオー記法を用いて答えよ。ただし、オーダーとして最も低いものを答えること。

(5) (3) のアルゴリズムに長さ n の降順に整列された整数列を与えたときの実行時間の漸近的上界をビッグオー記法を用いて答えよ。ただし、オーダーとして最も低いものを答えること。

Use one answer sheet for each of F1–1, F1–2, F2–1, and F2–2.

Q. A max-heap is a nearly complete binary tree in which, except for the root node, the value of each node is at most the value of its parent. A max-heap can be represented with an array with indices starting with 1, and the indices of the parent node, left child node, and right child node of a node with index `i` can be found with `Parent(i)`$=\lfloor \frac{i}{2} \rfloor$, `Left(i)=2i`, `Right(i)=2i+1`, respectively, where $\lfloor \cdot \rfloor$ is the floor function.

(1) Complete the following pseudocode of a recursive function `MaxHeapify` for turning the subtree rooted at a node with index `i` into a max-heap by filling the blanks (`(1-a)`, `(1-b)`, `(1-c)`, and `(1-d)`). The code assumes that the subtrees rooted at the left child node and right child node of the node with index `i` are max-heaps.

Algorithm 1 `MaxHeapify(A, i)`

```
lft ← Left(i)
rgt ← Right(i)
largest ← i
if lft ≤ A.heapsize and [ (1-a) ] then
    largest ← lft
end if
if rgt ≤ A.heapsize and [ (1-b) ] then
    largest ← rgt
end if
if [ (1-c) ] then
    tmp ← A[i]; A[i] ← A[largest]; A[largest] ← tmp
    [ (1-d) ]
end if
```

In the pseudocode, `A.heapsize` denotes the number of elements of the max-heap in array `A`. Note that `A.length` used in (3) is the size of the array `A` itself and `A.heapsize` ≤ `A.length`.

(2) Answer the asymptotic upper bound of the worst-case running time of `MaxHeapify` in (1) when given an integer sequence of length `n` and answer it with the big-O notation. Answer with the smallest order.

(continued on the next page)

(3) Complete the following pseudocode of an algorithm `HeapSort` that uses `MaxHeapify` in (1) to sort an arbitrary integer sequence in ascending order by filling the blanks (`(3-a)`, `(3-b)`, and `(3-c)`). Note that the smallest possible value should be given for `(3-a)`.

Algorithm 2 `HeapSort(A)`

```
A.heapsize ← A.length
for i ← [ (3-a) ],...,1 do
    [ (3-b) ]
end for
for i ← A.length,...,2 do
    tmp ← A[1]; A[1] ← A[i]; A[i] ← tmp
    A.heapsize ← A.heapsize - 1
    [ (3-c) ]
end for
```

(4) Answer the asymptotic upper bound of the worst-case running time of the algorithm in (3) when given an unsorted integer sequence of length n and answer it with the big-O notation. Answer with the smallest order.

(5) Answer the asymptotic upper bound of the running time of the algorithm in (3) when given an integer sequence of length n sorted in descending order and answer it with the big-O notation. Answer with the smallest order.

令和5年9月4日修正

F1-1, F1-2, F2-1, F2-21 それぞれ別の解答用紙を用いて解答すること。

設問1　２つの文字列の編集コストは、文字の置換・削除・挿入により、一方を他方に変換するのに必要最小限の操作（置換・削除・挿入）の数で定義できる。なお、２つの文字列の長さ(N)は同じとする。以下の問いに答えよ。

(1) 文字列“PARIS”を“PAIRS”に編集するコストを答えよ。

F1–1、F1–2、 F2–1、 F2–2 それぞれ別の解答用紙を用いて解答すること。

設問 2　すべての頂点の次数が 1 または 3 であり、辺に向きがない、根なし無順序木について考える。次数が 1 の頂点を葉と呼ぶ。すべての辺には正の値の長さが与えられ、葉 X, Y 間の最短路の長さを $d(X, Y)$ とする。

(1)　3 つの葉 A, B, C を持ち、それ以外には葉を持たない木を考える。$d(A, B) = 12$, $d(A, C) = 10, d(B, C) = 6$ を満たす木を重複なく全て導出して、各辺の長さとともに図示せよ。

(2)　4 つの葉 A, B, C, D を持ち、それ以外には葉を持たない木を考える。$d(A, B) = 12$, $d(A, C) = 10, d(A, D) = 7, d(B, C) = 6, d(B, D) = 9, d(C, D) = 7$ を満たす木を重複なく全て導出して、各辺の長さとともに図示せよ。

Correction: September 4, 2023

Use one answer sheet for each of F1-1, F1-2, F2-1, and F2-2.

Q.1 The edit cost of two strings can be defined by the minimum number of operations (substitutions, deletions, and insertions) necessary to make one into the other by substituting, deleting, or inserting characters. Here we assume the length (N) of the two strings is same. Answer the following questions.

(1) Answer the cost of editing the string of "PARIS" into "PAIRS".

Use one answer sheet for each of F1–1, F1–2, F2–1, and F2–2.

Q.2 Let us consider unrooted unordered trees where the degree of each node is 1 or 3, and each edge has no direction. A node with degree 1 is called a leaf. Each edge is given a length of a positive value. Let $d(X, Y)$ be the length of the shortest path between leaves X and Y.

(1) Let us consider all the trees with exactly three leaves A, B, and C satisfying $d(A, B) = 12$, $d(A, C) = 10$, and $d(B, C) = 6$. Without duplication, derive and draw such trees with the length of each edge.

(2) Let us consider all the trees with exactly four leaves A, B, C, and D satisfying $d(A, B) = 12$, $d(A, C) = 10$, $d(A, D) = 7$, $d(B, C) = 6$, $d(B, D) = 9$, and $d(C, D) = 7$. Without duplication, derive and draw such trees with the length of each edge.

2023 年度 10 月期入学 / 2024 年度 4 月期入学
京都大学 大学院情報学研究科
修士課程 知能情報学コース 入学者選抜試験問題
（専門科目）

2023 年 8 月 7 日 12:00～14:00

【注意】
1. 問題冊子はこの表紙を含めて 13 枚ある。
2. 試験開始の合図があるまで中を見てはいけない。
3. 試験開始後、枚数を確認し、落丁または印刷の不鮮明なものがあれば直ちに申し出ること。
4. 問題は下記 6 題であり、日本語と英語の両方で出題されている。このうちいずれか **2 題**を選択し、解答しなさい。

5. 特に指定のない限り、日本語または英語で解答すること。
6. 解答用紙に記載されている注意事項についても留意すること。

October 2023 Admissions / April 2024 Admissions
Entrance Examination for Master's Program
Intelligence Science and Technology Course
Graduate School of Informatics, Kyoto University
(Specialized Subjects)

August 7, 2023
12:00 - 14:00

NOTES
1. This is the Question Booklet in 13 pages including this front cover.
2. Do not open the booklet until you are instructed to start.
3. After the examination has started, check the number of pages and notify proctors (professors) immediately if you find missing pages or unclear printings.
4. There are 6 questions, written in Japanese and English. The questions are classified as listed below. **Choose and answer 2 questions.**

5. Write your answer in Japanese or English, unless otherwise specified.
6. Read carefully the notes on the Answer Sheets as well.

設問 1 以下の問いに答えよ。

(1) 人間の認知機能の脳神経メカニズムを調べるために用いられる非侵襲の脳活動計測法を2つ挙げ、それぞれの計測法の長所と短所を説明しなさい。

(2) 人間の脳活動計測以外の研究方法で、人間の認知機能と脳神経メカニズムの関係を調べるために用いられている研究方法を2つ挙げ、簡潔に説明しなさい。

設問 2 次の6つの心理学、神経科学の用語のペアの中から4つを選んで、各ペアの共通点と相違点に着目しながら意味を説明しなさい。

(1) Law of good continuation & Law of common fate

(2) Pop-out & Illusory conjunction

(3) Visual agnosia & Optic ataxia

(4) Change blindness & Attentional blink

(5) Episodic memory & Procedural memory

(6) Theory of mind & Empathy

Q.1 Answer the following questions.

(1) List and describe the advantages and disadvantages of two noninvasive brain activity recording methods used to investigate the neural mechanisms of human cognitive functions.

(2) List and briefly explain two research methods, other than the recordings of human brain activity, used to investigate the relationship between human cognitive functions and the brain's neural mechanisms.

Q.2 Select four pairs of psychology/neuroscience terms from the following six pairs, and explain their meanings, focusing on commonalities and differences within each pair.

(1) Law of good continuation & Law of common fate

(2) Pop-out & Illusory conjunction

(3) Visual agnosia & Optic ataxia

(4) Change blindness & Attentional blink

(5) Episodic memory & Procedural memory

(6) Theory of mind & Empathy

設問 1　$X_1, X_2, \cdots, X_n$ を平均 μ の母集団からの大きさ n の無作為標本とする。X_i の加重和 $\sum_{i=1}^{n} w_i X_i$ が μ の不偏推定量であるための必要十分条件を示せ。

設問 2　$X_1, X_2, \cdots, X_n$ を平均 μ, 分散 σ^2 の正規母集団からの大きさ n の無作為標本とする。
(1) μ と σ^2 の対数尤度関数 $L(\mu, \sigma^2)$ を導出せよ。
(2) $L(\mu, \sigma^2)$ を用いて、μ と σ^2 の最尤推定量を求めよ。

設問 3　確率変数 X と Y の平均がいずれも 1、分散がいずれも 2 であるとする。
$S = X + 2Y, T = X - 2Y$ のとき、S と T の共分散を求めよ。

設問 4　薬 A と薬 B の効果に差がないという帰無仮説のもとで、以下の分割表についてフィッシャーの正確検定を行うことを考える。

	効果あり	効果なし	計
薬 A	3	1	4
薬 B	2	3	5
計	5	4	9

(1) 帰無仮説のもとでこの表が得られる確率は、計 9 人から無作為に 5 人を「効果あり」として選んだときに、そのうち薬 A 群が 3 人、薬 B 群が 2 人となる確率として計算できる。この確率を求めよ。
(2) フィッシャーの正確検定の p 値（両側検定）は、周辺度数を固定したときに、観察された分割表と同じか低い確率をもつ表の合計確率として定義される。上の表が観察されたときの p 値を求めよ。

Q.1 Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a population with mean μ. Show the necessary and sufficient condition for the weighted sum $\sum_{i=1}^{n} w_i X_i$ to be an unbiased estimator of μ.

Q.2 Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a normal distribution with mean μ and variance σ^2.
(1) Derive the log-likelihood function $L(\mu, \sigma^2)$.
(2) Using $L(\mu, \sigma^2)$, find the maximum likelihood estimators of μ and σ^2.

Q.3 Suppose that the means of random variables X and Y are both 1 and the variances of X and Y are both 2. Find the covariance of S and T where $S = X + 2Y$ and $T = X - 2Y$.

Q.4 Consider performing Fisher's exact test on the following contingency table under the null hypothesis that there is no difference in the effect between Drug A and Drug B.

	Effective	Not Effective	Total
Drug A	3	1	4
Drug B	2	3	5
Total	5	4	9

(1) The probability of obtaining this table under the null hypothesis can be calculated as the probability that when 5 persons are randomly selected as 'Effective' from a total of 9 persons, 3 of them are from the Drug A group and 2 of them are from the Drug B group. Compute this probability.
(2) The p value of Fisher's exact test (two-tailed test) is defined as the sum of the probabilities of the tables each having a probability equal to or lower than that of the observed table with fixed marginal frequencies. Find the p value when the table above is observed.

設問　2 次元空間中のサンプルが下の表 1 によって与えられている。ただし、各サンプルは、$\boldsymbol{x}_i = (x_{i1}, x_{i2})^{\mathrm{T}}$, $i = 1, \ldots, n$ と表され、T はベクトルまたは行列の転置を表す。以下の問いに答えよ。

表 1: サンプル

(a)	$\boldsymbol{x}_1$	$(-1, 4)^{\mathrm{T}}$
	$\boldsymbol{x}_2$	$(1, 1)^{\mathrm{T}}$
	$\boldsymbol{x}_3$	$(0, 8)^{\mathrm{T}}$
	$\boldsymbol{x}_4$	$(4, 2)^{\mathrm{T}}$
	$\boldsymbol{x}_5$	$(-2, 6)^{\mathrm{T}}$
	$\boldsymbol{x}_6$	$(4, 6)^{\mathrm{T}}$

(b)	$\boldsymbol{x}_7$	$(0, 2)^{\mathrm{T}}$

(1) 表 1(a) のサンプルに対し、以下に示されるように、2 つのクラス C_1, C_2 への分類を行った。

$$\mathrm{Class}(\boldsymbol{x}_i) = \begin{cases} C_1 & (i = 1, 3, 5) \\ C_2 & (i = 2, 4, 6) \end{cases}$$

新しく与えられるサンプル $\boldsymbol{x}_j$ に対し、そのサンプルとクラス中心との間のユークリッド距離が小さい方のクラスに分類するように決定する線形識別関数を求めよ。

(2) 表 1(a) のサンプルに対し、以下に示されるように、C_3, C_4 へのクラス分類を行ったとする。

$$\mathrm{Class}(\boldsymbol{x}_i) = \begin{cases} C_3 & (i = 3, 5, 6) \\ C_4 & (i = 1, 2, 4) \end{cases}$$

この分類と (1) で与えられた分類の良さを比較したい。それぞれのクラス内のサンプルのばらつき (クラス内分散) とクラスのばらつき (クラス間分散) の値を用いて良さを比較せよ。

(3) (1) で与えられたそれぞれのクラスについて、サンプル $\boldsymbol{x}_i$ に対するマハラノビス距離を計算する式を求めよ。

(4) 表 1(b) のサンプルが新しく加えられたとき、それを (1) のどちらのクラスに分類するのが良いかを論ぜよ。ただし、その議論の基になった、サンプルの分布に関する仮定をはっきりと示すこと。

Q. Suppose that samples are given in a 2-dimensional feature space as shown in Table 1. The samples are denoted by $\boldsymbol{x}_i = (x_{i1}, x_{i2})^{\mathrm{T}},\ (i = 1, \dots, n)$, where T represents the transpose of a vector or a matrix. Answer the following questions.

Table 1: Samples

(a)	$\boldsymbol{x}_1$	$(-1,4)^{\mathrm{T}}$
	$\boldsymbol{x}_2$	$(1,1)^{\mathrm{T}}$
	$\boldsymbol{x}_3$	$(0,8)^{\mathrm{T}}$
	$\boldsymbol{x}_4$	$(4,2)^{\mathrm{T}}$
	$\boldsymbol{x}_5$	$(-2,6)^{\mathrm{T}}$
	$\boldsymbol{x}_6$	$(4,6)^{\mathrm{T}}$

(b)	$\boldsymbol{x}_7$	$(0,2)^{\mathrm{T}}$

(1) Suppose that a classification of the samples in Table 1 (a) is given by

$$\mathrm{Class}(\boldsymbol{x}_i) = \begin{cases} C_1 & (i = 1, 3, 5) \\ C_2 & (i = 2, 4, 6) \end{cases},$$

where C_1 and C_2 represent two classes. Find the linear classifier function that classifies a new sample $\boldsymbol{x}_j$ into the class whose center is closer to the sample in Euclid distance than the other class's center.

(2) Suppose that another classification of the samples in Table 1 (a) is given by

$$\mathrm{Class}(\boldsymbol{x}_i) = \begin{cases} C_3 & (i = 3, 5, 6) \\ C_4 & (i = 1, 2, 4) \end{cases},$$

where C_3 and C_4 represent two classes. Explain how we can compare the discriminability of the classifications given in (1) and (2) by using the within-class variance concerning the distribution of the samples within a class and the between-class variance concerning the distribution of the classes.

(3) For each class given in (1), give the formula for calculating Mahalanobis distance for $\boldsymbol{x}_i$.

(4) Explain which class given in (1) is suitable for the sample in Table 1 (b). The assumption on the sample distribution required for the discussion must be clearly stated.

以下ではすべて記憶のない定常情報源を考える。なお、解答には理由も明確に示すこと。

設問 1　$\{a, b\}$ をアルファベットとし、各記号の生起確率が $P(a) = p, P(b) = 1 - p$ で与えられる情報源を考える。p を変化させた時、この情報源のエントロピーの最大値と最小値を求めよ。

設問 2　m を正の定数とする。$\{a_1, \ldots, a_{2m}\}$ をアルファベットとし、各記号の生起確率が次式で与えられる情報源を考える。

$$P(a_i) = \begin{cases} p & \text{if } i \leq m+1 \\ q & \text{otherwise} \end{cases}$$

ただし、$(m+1)p + (m-1)q = 1$ とする。p, q を変化させた時、この情報源のエントロピーの最大値と最小値を求めよ。

設問 3　$\{a_1, a_2, a_3\}$ をアルファベットとし、各記号の生起確率が $P(a_1) = P(a_2) = P(a_3) = \frac{1}{3}$ である情報源を考える。この情報源に対する 2 元ハフマン符号の平均長を求めよ。

設問 4　m を正整数とし、$n = 2^m + 1$ とする。$\{a_1, \ldots, a_n\}$ をアルファベットとし、すべての記号の生起確率が $P(a_i) = \frac{1}{n}$ である情報源を考える。この情報源に対する 2 元ハフマン符号の平均長を求めよ。

設問 5　次の通信路行列により定まる通信路の容量を求めよ。なお、入力アルファベットのサイズは 2、出力アルファベットのサイズは 3 である。

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

設問 6　次の通信路行列により定まる通信路の容量を C とする。なお、入力アルファベットのサイズは 2、出力アルファベットのサイズは n である。

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2(n-1)} & \cdots & \frac{1}{2(n-1)} & \frac{1}{2(n-1)} \\ \frac{1}{2(n-1)} & \frac{1}{2(n-1)} & \cdots & \frac{1}{2(n-1)} & \frac{1}{2} \end{bmatrix}$$

すると、n に関する関数 $f(n)$ を用いて、$C = f(n) \log_2 n + \frac{1}{2} \log_2 (n-1) + \frac{n}{2(n-1)}$ と書ける。$f(n)$ を求めよ。

Consider a stationary memoryless source in each question. The reason must be clearly stated for each answer.

Q.1 Consider a source that has a source alphabet $\{a, b\}$ with the probabilities $P(a) = p$ and $P(b) = 1 - p$. Find the maximum and minimum values of the entropy of this source when varying p.

Q.2 Let m be a positive constant. Consider a source that has a source alphabet $\{a_1, \dots, a_{2m}\}$ with the probabilities given by

$$P(a_i) = \begin{cases} p & \text{if } i \leq m+1 \\ q & \text{otherwise} \end{cases},$$

where $(m+1)p + (m-1)q = 1$. Find the maximum and minimum values of the entropy of this source when varying p and q.

Q.3 Consider a source that has a source alphabet $\{a_1, a_2, a_3\}$ with the probabilities $P(a_1) = P(a_2) = P(a_3) = \frac{1}{3}$. Compute the average codeword length of the binary Huffman code for this source.

Q.4 Let m be a positive integer and $n = 2^m + 1$. Consider a source that has a source alphabet $\{a_1, \dots, a_n\}$ with the probability $P(a_i) = \frac{1}{n}$ for each i. Compute the average codeword length of the binary Huffman code for this source.

Q.5 Consider a communication channel with a probability transition matrix given by

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix},$$

where the sizes of the input and output alphabets are two and three, respectively. Compute the capacity of this channel.

Q.6 Consider a communication channel with a probability transition matrix given by

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2(n-1)} & \cdots & \frac{1}{2(n-1)} & \frac{1}{2(n-1)} \\ \frac{1}{2(n-1)} & \frac{1}{2(n-1)} & \cdots & \frac{1}{2(n-1)} & \frac{1}{2} \end{bmatrix},$$

where the sizes of the input and output alphabets are two and n, respectively. The capacity of this channel is given by $C = f(n) \log_2 n + \frac{1}{2} \log_2(n-1) + \frac{n}{2(n-1)}$, where $f(n)$ is a function of n. Show $f(n)$.

設問 1　2 次元信号 $f(x,y)$ の 2 次元フーリエ変換を

$$F(u,v) = \iint_{-\infty}^{\infty} f(x,y)\mathrm{e}^{-\mathrm{j}(ux+vy)}\mathrm{d}x\mathrm{d}y$$

とする。ただし j は虚数単位である。また $f(x,y)$ のある軸 ℓ への投影を、軸 ℓ 上の各点における、ℓ に垂直な直線に沿った $f(x,y)$ の線積分とする。以下の問いに答えよ。

(1) $f(x,y)$ を x 軸に投影した信号 $p(x)$ の 1 次元フーリエ変換を、$F(u,v)$ を用いて表せ。

(2) 原点を中心として x 軸を反時計回りに角度 θ 回転して得られた s 軸上に $f(x,y)$ を投影した信号を $p_\theta(s)$ とする。$p_\theta(s)$ の s についての 1 次元フーリエ変換を、$F(u,v)$ を用いて表せ。

設問 2　長さ N の離散時間信号 $x[n]$ の N 点離散フーリエ変換 $X[k]$ を

$$X[k] = \sum_{n=0}^{N-1} x[n]W_N^{kn}, \quad W_N = \mathrm{e}^{-\mathrm{j}\frac{2\pi}{N}}$$

とする。ただし j は虚数単位、$n,k = 0,\ldots,N-1$ であり、N は正の偶数とする。以下の問いに答えよ。

(1) 観測系列 $x_0[n] = \{x_0[0], x_0[1], x_0[2], x_0[3]\} = \{1,2,1,-2\}$ を、ある信号を 4000Hz で等間隔にサンプリングすることで得たとする。$x_0[n]$ の 4 点離散フーリエ変換を計算し、周波数 (Hz) に対する振幅スペクトルおよび位相スペクトルを図示せよ。

(2) 2 つの要素数 N の実数値系列 $x_1[n]$ および $x_2[n]$ の N 点離散フーリエ変換を、1 回の N 点離散フーリエ変換によって計算する方法を導出せよ。

(3) 要素数 $2N$ の実数値系列の $2N$ 点離散フーリエ変換を、1 回の N 点離散フーリエ変換によって計算する方法を導出せよ。

Q.1 Suppose the 2D Fourier transform of a 2D signal $f(x, y)$ is defined as

$$F(u, v) = \iint_{-\infty}^{\infty} f(x, y) \mathrm{e}^{-\mathrm{j}(ux+vy)} \mathrm{d}x\mathrm{d}y\,,$$

where j is the imaginary unit. Let the projection of $f(x, y)$ to an axis ℓ be defined as the line integral at each point on ℓ along the direction orthogonal to ℓ. Answer the following questions.

(1) $p(x)$ is the projection of $f(x, y)$ to the x axis. Express the 1D Fourier transform of $p(x)$ using $F(u, v)$.

(2) $p_\theta(s)$ is the projection of $f(x, y)$ to the s axis given by rotating the x axis around the origin counterclockwisely by angle θ. Express the 1D Fourier transform of $p_\theta(s)$ about s using $F(u, v)$.

Q.2 Let the N-point discrete Fourier transform $X[k]$ of a discrete-time signal $x[n]$ of length N be

$$X[k] = \sum_{n=0}^{N-1} x[n] W_N^{kn}\,, \quad W_N = \mathrm{e}^{-\mathrm{j}\frac{2\pi}{N}}\,,$$

where j is the imaginary unit, N is a positive even integer, and $n, k = 0, \ldots, N-1$. Answer the following questions.

(1) Consider an observation signal $x_0[n] = \{x_0[0], x_0[1], x_0[2], x_0[3]\} = \{1, 2, 1, -2\}$ given by sampling a signal regularly at 4000 Hz. Compute the 4-point discrete Fourier transform of $x_0[n]$, and draw its magnitude and phase spectra against frequency in Hz.

(2) Derive the N-point discrete Fourier transforms of two real-valued sequences $x_1[n]$ and $x_2[n]$ of length N using a single execution of the N-point discrete Fourier transform.

(3) Derive the $2N$-point discrete Fourier transform of a real-valued sequence of length $2N$ using a single execution of the N-point discrete Fourier transform.

文法 $G = (\Sigma, N, P, S)$ を考える。ここで、Σ, N, P, S はそれぞれ終端記号の有限集合、非終端記号の有限集合、生成規則の有限集合、開始記号である。

設問 1 以下は、文脈自由文法についての記述である。空欄 (1)(2)(3) を埋めよ。

> 文法 G が文脈自由文法であるとは、P の各生成規則 $\alpha \to \beta$ において $\alpha \in$ $\boxed{\quad(1)\quad}$ かつ $\beta \in$ $\boxed{\quad(2)\quad}$ であるときをいう。$\boxed{\quad(3)\quad}$ 上の言語 L は、それがある文脈自由文法 G によって生成されるとき、文脈自由言語という。

設問 2 $L_{ab} = \{a^m b^m c^n \mid m, n > 0\}$ および $L_{bc} = \{a^m b^n c^n \mid m, n > 0\}$ が文脈自由言語であることを示せ。

設問 3 文脈自由言語のクラスが連接 $L_1 \cdot L_2$ のもとで閉じていることを、L_1 と L_2 をそれぞれ生成する文法 $G_1 = (\Sigma, N_1, P_1, S_1)$ と $G_2 = (\Sigma, N_2, P_2, S_2)$ に対して $G_3 = (\Sigma, N_3, P_3, S_3)$ を構成することで証明せよ。

設問 4 文脈自由言語のクラスが和 $L_1 \cup L_2$ のもとで閉じていることを、文脈自由文法 $G_1 = (\Sigma, N_1, P_1, S_1)$ と $G_2 = (\Sigma, N_2, P_2, S_2)$ に対して $G_3 = (\Sigma, N_3, P_3, S_3)$ を構成することで証明せよ。

設問 5 $L_{abc} = \{a^k b^k c^k \mid k > 0\}$ は文脈自由言語ではない。このことを用いて文脈自由言語のクラスが補集合演算 $\overline{L}$ のもとで閉じていないことを証明せよ。

設問 6 文脈自由言語のクラスは差 $L_1 - L_2$ のもとで閉じていないことを証明せよ。

We consider a grammar $G = (\Sigma, N, P, S)$, where Σ, N, P, and S are a finite set of terminal symbols, a finite set of nonterminal symbols, a finite set of production rules, and the start symbol, respectively.

Q.1 The following is a description on context-free grammar. Fill the blanks (1), (2), and (3).

A grammar G is a context-free grammar when, for each production rule $\alpha \to \beta$ in P, $\alpha \in$ $\boxed{\quad (1) \quad}$ and $\beta \in$ $\boxed{\quad (2) \quad}$. A language L on $\boxed{\quad (3) \quad}$ is a context-free language when it is generated by a context-free grammar G.

Q.2 Prove that $L_{ab} = \{a^m b^m c^n \mid m, n > 0\}$ and $L_{bc} = \{a^m b^n c^n \mid m, n > 0\}$ are context-free languages.

Q.3 Prove that the class of context-free languages is closed under concatenation $L_1 \cdot L_2$ by constructing $G_3 = (\Sigma, N_3, P_3, S_3)$ from the context-free grammars $G_1 = (\Sigma, N_1, P_1, S_1)$ and $G_2 = (\Sigma, N_2, P_2, S_2)$ generating L_1 and L_2, respectively.

Q.4 Prove that the class of context-free languages is closed under union $L_1 \cup L_2$ by constructing $G_3 = (\Sigma, N_3, P_3, S_3)$ from the context-free grammars $G_1 = (\Sigma, N_1, P_1, S_1)$ and $G_2 = (\Sigma, N_2, P_2, S_2)$.

Q.5 $L_{abc} = \{a^k b^k c^k \mid k > 0\}$ is not a context-free language. By using this fact prove that the class of context-free languages is not closed under complement $\overline{L}$.

Q.6 Prove that the class of context-free languages is not closed under difference $L_1 - L_2$.