2020 年度

東京大学大学院情報理工学系研究科 コンピュータ科学専攻 入学試験問題

専門科目 II

2019 年 8 月 20 日

13:00 – 14:30

注意事項

(1) 試験開始の合図があるまで, この問題冊子を開けないこと.

(2) 6 題の中から 2 題を選んで答えよ. 問題ごとに 1 枚の解答用紙を使用すること.

(3) 解答用紙および問題冊子は持ち帰らないこと.

Specialized Subjects II

13:00 – 14:30, August 20, 2019

Entrance Examination (AY 2020)

Department of Computer Science, Graduate School of Information Science and Technology

The University of Tokyo

Notice:

(1) Do not open this problem booklet until the start of the examination is announced.

(2) Choose and answer 2 problems out of the following 6 problems. Use one answer sheet for one problem.

(3) Do not take the problem booklet or any answer sheet out of the examination room.

下欄に受験番号を記入すること.

Write your examinee's number in the box below.

受験番号	No.

問題 1

A は命題変数，L_i はリテラル（命題変数もしくは命題変数の否定）とする．この問題では，

$$L_1 \wedge \ldots \wedge L_n \supset A$$

という形の命題論理式を節と呼ぶ．ただし，$n=1$ のときは $L_1 \supset A$，$n=0$ のときは A という形になる．以下，Π を節の集合とし，M は命題変数の集合とする．M に属する命題変数を真，その他の命題変数を偽と解釈したときに，Π に属するすべての節が真になるならば，M を Π のモデルと呼ぶ．モデルの間には集合の包含関係が自然に定義される．

以下の問いに答えよ．

(1) $\Pi_0 = \{P, P \supset Q, Q \wedge \neg R \supset S, P \wedge \neg S \wedge \neg T \supset T\}$ としたとき，$\{P, Q, R, S, T\}$ の部分集合で Π_0 のモデルであるものをすべてあげよ．

以下では，(i) Π の節のうち，M に属する命題変数の否定を $\supset$ の左に含むものは除き，(ii) 残された節の否定リテラル（命題変数の否定）をすべて取り除いて得られる節の集合を Π_M と書く．

(2) 問い (1) の Π_0 に対して，$M_0 = \{P, Q, S\}$ のとき，Π_{0M_0} は何になるか．

(3) Π_M のモデル M' が $M \subseteq M'$ を満たすならば，M' は Π のモデルであることを示せ．

(4) Π のモデル M' が $M' \subseteq M$ を満たすならば，M' は Π_M のモデルであることを示せ．

(5) 問い (2) の Π_0 と M_0 に対して，Π_{0M_0} の最小モデルを求めよ．ただし，Π_M のモデル M' は，Π_M のすべてのモデル M'' に対して $M' \subseteq M''$ が成り立つとき，Π_M の最小モデルと呼ばれる．

(6) Π_M の最小モデルが M に一致するならば，M は Π の極小モデルであることを示せ．ただし，Π のモデル M' は，$M'' \subsetneq M'$ を満たす Π のモデル M'' が存在しないとき，Π の極小モデルと呼ばれる．

(7) Π の極小モデル M は常に Π_M の最小モデルか．そうならば証明し，違うならば反例を示せ．

Problem 1

Let A be a propositional variable, and L_i be a literal (i.e., a propositional variable or negation of a propositional variable). In this problem, a propositional formula of the following form is called a clause.

$$L_1 \wedge \ldots \wedge L_n \supset A$$

If $n = 1$, it is of the form $L_1 \supset A$, and if $n = 0$, it is of the form A. Hereinafter, Π is a set of clauses, and M is a set of propositional variables. If all clauses in Π are true under the interpretation in which all propositional variables in M are true and the other variables are false, then M is called a model of Π. The inclusion relation between sets is naturally defined between models.

Answer the following questions.

(1) Let $\Pi_0 = \{P, P \supset Q, Q \wedge \neg R \supset S, P \wedge \neg S \wedge \neg T \supset T\}$. Enumerate all the subsets of $\{P, Q, R, S, T\}$ that are models of Π_0.

We write Π_M for the set of clauses obtained from Π by (i) removing all the clauses that contain negation of a propositional variable in M in the left hand side of $\supset$, and then (ii) deleting all the negated literals (negation of propositional variables) from the remaining clauses.

(2) For Π_0 in question (1), if $M_0 = \{P, Q, S\}$, what is Π_{0M_0}?

(3) Show that if a model M' of Π_M satisfies $M \subseteq M'$, then M' is a model of Π.

(4) Show that if a model M' of Π satisfies $M' \subseteq M$, then M' is a model of Π_M.

(5) For Π_0 and M_0 in question (2), obtain the minimum model of Π_{0M_0}. Here, a model M' of Π_M is called a minimum model of Π_M if $M' \subseteq M''$ holds for every model M'' of Π_M.

(6) Show that if the minimum model of Π_M coincides with M, then M is a minimal model of Π. Here, a model M' of Π is called a minimal model of Π if there does not exist a model M'' of Π such that $M'' \subsetneq M'$.

(7) Is a minimal model M of Π always a minimum model of Π_M? If so, prove the fact. Otherwise, give a counterexample.

問題 2

自然言語の文（単語列）$\mathbf{w} = (w_1, \ldots, w_\ell)$ に対し，品詞タグ列 $\mathbf{t} = (t_1, \ldots, t_\ell)$ を求める問題を考える．例えば，以下の 4 単語の文

$$(w_1, w_2, w_3, w_4) = (\text{John, wrote, a, book})$$

に対し，以下のように品詞タグ列を出力したい．

$$(t_1, t_2, t_3, t_4) = (\mathsf{NOUN, VERB, DET, NOUN})$$

ただし，この例では NOUN, VERB, DET は品詞タグであり，それぞれ名詞，動詞，限定詞を表す．W を全ての単語の有限集合，T を全ての品詞タグの有限集合とする．また，学習データとして品詞タグ付きコーパス $D = \{(\mathbf{w}^{(k)}, \mathbf{t}^{(k)}) | k \in \{1, \ldots, N\}\}$（$\mathbf{w}^{(k)}$ は D の k 番目の文，$\mathbf{t}^{(k)}$ はその品詞タグ列，$N > 0$ は D の要素数）が与えられるとする．以下では，単語列 $\mathbf{w} = (w_1, \ldots, w_\ell)$ に対し，その長さ ℓ を $|\mathbf{w}|$ と表記する．

以下の問いに答えよ．

(1) 単語 $w \in W$ に対して品詞タグ $t \in T$ を与える確率 $p_u(t|w)$ を考え，単語列 $\mathbf{w}$ に対して品詞タグ列 $\mathbf{t}$ を与える確率関数 $p_u(\mathbf{t}|\mathbf{w})$ を以下のように定義する．

$$p_u(\mathbf{t}|\mathbf{w}) \equiv \prod_{i=1}^{|\mathbf{w}|} p_u(t_i|w_i)$$

学習データ D の各要素はそれぞれ独立に $p_u(\mathbf{t}|\mathbf{w})$ に従って分布すると仮定したとき，$p_u(t|w)$ の最尤推定量を求める方法を答えよ．

(2) 各 $w \in W, t \in T$ に対して $p_u(t|w)$ が与えられたとき，入力文 $\mathbf{w}$ に対して $p_u(\mathbf{t}|\mathbf{w})$ が最大となる品詞タグ列 $\mathbf{t}$ を求めるアルゴリズムを与えよ．

(3) 直前の単語が $v \in W$ の時に単語 $w \in W$ に対して品詞タグ $t \in T$ を与える確率 $p_b(t|v, w)$ を考える．ただし，文頭に出現する w については v は特殊な単語<s>とする．確率関数 $p_b(\mathbf{t}|\mathbf{w})$ を以下のように定義する．

$$p_b(\mathbf{t}|\mathbf{w}) \equiv p_b(t_1|\text{<s>}, w_1) \prod_{i=2}^{|\mathbf{w}|} p_b(t_i|w_{i-1}, w_i)$$

学習データ D の各要素はそれぞれ独立に $p_b(\mathbf{t}|\mathbf{w})$ に従って分布すると仮定したとき，$p_b(t|v, w)$ の最尤推定量を求める方法を答えよ．

また，各 $v \in W \cup \{\text{<s>}\}, w \in W, t \in T$ に対して $p_b(t|v, w)$ が与えられたとき，入力文 $\mathbf{w}$ に対して $p_b(\mathbf{t}|\mathbf{w})$ が最大となる品詞タグ列 $\mathbf{t}$ を求めるアルゴリズムを与えよ．

(4) 問い (1)〜(3) の手法と比べて，隠れマルコフモデルを用いることでより高精度な品詞タグ付けが期待できる理由を説明せよ．隠れマルコフモデルの定義とそれを用いた品詞タグ付けアルゴリズムについて述べ，問い (1)〜(3) の手法では誤った品詞タグが出力されるが隠れマルコフモデルでは正しい品詞タグが出力される具体例を挙げて説明を行うこと．

Problem 2

Consider the problem of obtaining a sequence of part-of-speech (POS) tags $\mathbf{t} = (t_1, \dots, t_\ell)$ for a given natural language sentence (word sequence) $\mathbf{w} = (w_1, \dots, w_\ell)$. For example, for the following four-word sentence

$$(w_1, w_2, w_3, w_4) = (\text{John, wrote, a, book}),$$

our goal is to output the following POS tag sequence.

$$(t_1, t_2, t_3, t_4) = (\mathsf{NOUN}, \mathsf{VERB}, \mathsf{DET}, \mathsf{NOUN})$$

In this example, NOUN, VERB, and DET are POS tags, denoting noun, verb, and determiner, respectively. W is a finite set of all words, and T is a finite set of all POS tags. Suppose that a POS-tagged corpus $D = \{(\mathbf{w}^{(k)}, \mathbf{t}^{(k)}) | k \in \{1, \dots, N\}\}$ ($\mathbf{w}^{(k)}$ is a k-th sentence in D, $\mathbf{t}^{(k)}$ is its POS tag sequence, and $N > 0$ is the number of elements in D) is given as training data. In the following, for a word sequence $\mathbf{w} = (w_1, \dots, w_\ell)$, its length ℓ is represented as $|\mathbf{w}|$.

Answer the following questions.

(1) Consider the probability $p_u(t|w)$ for assigning a POS tag $t \in T$ to a word $w \in W$, and define the probability function $p_u(\mathbf{t}|\mathbf{w})$ for assigning a POS tag sequence $\mathbf{t}$ to a word sequence $\mathbf{w}$ as follows.

$$p_u(\mathbf{t}|\mathbf{w}) \equiv \prod_{i=1}^{|\mathbf{w}|} p_u(t_i|w_i)$$

Supposing each element of trainining data D is independently distributed following $p_u(\mathbf{t}|\mathbf{w})$, answer a method for computing the maximum likelihood estimate of $p_u(t|w)$.

(2) Assume that $p_u(t|w)$ is given for each $w \in W, t \in T$. Describe an algorithm to obtain a POS tag sequence $\mathbf{t}$ that maximizes $p_u(\mathbf{t}|\mathbf{w})$ for an input sentence $\mathbf{w}$.

(3) Consider the probability $p_b(t|v, w)$ for assigning a POS tag $t \in T$ to a word $w \in W$ when a word $v \in W$ immediately precedes w in a sentence. Note that v is considered as a special word <s> when w is the first word of the sentence. The probability function $p_b(\mathbf{t}|\mathbf{w})$ is defined as follows.

$$p_b(\mathbf{t}|\mathbf{w}) \equiv p_b(t_1|\texttt{<s>}, w_1) \prod_{i=2}^{|\mathbf{w}|} p_b(t_i|w_{i-1}, w_i)$$

Supposing each element of trainining data D is independently distributed following $p_b(\mathbf{t}|\mathbf{w})$, answer a method for computing the maximum likelihood estimate of $p_b(t|v, w)$.

Also, assuming that $p_b(t|v, w)$ is given for each $v \in W \cup \{\texttt{<s>}\}, w \in W, t \in T$, describe an algorithm to obtain a POS tag sequence $\mathbf{t}$ that maximizes $p_b(\mathbf{t}|\mathbf{w})$ for an input sentence $\mathbf{w}$.

(4) Explain why POS tagging using hidden Markov models is expected to attain higher accuracy than the methods described in questions (1) to (3). You must describe the definition of hidden Markov models and the POS tagging algorithm using hidden Markov models, and provide an explanation including an example where the methods described in questions (1) to (3) output a wrong POS tag but the POS tagging using hidden Markov models outputs a correct POS tag.

問題 3

3次元空間における物体の姿勢を，ある標準姿勢からの回転として表現することについて考える．以下では，図1や図2に示す4つの立方体からなる物体を考える（ただし隣接する立方体同士は面を共有する）．図1の姿勢を標準姿勢とする．なお，図1では左下の立方体の中心が原点と一致し，各立方体の中心は x,y,z 軸のいずれかの軸上に置かれているものとする．図2はこの物体の別な姿勢であり，右下の立方体の中心が原点と一致し，各立方体の中心は x,y,z 軸のいずれかの軸上に置かれているものとする．物体を図示する際には，図1および図2と同様の視点から描くものとし，x,y,z 軸も描くこと．角度の単位にはラジアンを用いる．

以下の問いに答えよ．

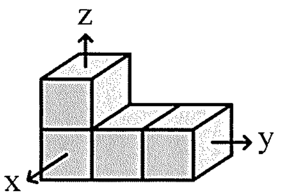

図1

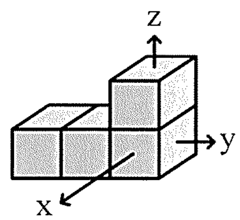

図2

物体の姿勢を，図1の標準姿勢からの回転角の3つ組 $(\theta_x, \theta_y, \theta_z)$（オイラー角と呼ぶ）で表現することを考える．具体的には，x 軸回りに角度 θ_x，y 軸回りに角度 θ_y，z 軸回りに角度 θ_z の回転をこの順に適用するものとし，回転角は回転軸の負から正の方向にみて時計回りを正とする．

(1) オイラー角 $(\pi, \pi, 0)$ の表す姿勢を図示せよ．

(2) 標準姿勢を表す $(0, 0, 0)$ と $(\pi, \pi, 0)$ の2つのオイラー角の要素ごとの算術平均として与えられるオイラー角を答えよ．そのオイラー角の表す姿勢を図示せよ．

物体の姿勢を，図1の標準姿勢からの回転を表す 3×3 の変換行列で表現することを考える．なお，回転前の物体上の点の3次元座標を v としたとき，変換行列 R による変換後のその点の3次元座標は Rv で与えられるものとする．

(3) 図2の姿勢を表現する変換行列を答えよ．

(4) 図1と図2の姿勢を表す変換行列の要素ごとの算術平均として与えられる変換行列を答えよ．また，その平均変換行列で，図1に示される物体を変換した結果の形状を図示して説明せよ．

物体の姿勢を，図1の標準姿勢からの回転を表すクオータニオンで表現することを考える．クオータニオンとは4次元単位ベクトルであり，ある3次元単位ベクトル $v = (v_x, v_y, v_z)$ 周りの原点を中心とする回転角 θ の3次元回転を表すクオータニオンは，$(v_x \sin\frac{\theta}{2}, v_y \sin\frac{\theta}{2}, v_z \sin\frac{\theta}{2}, \cos\frac{\theta}{2})$ で表される．なお，3次元単位ベクトル周りの回転の回転角は，その単位ベクトルを始点からみて時計回りを正とする．

(5) 図2の姿勢を表現するクオータニオンを答えよ．答えは2通りあるので，両方とも示せ．

(6) 図1と図2の姿勢を表す2つのクオータニオンの球面線形平均として与えられるクオータニオンを答えよ．答えは2通りあるので，両方とも示せ．また，それぞれの平均クオータニオンの表す姿勢を図示せよ．なお，2つのクオータニオン q_1,q_2 の球面線形平均は $\frac{\sin 0.5\varphi}{\sin\varphi} q_1 + \frac{\sin 0.5\varphi}{\sin\varphi} q_2$，$\varphi = \cos^{-1}(q_1 \cdot q_2)$　$(0 \leq \varphi \leq \pi)$ で与えられる（ただし $\sin\varphi = 0$ のときの球面線形平均は未定義とする）．

Problem 3

Let us consider representing an orientation of an object in a three-dimensional space as a rotation from a canonical orientation. In the following, we consider the orientation of the object consisting of four cubes as shown in Figures 1 and 2 (note that each pair of cubes placed side by side shares a surface). We also define the canonical orientation as the one illustrated in Figure 1. Note that the center of the cube at the lower left in Figure 1 is placed at the origin, and the center of each cube lies on the x, y, or z axis. Figure 2 shows another orientation of this object, in which the center of the cube at the lower right is placed at the origin, and the center of each cube lies on the x, y, or z axis. When you draw another orientation of this object, you need to draw it with the x,y,z axes from the same point of view as Figures 1 and 2. Angles must be represented in radian.

Answer the following questions.

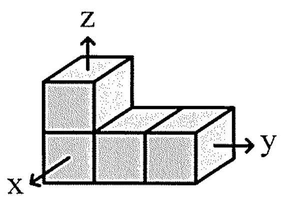

Figure 1

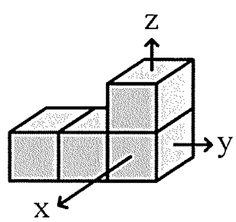

Figure 2

Let us represent an orientation of the object using a triplet $(\theta_x, \theta_y, \theta_z)$ (called, Euler angles) of the angles of rotations from the canonical orientation given in Figure 1. Namely, the object is rotated around the x axis by the angle θ_x, around the y axis by the angle θ_y, and around the z axis by the angle θ_z in this order, and the angle of a rotation around each axis is positive if the rotation is in the clockwise direction when viewed from the negative-to-positive direction of the axis.

(1) Draw the object with the orientation represented by Euler angles $(\pi, \pi, 0)$.

(2) Answer Euler angles given as the elementwise arithmetic mean of the two triplets of Euler angles, $(0, 0, 0)$ (corresponding to the canonical orientation) and $(\pi, \pi, 0)$. Also, draw the object with the orientation represented by the mean Euler angles.

Let us represent an orientation of the object using a 3×3 transformation matrix that corresponds to a rotation from the canonical orientation given in Figure 1. Note that a point on the object at a three-dimensional coordinate v moves to another three-dimensional coordinate Rv through the rotation by a transformation matrix R.

(3) Answer the transformation matrix that represents the orientation shown in Figure 2.

(4) Answer the transformation matrix given as the elementwise arithmetic mean of the two transformation matrices corresponding to the orientations shown in Figures 1 and 2. Also, draw and describe the shape of the object obtained by applying the mean transformation matrix to the object shown in Figure 1.

Let us represent an orientation of the object using a quaternion that expresses a rotation from the canonical orientation given in Figure 1. A quaternion is a four-dimensional unit vector, and a quaternion that expresses the three-dimensional rotation centered at the origin around a three-dimensional unit vector $v = (v_x, v_y, v_z)$ with an angle θ is represented as $(v_x \sin\frac{\theta}{2}, v_y \sin\frac{\theta}{2}, v_z \sin\frac{\theta}{2}, \cos\frac{\theta}{2})$.

Note that the angle of a rotation around the three-dimenstional unit vector is positive if the rotation is in the clockwise direction when the unit vector is viewed from its start point.

(5) Answer the quaternion that corresponds to the orientation shown in Figure 2. There are two answers for this question; give both of the two answers.

(6) Answer the quaternion given as the spherical linear average of the two quaternions corresponding to the orientations shown in Figures 1 and 2. There are two answers for this question; give both of the two answers. Also, draw the object with the orientation represented by each of the average quaternions. Note that the spherical linear average of two quaternions q_1 and q_2 is given as $\frac{\sin 0.5\varphi}{\sin\varphi} q_1 + \frac{\sin 0.5\varphi}{\sin\varphi} q_2$ where $\varphi = \cos^{-1}(q_1 \cdot q_2)$ $(0 \leq \varphi \leq \pi)$; the spherical linear average is undefined when $\sin\varphi = 0$.

問題 4

各辺がそれぞれ正の重みを持つ連結な無向グラフ $G=(V,E)$ を考える．G より辺をいくつか除去することで得られる G の部分グラフ $G'=(V,E')$ が木となっているとき，G' を G の全域木という．全域木の重みとは，全域木に含まれるすべての辺の重みの和のことをいう．G の最小全域木とは，G の重み最小の全域木のことをいう．なお，以下の問いにおいて，グラフや木のデータ表現については適切な表現を仮定してよい．

以下の問いに答えよ．

(1) グラフ G の任意の閉路 C に対し，C に含まれる辺のうち最大重みの辺（複数ある場合はそのうちの任意の 1 つ）を e とおく．このとき，e を含まない G の最小全域木があることを証明せよ．

(2) グラフ $G=(V,E)$ の頂点集合 V の任意の部分集合 V' $(V'\neq V, V'\neq\emptyset)$ に対し，$u\in V'$, $v\in V-V'$ であるような辺 $(u,v)\in E$ のうち最小の重みの辺（複数あるならばそのうちの任意の 1 つ）を e とおく．このとき，e を含む G の最小全域木があることを証明せよ．なお，$\emptyset$ は空集合を表す．

(3) グラフ $G=(V,E)$ 上の 2 頂点 $u,v\in V$ 間の任意の 1 本のパスを $O(|E|)$ 時間で求めるアルゴリズムを示せ．

(4) グラフ $G=(V,E)$ とその最小全域木 T が与えられているものとする．G に重み $w>0$ の新たな辺 $e=(u,v)\notin E$ $(u,v\in V)$ を追加して得られるグラフ G' の最小全域木を $O(|V|)$ 時間で求めるアルゴリズムを示せ．

(5) 問い (4) で示したアルゴリズムの正しさを証明せよ．

Problem 4

Consider a connected undirected graph $G = (V, E)$ with positive edge weights. A subgraph $G' = (V, E')$ of G obtained by removing some of the edges in G is called a spanning tree of G, if G' is a tree. The summation of weights of all the edges in a spanning tree is called the weight of the spanning tree. A minimum spanning tree of G is a spanning tree of G whose weight is minimum. You can assume appropriate data representation for graphs and trees in the questions below.

Answer the following questions.

(1) Let e be the edge (or arbitrary one of the edges if there are multiple such edges) with the maximum weight in some arbitrary cycle C in G. Prove that there is a minimum spanning tree of G that does not contain e.

(2) Consider an arbitrary vertex subset V' of V ($V' \neq V, V' \neq \emptyset$) for $G = (V, E)$. Let e be the edge (or arbitrary one of the edges if there are multiple such edges) with the minimum weight among the edges $(u, v) \in E$ such that $u \in V'$ and $v \in V - V'$. Prove that there is a minimum spanning tree that contains e. Note that $\emptyset$ denotes an empty set.

(3) Describe an $O(|E|)$-time algorithm that finds an arbitrary path between two nodes $u, v \in V$ on graph $G = (V, E)$.

(4) Assume that we are given a graph $G = (V, E)$ and its minimum spanning tree T. Let G' be the graph obtained by adding to G a new edge $e = (u, v) \notin E$ $(u, v \in V)$ with weight $w > 0$. Describe an $O(|V|)$-time algorithm that finds a minimum spanning tree of G'.

(5) Prove the correctness of the algorithm described in question (4).

問題 5

$f(x)$ を a から b $(a < b)$ までの閉区間で定義された実数の関数とせよ．n を 2 以上の整数とし，$h = (b-a)/n$ と定義する．各整数 $i = 0, 1, \ldots, n$ について，$x_i = a + ih$ および $f_i = f(x_i)$ とそれぞれ定義する．つまり $x_0, \ldots, x_n$ は区間 a から b までを n 等分する点であり，f_i は $x = x_i$ での関数 $f(x)$ の値である．

次に $J = \int_a^b f(x)dx$ と定義し，区間 a から b を n 等分する点を用いた複合台形公式による J の近似値を J_n と定義する．

以下の問いに答えよ．

(1) $f(x)$ を 4 回連続微分可能な関数とせよ．整数 k は $0 < k < n$ と仮定し，f''_k を x_k における $f(x)$ の 2 階微分と定義せよ．f''_k の近似値であってその誤差が $O(h^2)$ であるものを f_{k-1}, f_k, f_{k+1} の線型結合で構成せよ．

(2) 問い (1) で得られた近似式は h が 0 に近づくほど高精度になると思われる．IEEE 754 倍精度浮動小数点演算を用いてこの計算をする時，これが事実であるかどうか理由とともに述べよ．

(3) n，h と f_i $(i = 0, \ldots, n)$ を用いて J_n を書き表せ．

(4) n 等分されてできる各区間の中で $f(x)$ はそれぞれ 2 次関数で表せると仮定せよ．そして，元の各区間をさらに 2 等分してできる $2n$ 分割に対して J_n と同様に J_{2n} を定義する．J_{2n} と J_n を用いて $E_n = J_n - J$ を表せ．

Problem 5

Suppose that $f(x)$ is a real function defined on a closed interval from a to b $(a < b)$. Suppose that n is an integer that is no less than 2, and define $h = (b-a)/n$. Then, for each integer $i = 0, 1, \ldots, n$, define $x_i = a + ih$ and $f_i = f(x_i)$, respectively. Namely, $x_0, \ldots, x_n$ are the points that divide the interval from a to b into n equal parts, and f_i is the value of the function $f(x)$ at $x = x_i$.

Next, define $J = \int_a^b f(x)dx$, and define J_n as the approximate value calculated by the composite trapezoid rule applied on J using the points which divide the interval from a to b into n equal parts.

Answer the following questions.

(1) Assume that $f(x)$ is a four times continuously differentiable function. Let k be an integer such that $0 < k < n$ and define f_k'' as the second order differential of $f(x)$ at x_k. Express an approximate value of f_k'' whose error is $O(h^2)$, as a linear combination of f_{k-1}, f_k, and f_{k+1}.

(2) The approximation obtained by question (1) seems to become accurate when h approaches zero. Answer, with a reason, whether this is correct or not in the calculation with the IEEE 754 double precision floating point operations.

(3) Express J_n using n, h, and f_i $(i = 0, \ldots, n)$.

(4) Assume that $f(x)$ can be expressed by a quadratic function in each interval formed by the division into n equal parts. Then, define J_{2n} similarly using the division into $2n$ equal parts composed by the division of each original part into two halves. Express $E_n = J_n - J$ using J_{2n} and J_n.

問題 6

平均 $\mu \in \mathbb{R}$，分散 $\sigma^2 > 0$ の正規分布 $N(\mu, \sigma^2)$ の確率密度関数は

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

で与えられる．X, Z を互いに独立でそれぞれ $N(\mu, 1)$, $N(0, 1)$ に従う確率変数とし，ある定数 $\theta \in \mathbb{R}$ に対して $Y = \theta X + Z$ と定義する．また，整数 $n > 1$ に対して $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ を互いに独立かつ (X, Y) と同一分布に従う 2 次元確率変数とし，$X^{(n)} = (X_1, X_2, \ldots, X_n)$, $Y^{(n)} = (Y_1, Y_2, \ldots, Y_n)$ と表す．

以下の問いに答えよ．

(1) Y の期待値 $\mathbb{E}[Y]$ および分散 $\mathbb{V}[Y]$ を μ, θ を用いて表せ．

(2) Y が与えられたときの X の条件付き確率分布が正規分布となることを示し，その期待値 $\mathbb{E}[X|Y]$ および分散 $\mathbb{V}[X|Y]$ を μ, θ, Y を用いて表せ．

(3) $(X^{(n)}, Y^{(n)})$ の実現値を $(x^{(n)}, y^{(n)})$ で表す．$(X^{(n)}, Y^{(n)})$ の同時確率密度関数 $p_{\mu,\theta}(x^{(n)}, y^{(n)})$ を $\mu, \theta, x^{(n)} = (x_1, x_2, \ldots, x_n), y^{(n)} = (y_1, y_2, \ldots, y_n)$ を用いて表せ．

(4) $(X^{(n)}, Y^{(n)})$ のうち X_n の観測が欠損している場合，すなわち $(X^{(n-1)}, Y^{(n)})$ を観測した場合において (μ, θ) の最尤推定を考える．このとき EM アルゴリズムによる (μ, θ) の推定量の更新則は，適当な初期値 $(\mu_0, \theta_0) \in \mathbb{R}^2$ に対して

$$(\mu_{t+1}, \theta_{t+1}) = \underset{(\mu,\theta)\in\mathbb{R}^2}{\operatorname{argmax}}\ \mathbb{E}_{X_n \sim N(\bar{\mu}, \bar{\sigma}^2)}[\log p_{\mu,\theta}(X^{(n)}, Y^{(n)})], \qquad t = 0, 1, \ldots,$$

で与えられる．ここで，$\bar{\mu}, \bar{\sigma}^2$ は問い (2) で得られた $\mathbb{E}[X|Y]$, $\mathbb{V}[X|Y]$ の式に対して $(\mu,\ \theta,\ Y) := (\mu_t,\ \theta_t,\ Y_n)$ を代入したものであり，また $\mathbb{E}_{X_n \sim N(\bar{\mu}, \bar{\sigma}^2)}$ は X_n が $N(\bar{\mu}, \bar{\sigma}^2)$ に従い $(X^{(n-1)},\ Y^{(n)})$ が固定されている場合の期待値を表す．

(i) $\mathbb{E}_{X_n \sim N(\bar{\mu}, \bar{\sigma}^2)}[\log p_{\mu,\theta}(X^{(n)}, Y^{(n)})]$ を $n, \mu, \theta, \bar{\mu}, \bar{\sigma}^2, X^{(n-1)}, Y^{(n)}$ を用いて表せ．

(ii) $(\mu_{t+1}, \theta_{t+1})$ を $n, \bar{\mu}, \bar{\sigma}^2, X^{(n-1)}, Y^{(n)}$ を用いて表せ．

Problem 6

The probability density function of the normal distribution $N(\mu, \sigma^2)$ with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Let X and Z be random variables that independently follow $N(\mu, 1)$ and $N(0, 1)$, respectively, and define $Y = \theta X + Z$ for some constant $\theta \in \mathbb{R}$. For an integer $n > 1$, let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be two-dimensional random variables that independently follow the same distribution as (X, Y), for which we write $X^{(n)} = (X_1, X_2, \ldots, X_n)$ and $Y^{(n)} = (Y_1, Y_2, \ldots, Y_n)$.

Answer the following questions.

(1) Express the expectation $\mathbb{E}[Y]$ and variance $\mathbb{V}[Y]$ of Y using μ and θ.

(2) Show that the conditional distribution of X given Y is a normal distribution, and express its expectation $\mathbb{E}[X|Y]$ and variance $\mathbb{V}[X|Y]$ using μ, θ and Y.

(3) Let $(x^{(n)}, y^{(n)})$ denote a realization of $(X^{(n)}, Y^{(n)})$. Express the joint probability density function $p_{\mu,\theta}(x^{(n)}, y^{(n)})$ of $(X^{(n)}, Y^{(n)})$ using $\mu, \theta, x^{(n)} = (x_1, x_2, \ldots, x_n)$ and $y^{(n)} = (y_1, y_2, \ldots, y_n)$.

(4) Consider maximum-likelihood estimation of (μ, θ) by the EM algorithm for the case where the observation of X_n is missing from $(X^{(n)}, Y^{(n)})$, that is, the case where $(X^{(n-1)}, Y^{(n)})$ is observed. Then the update rule of estimators of (μ, θ) by the EM algorithm for some initial value $(\mu_0, \theta_0) \in \mathbb{R}^2$ is given by

$$(\mu_{t+1}, \theta_{t+1}) = \operatorname*{argmax}_{(\mu,\theta)\in\mathbb{R}^2} \mathbb{E}_{X_n \sim N(\bar{\mu}, \bar{\sigma}^2)}[\log p_{\mu,\theta}(X^{(n)}, Y^{(n)})], \qquad t = 0, 1, \ldots,$$

where $\bar{\mu}$ and $\bar{\sigma}^2$ are the values obtained by the substitution $(\mu,\ \theta,\ Y) := (\mu_t,\ \theta_t,\ Y_n)$ in the expressions of $\mathbb{E}[X|Y]$ and $\mathbb{V}[X|Y]$ obtained in question (2), respectively, and $\mathbb{E}_{X_n \sim N(\bar{\mu}, \bar{\sigma}^2)}$ denotes the expectation when X_n follows $N(\bar{\mu}, \bar{\sigma}^2)$ and $(X^{(n-1)},\ Y^{(n)})$ is fixed.

(i) Express $\mathbb{E}_{X_n \sim N(\bar{\mu}, \bar{\sigma}^2)}[\log p_{\mu,\theta}(X^{(n)}, Y^{(n)})]$ using $n, \mu, \theta, \bar{\mu}, \bar{\sigma}^2, X^{(n-1)}$ and $Y^{(n)}$.

(ii) Express $(\mu_{t+1}, \theta_{t+1})$ using $n, \bar{\mu}, \bar{\sigma}^2, X^{(n-1)}$ and $Y^{(n)}$.