平成 30 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

平成 29 年 8 月 21 日

15:45 – 17:15

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

15:45 – 17:15, August 21, 2017

Entrance Examination (AY 2018)

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

実数の独立変数 t に対する n 個の実数関数の組

$$x_i(t), \quad i = 1, 2, \cdots, n$$

についての，連立 1 階線形常微分方程式

$$\frac{dx_i}{dt}(t) = \sum_{j=1}^{n} a_{i,j}(t) x_j(t), \qquad t \geq 0 \tag{†}$$

を数値的に解くことを考える．ここで，$a_{i,j}(t)$ は $n \times n$ 係数行列 $A(t)$ の i, j 成分である．$h > 0$ を数値解法で用いる差分幅とせよ．数値解は初期値

$$x_i(0) \in \mathbb{R}, \quad i = 1, 2, \cdots, n$$

から計算する．そして任意の t について行列 $A(t)$ は n 個の相異なる固有値を持つとする．

以下の問いに答えよ．

(1) $A(t)$ が t に依存しない定数行列 A_0 であるとき，任意の初期値に対して，

$$\forall\, i = 1, 2, \cdots, n, \quad \lim_{t \to \infty} x_i(t) = 0$$

が満たされるための A_0 の固有値についての必要十分条件を答えよ．

(2) 前進差分による Euler 法によって式 (†) を解くための漸化式を答えよ．さらにその局所打ち切り誤差のオーダーを差分幅 h の O 記法で答えよ．

(3) 前進差分による Euler 法によって式 (†) を解くとする．得られた $x_i(kh)$ の近似を $X_i(k)$ と書く．ここで k はこの方法のステップ数を表す非負整数である．

$A(t)$ が t に依存しない定数行列 A_0 であるとき，任意の初期値に対して，

$$\forall\, i = 1, 2, \cdots, n, \quad \lim_{k \to \infty} X_i(k) = 0$$

が満たされるための A_0 の固有値についての必要十分条件を求めよ．

(4) 後退差分による Euler 法が前進差分による Euler 法よりも優れている点を，簡単な理由とともに説明せよ．

(5) 式 (†) を前進差分による Euler 法で解くための 1 ステップあたりの計算量を後退差分による Euler 法のそれと，n に関する O 記法を用いて比較せよ．ここで各 $a_{i,j}(t)$ を読む計算量は $O(1)$ とせよ．

Problem 1

Consider solving numerically a system of first-order linear ordinary differential equations

$$\frac{dx_i}{dt}(t) = \sum_{j=1}^{n} a_{i,j}(t) x_j(t), \qquad t \geq 0 \tag{†}$$

on a set of n real functions of a real independent variable t,

$$x_i(t), \quad i = 1, 2, \cdots, n,$$

where $a_{i,j}(t)$ is the i, j component of an $n \times n$ coefficient matrix $A(t)$. Let $h > 0$ be the step size used in the numerical methods. Suppose that the numerical solution is computed from initial values

$$x_i(0) \in \mathbb{R}, \quad i = 1, 2, \cdots, n.$$

Assume that matrix $A(t)$ has n distinct eigenvalues for any t.

Answer the following questions.

(1) Suppose that $A(t)$ is a constant matrix A_0, which is independent of t. Answer a necessary and sufficient condition for the eigenvalues of A_0 to satisfy

$$\forall\, i = 1, 2, \cdots, n, \quad \lim_{t \to \infty} x_i(t) = 0,$$

for any initial value.

(2) Answer a recurrence formula to solve Eq. (†) by the forward Euler method. In addition, express the order of the local truncation error by big O notation of the step size h.

(3) Suppose that Eq. (†) is solved by the forward Euler method. The obtained approximation of $x_i(kh)$ is expressed as $X_i(k)$, where k is a non-negative integer to express the step number of the method.

Suppose that $A(t)$ is a constant matrix A_0, which is independent of t. Answer a necessary and sufficient condition for the eigenvalues of A_0 to satisfy

$$\forall\, i = 1, 2, \cdots, n, \quad \lim_{k \to \infty} X_i(k) = 0,$$

for any initial value.

(4) Give an advantage of the backward Euler method compared to the forward Euler method with a brief explanation.

(5) Compare the computational complexity of one step to solve Eq. (†) by the forward Euler method with that by the backward Euler method, by using big O notation of n. Assume that the computational complexity to read each $a_{i,j}(t)$ is $O(1)$.

問題 2

V を変数記号の有限集合，F_0 を定数記号の有限集合，F_2 を二引数の関数記号の有限集合とし，V と F_0 と F_2 の記号から作られる項の集合を T とする．また，G は $x \to c$ または $x \to f(y,z)$ という形の書き換え規則の有限集合とする．ここで，$x, y, z \in V$, $c \in F_0$, $f \in F_2$ とし，各 $x \in V$ は G のちょうど一つの書き換え規則の左辺に現れるものとする．$\alpha, \beta \in T$ に対して関係 $\alpha \Rightarrow \beta$ を以下のように定義する:

> α の中の $x \in V$ の一つの出現に対して G の規則 $x \to \gamma$ (γ は c または $f(y,z)$) を適用して，その出現を γ で置き換えて β が得られる．

関係 $\Rightarrow^*$ を $\Rightarrow$ の反射推移閉包，関係 $\Leftrightarrow^*$ を $\Rightarrow$ の反射推移対称閉包とする．

以下の問いに答えよ．

(1) $\alpha, \beta \in T$ かつ $\alpha \Leftrightarrow^* \beta$ ならば $\alpha \Rightarrow^* \delta$ かつ $\beta \Rightarrow^* \delta$ を満たす $\delta \in T$ が存在することを示せ．

(2) $\alpha, \beta \in T$ に対して，$\alpha \Leftrightarrow^* \beta$ かどうかを判定するアルゴリズムを示し，その正しさを説明せよ．

次に，$\alpha \in T$ と $w \in \{L, R\}^*$ に対して，α の部分項を取り出す演算 $\alpha.w \in T$ を以下のように帰納的に定義する:

$$
\begin{aligned}
\alpha.\epsilon &= \alpha \\
\alpha.Lw &= \begin{cases} \alpha_1.w & \text{if } \alpha = f(\alpha_1, \alpha_2) \\ \text{undefined} & \text{otherwise} \end{cases} \\
\alpha.Rw &= \begin{cases} \alpha_2.w & \text{if } \alpha = f(\alpha_1, \alpha_2) \\ \text{undefined} & \text{otherwise} \end{cases}
\end{aligned}
$$

ここで ϵ は空列を表す．そして，$\alpha, \beta \in T$ に対して関係 $\alpha \approx \beta$ を以下のように定義する:

> 任意の $\alpha', \beta' \in T$ と $w \in \{L, R\}^*$ に対して，$\alpha \Leftrightarrow^* \alpha'$ かつ $\beta \Leftrightarrow^* \beta'$ かつ $\alpha'.w$ と $\beta'.w$ が定義されていてそれらの最初の記号が $F_0 \cup F_2$ に属しているならば，それらの記号は等しい．

以下の問いに答えよ．

(3) $\alpha, \beta \in T$ に対して，$\alpha \approx \beta$ かどうかを判定するアルゴリズムを示し，その正しさを説明せよ．

Problem 2

Let V be a finite set of symbols denoting variables, F_0 a finite set of constant symbols, F_2 a finite set of binary function symbols, and T the set of terms constructed from symbols in V, F_0 and F_2. Let G be a finite set of rewrite rules of the form $x \to c$ or of the form $x \to f(y, z)$, where $x, y, z \in V$, $c \in F_0$ and $f \in F_2$, and each $x \in V$ appears in the left-hand side of exactly one rewrite rule in G. For $\alpha, \beta \in T$, the relation $\alpha \Rightarrow \beta$ is defined as follows:

> β is obtained by applying a rule $x \to \gamma$ (γ is c or $f(y, z)$) in G to one occurrence of $x \in V$ in α and replacing the occurrence with γ.

The relation $\Rightarrow^*$ is defined as the reflexive and transitive closure of $\Rightarrow$, and the relation $\Leftrightarrow^*$ is defined as the reflexive, transitive and symmetric closure of $\Rightarrow$.

Answer the following questions.

(1) Show that if $\alpha, \beta \in T$ and $\alpha \Leftrightarrow^* \beta$, then there exists $\delta \in T$ such that $\alpha \Rightarrow^* \delta$ and $\beta \Rightarrow^* \delta$.

(2) Give an algorithm that judges $\alpha \Leftrightarrow^* \beta$ for $\alpha, \beta \in T$, and explain its correctness.

For $\alpha \in T$ and $w \in \{L, R\}^*$, the operation $\alpha.w \in T$ that extracts a subterm of α is inductively defined as follows:

$$\begin{aligned}
\alpha.\epsilon &= \alpha \\
\alpha.Lw &= \begin{cases} \alpha_1.w & \text{if } \alpha = f(\alpha_1, \alpha_2) \\ \text{undefined} & \text{otherwise} \end{cases} \\
\alpha.Rw &= \begin{cases} \alpha_2.w & \text{if } \alpha = f(\alpha_1, \alpha_2) \\ \text{undefined} & \text{otherwise} \end{cases}
\end{aligned}$$

Here ϵ is an empty word. For $\alpha, \beta \in T$, the relation $\alpha \approx \beta$ is then defined as follows:

> for any $\alpha', \beta' \in T$ and $w \in \{L, R\}^*$, if $\alpha \Leftrightarrow^* \alpha'$ and $\beta \Leftrightarrow^* \beta'$, and $\alpha'.w$ and $\beta'.w$ are both defined and their first symbols belong to $F_0 \cup F_2$, then those symbols are identical.

Answer the following question.

(3) Give an algorithm that judges $\alpha \approx \beta$ for $\alpha, \beta \in T$, and explain its correctness.

問題 3

木において節点 v の子孫である葉の集合を $L(v)$ とおき，$p(v, w)$ を節点 v から節点 w に至る単純パス上の枝数とする．このとき，葉ではない節点 v に対し，$\max_{w \in L(v)} p(v, w)$ を v の高さとよぶ．葉節点の高さは 0 であるものとする．また，根節点の高さをその木の高さとよぶものとする．

ここで，高さ $n \geq 0$ の二分木 T_n における各節点 v が以下のいずれかの性質を必ず持つものとする．

- v は葉である．
- v は子を 1 つのみ持ち，v の高さは 1 である．
- v は子を 2 つ持ち，v の 2 つの子の高さは 1 異なる．

$n \geq 0$ に対し，T_n の節点数を N_n とする．また，$r = (1 + \sqrt{5})/2$ とする．

以下の問いに答えよ．

(1) N_5 を求めよ．

(2) $n \geq 2$ に対し，N_n を，N_{n-1} と N_{n-2} を用いて表せ．

(3) すべての $n \geq 0$ に対し，$N_n \geq r^n$ であることを証明せよ．

(4) すべての $n \geq 0$ に対し，$N_n \leq r^{n+2}$ であることを証明せよ．

(5) 与えられた N_n 個の整数をそれぞれ T_n の相異なる節点に配置する問題を考える．このとき，T_n の各節点 v に配置した整数が v のいずれの子節点に配置した整数よりも小さくなることがないようにしたい．そのような配置を求める $O(r^n)$ のアルゴリズムを，そのアルゴリズムの計算時間が実際に $O(r^n)$ であることの証明とともに示せ．なお，入力される N_n 個の整数はソートされているとは限らないことに注意せよ．

Problem 3

Let $L(v)$ denote the set of leaves in the descendants of node v in a tree, and let $p(v, w)$ denote the number of edges of the simple path from node v to node w. For a non-leaf node v, $\max_{w \in L(v)} p(v, w)$ is called the height of v. Let the height of a leaf be 0. The height of the root of a tree is called the height of the tree.

Here, we have a binary tree T_n with height $n \geq 0$, in which each node v must have one of the following properties.

- v is a leaf.
- v has only one child, and the height of v is 1.
- v has two children, and the heights of the two children of v differ by 1.

Let N_n denote the number of nodes in T_n for $n \geq 0$. Let $r = (1 + \sqrt{5})/2$.

Answer the following questions.

(1) Calculate N_5.

(2) Express N_n in terms of N_{n-1} and N_{n-2} for $n \geq 2$.

(3) Prove that $N_n \geq r^n$ for every $n \geq 0$.

(4) Prove that $N_n \leq r^{n+2}$ for every $n \geq 0$.

(5) Consider the problem of assigning each of given N_n integers to a distinct node of T_n. The integer assigned to each node v must be no smaller than any of the integers assigned to v's children. Show an $O(r^n)$ algorithm that computes such an assignment, with a proof that the algorithm runs indeed in $O(r^n)$ time. Note that the N_n integers may not be sorted in the input.

問題 4

メモリアドレッシングが 32 ビットのマイクロプロセッサのキャッシュメモリに関する以下の問いに答えよ．

(1) 容量が 2^{15} バイト，ブロックサイズが 64 バイトのキャッシュメモリを考える．フルアソシアティブ方式，2 ウェイ・セットアソシアティブ方式，ダイレクトマップ方式を採用した場合のそれぞれに対して，タグ，インデックス，オフセットは各々何ビットになるか求めよ．

(2) 容量が 64 バイト，ブロックサイズが 8 バイトのキャッシュメモリを考える．フルアソシアティブ方式，2 ウェイ・セットアソシアティブ方式，ダイレクトマップ方式を採用した場合のそれぞれに対して，以下の 16 進数のメモリアドレスに対し，順番に 4 バイトの読み込みアクセスをしたときのキャッシュヒットの回数を求めよ．なお，キャッシュメモリの初期状態は空であり，キャッシュブロックの置換には LRU（Least Recently Used）アルゴリズムを用いることを仮定せよ．

0x20, 0x48, 0x40, 0x4C, 0x58, 0x80, 0xB8, 0xC8, 0x40, 0x44, 0x48, 0x4C, 0x50, 0x54, 0x58, 0x30, 0x28

Problem 4

Answer the following questions regarding cache memory of a microprocessor with 32-bit memory-addressing.

(1) Consider cache memory with a capacity of 2^{15} bytes and a block size of 64 bytes. The cache memory uses a full associative scheme, a two-way set-associative scheme, or a direct mapping scheme. For each of those schemes, obtain the bit length for each of a tag, an index, and an offset.

(2) Consider cache memory with a capacity of 64 bytes and a block size of 8 bytes. Obtain the number of cache hits, when the hexadecimal memory addresses below are accessed by 4-byte read operations in this order, in case that the cache memory uses a full associative scheme, a two-way set-associative scheme, and a direct mapping scheme, respectively. Assume that the cache memory is empty at the beginning, and the cache block is replaced based on the LRU (Least Recently Used) algorithm.

0x20, 0x48, 0x40, 0x4C, 0x58, 0x80, 0xB8, 0xC8, 0x40, 0x44, 0x48, 0x4C, 0x50, 0x54, 0x58, 0x30, 0x28

問題 5

$\mathbb{R}$ を実数の集合，$|w|$ を実数 w の絶対値とする．d 次元の実列ベクトル $\boldsymbol{w}$ の第 i 要素を w_i と書き，$\|\boldsymbol{w}\|_1 = |w_1| + |w_2| + \cdots + |w_d|$ および $\|\boldsymbol{w}\|_2 = \sqrt{w_1^2 + w_2^2 + \cdots + w_d^2}$ を定義する．また $\boldsymbol{w}^\top$ は $\boldsymbol{w}$ の転置を表す．

以下を満たすとき，ベクトル $\boldsymbol{g} \in \mathbb{R}^d$ を凸関数 f の $\boldsymbol{x} \in \mathbb{R}^d$ における劣勾配という．

$$\forall \boldsymbol{z} \in \mathbb{R}^d,\ f(\boldsymbol{z}) \geq f(\boldsymbol{x}) + \boldsymbol{g}^\top(\boldsymbol{z} - \boldsymbol{x})$$

点 $\boldsymbol{x}$ における凸関数 f の劣勾配の集合 $\{\boldsymbol{g} \in \mathbb{R}^d \mid \forall \boldsymbol{z} \in \mathbb{R}^d,\ f(\boldsymbol{z}) \geq f(\boldsymbol{x}) + \boldsymbol{g}^\top(\boldsymbol{z} - \boldsymbol{x})\}$ を，$\boldsymbol{x}$ における f の劣微分と呼び，$\partial f(\boldsymbol{x})$ と表現する．解答には事実 (i), (ii), (iii) を用いてもよい．(i) 微分可能な凸関数 $f(\boldsymbol{x})$ に対して

$$\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}, \qquad \nabla f(\boldsymbol{x}) = \begin{pmatrix} \partial f(\boldsymbol{x})/\partial x_1 \\ \vdots \\ \partial f(\boldsymbol{x})/\partial x_d \end{pmatrix}$$

が成り立つ．(ii) 凸関数 f_1, f_2 に対し $\partial(f_1 + f_2)(\boldsymbol{x}) = \{\boldsymbol{g}_1 + \boldsymbol{g}_2 \mid \boldsymbol{g}_1 \in \partial f_1(\boldsymbol{x}),\ \boldsymbol{g}_2 \in \partial f_2(\boldsymbol{x})\}$ が成り立つ．(iii) 凸関数 $f(\boldsymbol{w})$ が $\boldsymbol{w} = \boldsymbol{w}^*$ で最小となる必要十分条件は $\boldsymbol{0} \in \partial f(\boldsymbol{w}^*)$ である．

以下の問いに答えよ．

(1) (a) $f(w) = |w|$ $(w \in \mathbb{R})$ のとき，$\partial f(w)$ を求めよ．(b) $f(\boldsymbol{w}) = \|\boldsymbol{w}\|_1$ $(\boldsymbol{w} \in \mathbb{R}^d)$ のとき，$\partial f(\boldsymbol{w})$ を求めよ．

(2) $f(w) = \frac{1}{2}(w - z)^2 + \beta|w|$ $(w, z \in \mathbb{R},\ 0 < \beta \in \mathbb{R})$ のとき，$\partial f(w)$ を求め，$f(w)$ を最小にする $w^* \in \mathbb{R}$ を求めよ．

(3) $f(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w} - \boldsymbol{z}\|_2^2 + \beta\|\boldsymbol{w}\|_1$ $(\boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d, 0 < \beta \in \mathbb{R})$ のとき，$\partial f(\boldsymbol{w})$ を求めよ．また，$\boldsymbol{w} = \boldsymbol{w}^*$ で $f(\boldsymbol{w})$ が最小となるとし，j を $1 \leq j \leq d$ なる整数としたとき，$w_j^* = 0$ となる必要十分条件を求めよ．

d 次元の実ベクトル $\boldsymbol{x} \in \mathbb{R}^d$ から 1 次元の実数値ラベル $y \in \mathbb{R}$ を予測する問題を考える．n 個の訓練標本からなる集合

$$\{(\boldsymbol{x}_i, y_i) \mid \boldsymbol{x}_i \in \mathbb{R}^d,\ y_i \in \mathbb{R},\ i = 1, 2, \ldots, n\}$$

が与えられるとする．ここで，$(\boldsymbol{x}_i, y_i)$ は $\boldsymbol{x}_i$ の実数値ラベルが y_i であることを意味する．

d 次元パラメータ $\boldsymbol{w} \in \mathbb{R}^d$ を用いて損失関数を

$$L(\boldsymbol{w}) = \frac{1}{2n}\sum_{i=1}^{n}(y_i - \boldsymbol{w}^\top \boldsymbol{x}_i)^2$$

と定義する．また，正の実数 λ を用いて，予測器の学習を以下の最適化問題によって定式化する．

$$\boldsymbol{w}^* = \operatorname*{argmin}_{\boldsymbol{w} \in \mathbb{R}^d} \{L(\boldsymbol{w}) + \lambda\|\boldsymbol{w}\|_1\}$$

最適解 $\boldsymbol{w}^*$ を求めるアルゴリズムとして以下のものが知られている．すなわち，最適化問題 (†) を初期値 $\boldsymbol{w}^{(0)} \in \mathbb{R}^d$ から始めて，ステップサイズ $\eta_t > 0$ を用いて繰り返し解く．

$$\boldsymbol{w}^{(t+1)} = \operatorname*{argmin}_{\boldsymbol{w} \in \mathbb{R}^d} \left\{ \nabla L(\boldsymbol{w}^{(t)})^\top (\boldsymbol{w} - \boldsymbol{w}^{(t)}) + \lambda\|\boldsymbol{w}\|_1 + \frac{1}{2\eta_t}\|\boldsymbol{w} - \boldsymbol{w}^{(t)}\|_2^2 \right\} \qquad t = 0, 1, 2, \ldots \qquad (\dagger)$$

以下の問いに答えよ．

(4) j が $1 \leq j \leq d$ なる整数としたとき，$w_j^{(t+1)} = 0$ となる必要十分条件は $w_j^{(t)} - \eta_t \dfrac{\partial L}{\partial w_j}(\boldsymbol{w}^{(t)}) \in [-a, a]$ と表される．$a \in \mathbb{R}$ を η_t および λ を用いて表せ．

Problem 5

Denote the set of real numbers with $\mathbb{R}$ and the absolute value of a real value w with $|w|$. For a d-dimensional real column vector $\boldsymbol{w}$, we write its i-th element as w_i, and define $\|\boldsymbol{w}\|_1 = |w_1| + |w_2| + \cdots + |w_d|$ and $\|\boldsymbol{w}\|_2 = \sqrt{w_1^2 + w_2^2 + \cdots + w_d^2}$. The transpose of $\boldsymbol{w}$ is written as $\boldsymbol{w}^\top$.

A vector $\boldsymbol{g} \in \mathbb{R}^d$ is a *subgradient* of a convex function f at $\boldsymbol{x} \in \mathbb{R}^d$ if

$$\forall \boldsymbol{z} \in \mathbb{R}^d,\ f(\boldsymbol{z}) \geq f(\boldsymbol{x}) + \boldsymbol{g}^\top(\boldsymbol{z} - \boldsymbol{x})$$

holds. The set of subgradients of a convex function f at $\boldsymbol{x}$, $\{\boldsymbol{g} \in \mathbb{R}^d \mid \forall \boldsymbol{z} \in \mathbb{R}^d,\ f(\boldsymbol{z}) \geq f(\boldsymbol{x}) + \boldsymbol{g}^\top(\boldsymbol{z} - \boldsymbol{x})\}$, is called the *subdifferential* of f at $\boldsymbol{x}$, and is denoted by $\partial f(\boldsymbol{x})$. You may use the following facts (i), (ii) and (iii). (i) A differentiable convex function $f(\boldsymbol{x})$ satisfies

$$\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}, \qquad \nabla f(\boldsymbol{x}) = \begin{pmatrix} \partial f(\boldsymbol{x})/\partial x_1 \\ \vdots \\ \partial f(\boldsymbol{x})/\partial x_d \end{pmatrix}.$$

(ii) For convex functions f_1 and f_2, it holds that $\partial(f_1 + f_2)(\boldsymbol{x}) = \{\boldsymbol{g}_1 + \boldsymbol{g}_2 \mid \boldsymbol{g}_1 \in \partial f_1(\boldsymbol{x}),\ \boldsymbol{g}_2 \in \partial f_2(\boldsymbol{x})\}$. (iii) $\boldsymbol{0} \in \partial f(\boldsymbol{w}^*)$ is a necessary and sufficient condition that a convex function $f(\boldsymbol{w})$ is minimized by $\boldsymbol{w} = \boldsymbol{w}^*$.

Answer the following questions.

(1) (a) For $f(w) = |w|$ $(w \in \mathbb{R})$, obtain $\partial f(w)$. (b) For $f(\boldsymbol{w}) = \|\boldsymbol{w}\|_1$ $(\boldsymbol{w} \in \mathbb{R}^d)$, obtain $\partial f(\boldsymbol{w})$.

(2) For $f(w) = \frac{1}{2}(w - z)^2 + \beta|w|$ $(w, z \in \mathbb{R},\ 0 < \beta \in \mathbb{R})$, obtain $\partial f(w)$. Also obtain $w^* \in \mathbb{R}$ that minimizes $f(w)$.

(3) For $f(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w} - \boldsymbol{z}\|_2^2 + \beta\|\boldsymbol{w}\|_1$ $(\boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d,\ 0 < \beta \in \mathbb{R})$, obtain $\partial f(\boldsymbol{w})$. Also, assuming that $\boldsymbol{w} = \boldsymbol{w}^* \in \mathbb{R}^d$ minimizes $f(\boldsymbol{w})$, and letting j be an integer satisfying $1 \leq j \leq d$, obtain a necessary and sufficient condition for $w_j^* = 0$.

Consider the problem of predicting one dimensional real-valued label $y \in \mathbb{R}$ from a d-dimensional real vector $\boldsymbol{x} \in \mathbb{R}^d$. Suppose that a set of n training samples

$$\{(\boldsymbol{x}_i, y_i) \mid \boldsymbol{x}_i \in \mathbb{R}^d,\ y_i \in \mathbb{R},\ i = 1, 2, \ldots, n\}$$

is given where $(\boldsymbol{x}_i, y_i)$ means that y_i is the real-valued label of $\boldsymbol{x}_i$.

By using a d-dimensional parameter $\boldsymbol{w} \in \mathbb{R}^d$, define a loss function as

$$L(\boldsymbol{w}) = \frac{1}{2n}\sum_{i=1}^{n}(y_i - \boldsymbol{w}^\top \boldsymbol{x}_i)^2.$$

We formulate the training of a predictor as the following optimization problem with a positive real value λ:

$$\boldsymbol{w}^* = \operatorname*{argmin}_{\boldsymbol{w} \in \mathbb{R}^d} \{L(\boldsymbol{w}) + \lambda\|\boldsymbol{w}\|_1\}.$$

The following algorithm is known for obtaining the optimal solution $\boldsymbol{w}^*$. It iteratively solves the optimization problem (†) from an initial value $\boldsymbol{w}^{(0)} \in \mathbb{R}^d$ and using the step size $\eta_t > 0$:

$$\boldsymbol{w}^{(t+1)} = \operatorname*{argmin}_{\boldsymbol{w} \in \mathbb{R}^d} \left\{\nabla L(\boldsymbol{w}^{(t)})^\top(\boldsymbol{w} - \boldsymbol{w}^{(t)}) + \lambda\|\boldsymbol{w}\|_1 + \frac{1}{2\eta_t}\|\boldsymbol{w} - \boldsymbol{w}^{(t)}\|_2^2\right\} \qquad t = 0, 1, 2, \ldots. \qquad (\dagger)$$

Answer the following question.

(4) Express $a \in \mathbb{R}$ using η_t and λ such that $w_j^{(t)} - \eta_t \frac{\partial L}{\partial w_j}(\boldsymbol{w}^{(t)}) \in [-a, a]$ is a necessary and sufficient condition for $w_j^{(t+1)} = 0$, where j is an integer satisfying $1 \leq j \leq d$.

問題 6

あるRNA分子の分解時間を考える．分解にかかる時間 T の確率密度関数が

$$f_T(t) = \lambda e^{-\lambda t}, \qquad t \geq 0$$

であるとする．ただし λ は正の実数定数である．

以下の問いに答えよ．

(1) T の累積分布関数

$$F_T(t) = \int_0^t f_T(x)dx$$

を求めよ．また T の中央値を求めよ．

(2) n 本の RNA 分子の分解時間 T_i $(i = 1, 2, \ldots, n)$ を計測した．各 RNA 分子の分解時間は互いに独立に同一の確率密度関数 $f_T(t)$ に従って分布するとする．このとき

$$\mu_T = \frac{\sum_{i=1}^{n} T_i}{n}$$

の期待値および分散を求めよ．

(3) 問い (2) の観測時間 T_i の最大値を $T_{\max} = \max\{T_1, \ldots, T_n\}$ とし，$T_{\max} > t$ である確率を $\mathrm{Prob}(T_{\max} > t)$ と表す．$\mathrm{Prob}(T_{\max} > t)$ を $F_T(t)$ を用いて表せ．

(4) $T_{\max}$ の確率密度関数 $f_{T_{\max}}(t)$ および $T_{\max}$ の期待値を求めよ．

Problem 6

Consider the decomposition time of an RNA molecule. Assume that the probability density function of the decomposition time T is

$$f_T(t) = \lambda e^{-\lambda t}, \qquad t \geq 0,$$

where λ is a positive real constant.

Answer the following questions.

(1) Calculate the cumulative distribution function

$$F_T(t) = \int_0^t f_T(x)dx.$$

Also compute the median of T.

(2) We measured the decomposition times T_i $(i = 1, \dots, n)$ of n RNA molecules. Assume that the decomposition time of each RNA molecule follows the probability density function $f_T(t)$ independently and identically. Calculate the expected value and the variance of

$$\mu_T = \frac{\sum_{i=1}^{n} T_i}{n}.$$

(3) Let $T_{\max} = \max\{T_1, \dots, T_n\}$, which is the maximum of the measured times T_i in question (2). Let $\mathrm{Prob}(T_{\max} > t)$ denote the probability that $T_{\max} > t$. Give an expression for $\mathrm{Prob}(T_{\max} > t)$ in terms of $F_T(t)$.

(4) Calculate the probability density function $f_{T_{\max}}(t)$ of $T_{\max}$, and the expected value of $T_{\max}$.