問題 1

無向グラフにおいて，同じ点を結ぶ枝を自己閉路といい，同じ両端点を結ぶ複数の枝を多重枝という．以下では，自己閉路はもたないが，多重枝は存在する場合もある無向グラフを考える．グラフ G が A グラフであるとは，次の 2 つの操作を繰り返し適用することで 1 本の枝からなるグラフが G から得られることをいう．

B 操作 ある 2 点を結ぶ 2 本の多重枝があるとき，その多重枝を両端点を結ぶ 1 本の枝に置き換える

C 操作 1 つの枝が点 u と v を結び，他の枝が点 v と w を結び，$u \neq w$ であり，点 v に接続する枝はそれ以外にないとき，点 v を削除して 2 本の枝を点 u と w を結ぶ新たな枝に置き換える

以下の問いに答えよ．

(1) n 点からなる完全グラフを K_n で表す．K_3, K_4 はそれぞれ A グラフであるかないかを答えよ．

(2) すべての A グラフは平面グラフであることを示せ．

(3) 点の数が n で多重枝を持たない A グラフの枝数の最大値を，証明とともに与えよ．さらに，一般の n について，そのような最大値を達成する A グラフを説明とともに与えよ．

(4) n 点, m 枝の無向グラフを入力として，それが A グラフかどうかを判定する $O(m+n)$ 時間のアルゴリズムを与えよ．B 操作および C 操作を実現するためにアルゴリズム中で用いるグラフデータ構造も説明すること．

Problem 1

In undirected graphs, a self-loop is an edge connecting the same vertex, and multi-edges are multiple edges connecting the same pair of vertices. From now on, we consider undirected graphs without self-loops and possibly with multi-edges. We say that a graph G is an A-graph if a graph consisting of a single edge can be obtained from G by repeatedly applying the following two operations.

B-operation When two multi-edges connect a pair of vertices, replace the multi-edges with a single edge connecting the pair of vertices.

C-operation When one edge connects vertices u and v, another edge connects v and w (where $u \neq w$), and there is no other edge incident to v, remove the vertex v and replace the two edges with a new edge connecting u and w.

Answer the following questions.

(1) Let K_n be a complete graph of n vertices. Answer whether each of K_3 and K_4 is an A-graph or not.

(2) Show that every A-graph is planar.

(3) Give the maximum number of edges of an A-graph with n vertices without multi-edges, with a proof. Also, give such an A-graph attaining the maximum for general n, with an explanation.

(4) Give an $O(m + n)$-time algorithm which, given an undirected graph with n vertices and m edges as an input, determines whether it is an A-graph or not. Explain also the graph data structures used in the algorithm for realizing B-operations and C-operations.

問題 2

Σ を文字の集合 $\{a,b\}$ とする．Σ 上の語 $w \in \Sigma^*$ および 2 つの言語 $L_a, L_b \subseteq \Sigma^*$ について，言語 $w\{a \mapsto L_a, b \mapsto L_b\} \subseteq \Sigma^*$ を以下のように w に関する帰納法によって定義する．

$$
\begin{aligned}
&\epsilon\{a \mapsto L_a, b \mapsto L_b\} = \{\epsilon\} \\
&(aw)\{a \mapsto L_a, b \mapsto L_b\} = \{w_1w_2 \mid w_1 \in L_a, w_2 \in w\{a \mapsto L_a, b \mapsto L_b\}\} \\
&(bw)\{a \mapsto L_a, b \mapsto L_b\} = \{w_1w_2 \mid w_1 \in L_b, w_2 \in w\{a \mapsto L_a, b \mapsto L_b\}\}.
\end{aligned}
$$

ただし ϵ は空列を表す．例えば，$w = aba$, $L_a = \{b^n \mid n \geq 0\}$, $L_b = \{a^n \mid n \geq 0\}$ ならば $w\{a \mapsto L_a, b \mapsto L_b\} = \{b^\ell a^m b^n \mid \ell, m, n \geq 0\}$ である．さらに，言語 $L, L_a, L_b \subseteq \Sigma^*$ に対し，$L\{a \mapsto L_a, b \mapsto L_b\}$ を $\bigcup_{w \in L} w\{a \mapsto L_a, b \mapsto L_b\}$ と定義する．例えば，$L = \{a^n b \mid n \geq 0\}$, $L_a = \{ab\}$, $L_b = \{a^n \mid n \geq 0\}$ ならば $L\{a \mapsto L_a, b \mapsto L_b\} = \{(ab)^m a^n \mid m, n \geq 0\}$ である．

以下の問いに答えよ．

(1) $L = \{(ab)^m a^n \mid m, n \geq 0\}$, $L_a = \{bb\}$, $L_b = \{ab, a\}$ とする．$L\{a \mapsto L_a, b \mapsto L_b\}$ を正規表現を用いて表せ．

(2) $L' = \{a^m b^n \mid m \geq n \geq 0\}$, $L'_a = \{a^n \mid n \geq 0\}$, $L'_b = \{a^m b^m \mid m \geq 0\}$ とする．$\{w \in \Sigma^* \mid w\{a \mapsto L'_a, b \mapsto L'_b\} \subseteq L'\}$ を正規表現を用いて表せ．

(3) $\mathcal{A}_0 = (Q_0, \Sigma, \delta_0, q_{0,0}, F_0)$, $\mathcal{A}_1 = (Q_1, \Sigma, \delta_1, q_{1,0}, F_1)$, $\mathcal{A}_2 = (Q_2, \Sigma, \delta_2, q_{2,0}, F_2)$ は決定性有限オートマトンであり，各 $i \in \{0,1,2\}$ について，$\mathcal{A}_i$ が受理する言語を L_i とする．ただし $Q_i, \delta_i, q_{i,0}, F_i$ は $\mathcal{A}_i$ $(i \in \{0,1,2\})$ の状態集合，遷移関数，初期状態，受理状態の集合をそれぞれ表すものとし，遷移関数 $\delta_i \in Q_i \times \Sigma \to Q_i$ $(i \in \{0,1,2\})$ は全域関数であるとする．$L_0\{a \mapsto L_1, b \mapsto L_2\}$ を受理する非決定性有限オートマトンを，簡単な説明とともに与えよ．ϵ 遷移を用いてもよい．

(4) 問い (3) の $\mathcal{A}_i, L_i$ $(i \in \{0,1,2\})$ について，$\{w \in \Sigma^* \mid w\{a \mapsto L_1, b \mapsto L_2\} \subseteq L_0\}$ を受理する決定性有限オートマトンを，簡単な説明とともに与えよ．

Problem 2

Let Σ be the set $\{a, b\}$ of letters. For a word $w \in \Sigma^*$ and two languages $L_a, L_b \subseteq \Sigma^*$ over Σ, we define the language $w\{a \mapsto L_a, b \mapsto L_b\} \subseteq \Sigma^*$ as follows, by induction on w.

$$\begin{aligned}
\epsilon\{a \mapsto L_a, b \mapsto L_b\} &= \{\epsilon\} \\
(aw)\{a \mapsto L_a, b \mapsto L_b\} &= \{w_1 w_2 \mid w_1 \in L_a, w_2 \in w\{a \mapsto L_a, b \mapsto L_b\}\} \\
(bw)\{a \mapsto L_a, b \mapsto L_b\} &= \{w_1 w_2 \mid w_1 \in L_b, w_2 \in w\{a \mapsto L_a, b \mapsto L_b\}\}.
\end{aligned}$$

Here, ϵ represents the empty word. For example, if $w = aba$, $L_a = \{b^n \mid n \geq 0\}$, and $L_b = \{a^n \mid n \geq 0\}$, then $w\{a \mapsto L_a, b \mapsto L_b\} = \{b^\ell a^m b^n \mid \ell, m, n \geq 0\}$. Furthermore, for languages $L, L_a, L_b \subseteq \Sigma^*$, we define $L\{a \mapsto L_a, b \mapsto L_b\}$ as $\bigcup_{w \in L} w\{a \mapsto L_a, b \mapsto L_b\}$. For example, if $L = \{a^n b \mid n \geq 0\}$, $L_a = \{ab\}$, and $L_b = \{a^n \mid n \geq 0\}$, then $L\{a \mapsto L_a, b \mapsto L_b\} = \{(ab)^m a^n \mid m, n \geq 0\}$.

Answer the following questions.

(1) Let $L = \{(ab)^m a^n \mid m, n \geq 0\}$, $L_a = \{bb\}$, and $L_b = \{ab, a\}$. Express $L\{a \mapsto L_a, b \mapsto L_b\}$ using a regular expression.

(2) Let $L' = \{a^m b^n \mid m \geq n \geq 0\}$, $L'_a = \{a^n \mid n \geq 0\}$, and $L'_b = \{a^m b^m \mid m \geq 0\}$. Express $\{w \in \Sigma^* \mid w\{a \mapsto L'_a, b \mapsto L'_b\} \subseteq L'\}$ using a regular expression.

(3) Let $\mathcal{A}_0 = (Q_0, \Sigma, \delta_0, q_{0,0}, F_0)$, $\mathcal{A}_1 = (Q_1, \Sigma, \delta_1, q_{1,0}, F_1)$, and $\mathcal{A}_2 = (Q_2, \Sigma, \delta_2, q_{2,0}, F_2)$ be deterministic finite automata, and for each $i \in \{0, 1, 2\}$, let L_i be the language accepted by $\mathcal{A}_i$. Here, $Q_i, \delta_i, q_{i,0}$, and F_i are the set of states, the transition function, the initial state, and the set of final states of $\mathcal{A}_i$ ($i \in \{0, 1, 2\}$), respectively. Assume that the transition functions $\delta_i \in Q_i \times \Sigma \to Q_i$ ($i \in \{0, 1, 2\}$) are total. Give a non-deterministic finite automaton that accepts $L_0\{a \mapsto L_1, b \mapsto L_2\}$, with a brief explanation. You may use ϵ-transitions.

(4) For $\mathcal{A}_i$ and L_i ($i \in \{0, 1, 2\}$) in question (3), give a deterministic finite automaton that accepts $\{w \in \Sigma^* \mid w\{a \mapsto L_1, b \mapsto L_2\} \subseteq L_0\}$, with a brief explanation.

問題 3

ノイズがある通信路を介して，5ビットを1ビットずつ送受信するシリアル通信回路について考える．その5ビットは，2ビットのデータ，スタート信号2ビット，奇数パリティ信号1ビットで構成される．

送信回路は，初期状態では '0' を出力し続け，送信開始時には2ビットのスタート信号 '11' を1ビットずつ出力する．その次に，2ビットのペイロードデータを最上位ビットから順に1ビットずつ出力する．最後に，スタート信号2ビット，データ信号2ビット，およびパリティ信号を合わせた，送信ビット列中の '1' の個数が奇数になるようなパリティ信号1ビットを出力する．データ送信完了後は初期状態に戻り，次のデータを送信するまで '0' を出力し続ける．

受信回路は，送信回路からの1ビット入力A，パリティチェック結果の1ビット出力B，受信ペイロードデータの2ビット出力を持つ．受信回路は，受信ビット列からペイロードデータを取得し，受信ビット列の奇数パリティチェックを行う．

受信回路は，スタート信号の1ビット目の '1' を受信するまで初期状態で待機する．スタート信号の1ビット目を受信した次のクロックサイクルでは，スタート信号の2ビット目に対応する値を受信する．もし，スタート信号の2ビット目に対応する値が '0' の場合，1ビット目に受信した '1' の値は，ノイズにより誤って発生したものであると判断して，受信回路は初期状態に戻る．そうでなければ，その次のクロックサイクルから，入力Aの値を受信ペイロードデータとして2クロックサイクル連続で保存する．その次のクロックサイクルで，パリティ信号を受け取り，スタート信号2ビット，ペイロードデータ2ビット，パリティ信号を合わせた5ビットの受信ビット列中の '1' の個数が奇数かどうかを確認する．もし '1' の個数が奇数ならば出力Bの値を '1' とし，奇数でなければ '0' とする．また，パリティ信号を受信したクロックサイクル以外は出力Bを '0' とする．そして，パリティチェックの結果に関わらず，受信回路は初期状態に戻る．

以下の問いに答えよ．

(1) 受信回路における，入力Aと出力Bに関するパリティチェック回路の振る舞いを表す，7つの状態で構成される Mealy 型の有限状態機械 (FSM) の状態遷移図を示せ．また，その状態遷移図に対応するワンホットエンコーディングに基づく状態遷移表と出力表を構成せよ．ワンホットエンコーディングとは，ビット列中の1ビットのみが '1' であり，他のビットはすべて '0' であるビット列のみを用いる符号化方式である．

(2) 問い (1) で求めた状態遷移表および出力表に基づき，出力Bを，入力AとFSMの現在の状態のワンホットエンコーディング表現の各ビットを用いたブール論理式として表せ．また，求めたブール論理式から，入力AとFSMの現在の状態のワンホットエンコーディング表現の各ビットを入力とし，Bを出力するパリティチェック回路のゲートレベル回路を，2入力 AND ゲート，2入力 OR ゲート，NOT ゲートのみを用いて構成せよ．ただし，使用する論理ゲートの個数に制限はない．未使用の入力信号は記載しなくて良い．

(3) 問い (2) で求めたブール論理式から，入力AとFSMの現在の状態のワンホットエンコーディング表現の各ビットを入力とし，Bを出力するパリティチェック回路の CMOS トランジスタレベル回路図を示せ．ただし，回路のトランジスタ数は12個以下とすること．インバータ記号を用いても良いが，そのインバータを構成するトランジスタ数も総数に含めること．未使用の入力信号は記載しなくて良い．

Problem 3

Consider bit-serial communication circuits which send and receive 5-bit information bit-by-bit in a noisy environment. The 5-bit information consists of a 2-bit start-bit signal, 2-bit payload data, and a 1-bit odd-parity signal.

The sender circuit always outputs '0' in the initial state. At the beginning of a communication, the sender outputs 2-bit data '11' bit-by-bit. It subsequently outputs 2-bit payload data bit-by-bit from the most significant bit. It finally outputs an odd-parity signal such that the number of '1's in the sent bit sequence including the 2-bit start-bit signal, the payload data, and the parity signal itself is an odd number. After sending the parity signal, the sender circuit goes to the initial state, and it outputs '0' until the next sending.

The receiver circuit has a 1-bit input A from the sender circuit, a 1-bit output B for the parity check result, and a 2-bit output for the received payload data. It obtains payload data from a received bit sequence and does the odd parity checking.

In the initial state, the receiver circuit waits for '1' corresponding to the first bit of a start-bit signal. In the next clock cycle after receiving the first bit of a start-bit signal, it receives a value corresponding to the second bit of a start-bit signal. If the received value corresponding to the second bit of a start-bit signal is '0', the receiver circuit judges that the first received bit '1' was an error caused by a noise, and goes back to the initial state. Otherwise, in the next 2 clock cycles, it stores each value of the input A as payload data. At the next clock cycle, it receives a parity-bit, and it verifies that the number of '1's in the received 5-bit sequence consisting of the 2-bit start-bit signal, the 2-bit payload data, and the parity-bit is odd. It assigns '1' to the output B if the number of '1's is odd, and it assigns '0' otherwise. The value of the output B is always '0', except in the clock cycles for receiving a parity-bit. The receiver circuit then goes to the initial state, regardless of the parity-check result.

Answer the following questions.

(1) Give the state transition diagram of a Mealy-type finite state machine (FSM), consisting of 7 states, for the parity check circuit with the input A and the output B in the receiver circuit. Based on the state transition diagram, give also a corresponding state transition table and an output table by using the one-hot encoding. One-hot encoding is a method for encoding each state as a bit sequence where only one bit is '1' and the other bits are '0'.

(2) Based on the state transition table and the output table in question (1), express the output B as a Boolean expression in terms of the input A and the one-hot encoding representation of the current state of the FSM. Based on the Boolean expression, give also a corresponding gate-level circuit of the parity check circuit that outputs B, given A and the one-hot encoding representation of the current state of the FSM as inputs. You are allowed to use only 2-input AND gates, 2-input OR-gates, and NOT-gates. There is no limitation on the number of gates. You need not describe unused input signals.

(3) According to the Boolean expression answered in question (2), give a CMOS transistor level circuit that outputs B, given A and the one-hot encoding representation of the current state of the FSM as inputs. You are not allowed to use more than 12 transistors. You may use the inverter mark, but the number of transistors required for the inverters must be included in the total number of transistors. You need not describe unused input signals.

問題 4

$\mathbb{R}$ を実数の集合とする．ベクトル及び行列の転置を $\top$ で表す．$\boldsymbol{w} = (w_1, w_2, \ldots, w_d)^\top \in \mathbb{R}^d$ が d 次元の列ベクトルであるとき，そのノルム $\|\boldsymbol{w}\|_2$ を $\|\boldsymbol{w}\|_2 = \sqrt{w_1^2 + w_2^2 + \ldots + w_d^2}$ と定義する．2 つの列ベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^d$ の内積を $\boldsymbol{x}_1^\top \boldsymbol{x}_2 \in \mathbb{R}$ と定義する．$d \times d$ 行列 $A \in \mathbb{R}^{d \times d}$ に対して，$\|\boldsymbol{w}\|_A = \sqrt{\boldsymbol{w}^\top A \boldsymbol{w}}$ と定義する．$\mathrm{tr}(B)$ を行列 B のトレースとする．

d 次元の実ベクトル $\boldsymbol{x} \in \mathbb{R}^d$ から実数値ラベル $y \in \mathbb{R}$ を予測する問題を考える．予測器の学習のために，n 個の訓練標本

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

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．また，d 次元実ベクトル $\boldsymbol{w}^* \in \mathbb{R}^d$ および独立同分布にしたがう観測ノイズ ϵ_i $(i = 1, 2, \ldots, n)$ により訓練標本の生成過程を

$$y_i = \boldsymbol{w}^{*\top} \boldsymbol{x}_i + \epsilon_i \ (i = 1, 2, \ldots, n)$$

と仮定する．ここで，期待値 $\mathbb{E}[\epsilon_i] = 0$，分散 $\mathbb{V}[\epsilon_i] = \sigma^2 > 0$ $(i = 1, \ldots, n)$ とする．

$$X = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top \in \mathbb{R}^{n \times d},\ Y = [y_1, \ldots, y_n]^\top \in \mathbb{R}^n,\ \boldsymbol{\epsilon} = [\epsilon_1, \ldots, \epsilon_n]^\top \in \mathbb{R}^n$$

という記号を導入する．また，$\Phi = \frac{1}{n} X^\top X \in \mathbb{R}^{d \times d}$ とし，Φ は正則行列であると仮定する．観測ノイズに関する期待値計算を $\mathbb{E}_{\boldsymbol{\epsilon}}[\cdot]$ で表現する．

予測器 $f(\boldsymbol{x}) = \hat{\boldsymbol{w}}^\top \boldsymbol{x}$ の学習を以下の最適化問題によって定式化する．

$$\hat{\boldsymbol{w}} = \operatorname*{argmin}_{\boldsymbol{w} \in \mathbb{R}^d} L(\boldsymbol{w})$$

$$L(\boldsymbol{w}) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - \boldsymbol{w}^\top \boldsymbol{x}_i)^2 = \frac{1}{2n} \|Y - X\boldsymbol{w}\|_2^2.$$

以下の問いに答えよ．なお，答えだけでなく導出の過程も記述すること．

(1) $\hat{\boldsymbol{w}}$ を X, Y, Φ, n を用いて表せ．

(2) $\mathbb{E}_{\boldsymbol{\epsilon}}[L(\boldsymbol{w})]$ を $\frac{1}{2}\|\boldsymbol{w} - \boldsymbol{w}^*\|_A^2 + b$ の形で表現したい．Φ, σ^2 を用いて行列 $A \in \mathbb{R}^{d \times d}$ および正の実数 $b > 0$ を表せ．

(3) $\mathbb{E}_{\boldsymbol{\epsilon}}[L(\hat{\boldsymbol{w}})] - \mathbb{E}_{\boldsymbol{\epsilon}}[L(\boldsymbol{w}^*)]$ を $\frac{\sigma^2}{2n} \mathrm{tr}(B)$ の形で表現したい．行列 X を用いて行列 $B \in \mathbb{R}^{d \times d}$ を表せ．

(4) Φ が正則行列とならないとき，どのような問題が起こるか説明し，その対策案について答えよ．

Problem 4

Let $\mathbb{R}$ be the set of real numbers. Denote by $\top$ the transposition operator of a vector and a matrix. When $\boldsymbol{w} = (w_1, w_2, \ldots, w_d)^\top \in \mathbb{R}^d$ is a d-dimensional column vector, the norm $\|\boldsymbol{w}\|_2$ is defined by $\|\boldsymbol{w}\|_2 = \sqrt{w_1^2 + w_2^2 + \ldots + w_d^2}$. Define the inner product of two column vectors $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^d$ as $\boldsymbol{x}_1^\top \boldsymbol{x}_2 \in \mathbb{R}$. For a $d \times d$ matrix $A \in \mathbb{R}^{d\times d}$, define $\|\boldsymbol{w}\|_A = \sqrt{\boldsymbol{w}^\top A \boldsymbol{w}}$. Let $\mathrm{tr}(B)$ be the trace of the matrix B.

Consider the problem of predicting a real-valued label $y \in \mathbb{R}$ from a d-dimensional real vector $\boldsymbol{x} \in \mathbb{R}^d$. For learning a predictor, suppose that n training samples

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

are given where $(\boldsymbol{x}_i, y_i)$ means that y_i is the real-valued label of $\boldsymbol{x}_i$. In addition, by using a d-dimensional vector $\boldsymbol{w}^* \in \mathbb{R}^d$ and observational noise ϵ_i $(i = 1, 2, \ldots, n)$ that is independent and identically distributed, assume the data generation process as

$$y_i = \boldsymbol{w}^{*\top} \boldsymbol{x}_i + \epsilon_i \ (i = 1, 2, \ldots, n),$$

where the expectation $\mathbb{E}[\epsilon_i] = 0$ and variance $\mathbb{V}[\epsilon_i] = \sigma^2 > 0$ $(i = 1, \ldots, n)$. Let us introduce the symbols

$$X = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]^\top \in \mathbb{R}^{n\times d},\ Y = [y_1, \ldots, y_n]^\top \in \mathbb{R}^n,\ \boldsymbol{\epsilon} = [\epsilon_1, \ldots, \epsilon_n]^\top \in \mathbb{R}^n.$$

We also use the symbol $\Phi = \frac{1}{n} X^\top X \in \mathbb{R}^{d\times d}$ where Φ is assumed to be a regular matrix. The expectation over the observational noises is expressed by $\mathbb{E}_{\boldsymbol{\epsilon}}[\cdot]$.

We formulate the learning of a predictor $f(\boldsymbol{x}) = \hat{\boldsymbol{w}}^\top \boldsymbol{x}$ as the following optimization problem.

$$\hat{\boldsymbol{w}} = \operatorname*{argmin}_{\boldsymbol{w}\in\mathbb{R}^d} L(\boldsymbol{w})$$
$$L(\boldsymbol{w}) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - \boldsymbol{w}^\top \boldsymbol{x}_i)^2 = \frac{1}{2n} \|Y - X\boldsymbol{w}\|_2^2.$$

Answer the following questions. Describe not only an answer but also the derivation process.

(1) Express $\hat{\boldsymbol{w}}$ using X, Y, Φ, and n.

(2) Suppose we wish to express $\mathbb{E}_{\boldsymbol{\epsilon}}[L(\boldsymbol{w})]$ in the form of $\frac{1}{2}\|\boldsymbol{w} - \boldsymbol{w}^*\|_A^2 + b$. Express the matrix $A \in \mathbb{R}^{d\times d}$ and the positive real number $b > 0$ using Φ and σ^2.

(3) Suppose we wish to express $\mathbb{E}_{\boldsymbol{\epsilon}}[L(\hat{\boldsymbol{w}})] - \mathbb{E}_{\boldsymbol{\epsilon}}[L(\boldsymbol{w}^*)]$ in the form of $\frac{\sigma^2}{2n}\,\mathrm{tr}(B)$. Express the matrix $B \in \mathbb{R}^{d\times d}$ using the matrix X.

(4) Explain what problem arises when Φ is not a regular matrix and suggest a way to remedy the problem.