1　課題 6(重回帰問題)

各入力 $x_i \in \mathbb{R}^2,\ i = 1, 2, \cdots, 10000$ に対し、次の式で観測データ $y_i \in \mathbb{R}$ が生成されているとする。

$$y_i = x_i^T \theta + w_i \tag{1.1}$$

ここで観測誤差は $\mathcal{N}(0, 1)$ に従って発生している（ただし、平均の情報のみ知っているとし、分散も分布も知らないものとする)。このとき、次の問に答えよ。

1. θ の最小二乗誤差推定量を、与えられた全てのデータ $\{(x_i, y_i)\}_{i=1}^{10000}$ を使って求めよ。また、そのときの推定誤差共分散行列を求めよ。
2. 用いるデータ $\{(x_i, y_i)\}_{i=1}^{N}$ を増やしたとき、$\hat{\theta}_N$ の各要素が収束していくことを片対数グラフでプロットして確認せよ。ただし、$N = 2, 4, 8, \cdots, 2^{13} = 8192$ のときのみでよい。
3. 全てのデータを用いたときの決定変数を求めよ。

2　課題 7(多項式回帰)

各入力 $x_i \in \mathbb{R},\ i = 1, 2, \cdots, 10000$ に対し、次の式で観測データ $y_i \in \mathbb{R}$ が生成されている。

$$y_i = \varphi(x_i)\theta + w_i, \quad \varphi(x) := \begin{bmatrix} 1 & x & x^2 & x^3 \end{bmatrix} \tag{2.1}$$

ここで観測誤差は $\mathcal{N}(0, 9)$ に従って発生している（ただし、平均の情報のみ知っているとし、分散も分布も知らないものとする)。このとき、次の問いに答えよ。

1. θ の最小二乗誤差推定量を、与えられた全てのデータ $\{(x_i, y_i)\}_{i=1}^{10000}$ を使って求めよ。また、そのときの推定誤差共分散行列を求めよ。
2. 用いるデータ $\{(x_i, y_i)\}_{i=1}^{N}$ を増やしたとき、$\hat{\theta}_N$ の各要素が収束していくことを片対数グラフでプロットして確認せよ。ただし、$N = 4, 8, \cdots, 2^{13} = 8192$ のときのみでよい。
3. 全てのデータを用いたときの決定変数を求めよ.

3　課題 10(重み付き最小二乗法)

$x_i \in \mathbb{R},\ i = 1, \cdots, 10000$ に対し、観測値の次元が 2 次元となる場合を考える。

$$y_i = \varphi(x_i)\theta + w_i \tag{3.1}$$

ここで w_i は $\mathcal{N}(0, V)$ の独立同一分布に従うとし、

$$\varphi = \begin{bmatrix} 1 & x \\ 1 & x^2 \end{bmatrix}, \quad V = \begin{bmatrix} 100 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.2}$$

とする。このとき、

$$\hat{\theta}_N = \left(\sum_{i=1}^{N} \varphi(x_i)^T \varphi(x_i)\right)^{-1} \sum_{j=1}^{N} \varphi(x_j)^T y_j \tag{3.3}$$

$$\hat{\theta}_N = \left(\sum_{i=1}^{N} \varphi_i^T Q_i \varphi_i\right)^{-1} \sum_{j=1}^{N} \varphi_j^T Q_j y_j \tag{3.4}$$

をそれぞれ用いて推定値を求め、それぞれの推定誤差共分散を求めよ。ただし、$Q_i = V^{-1}$ とする。また、$\hat{\theta}_N$ の各要素の収束の仕方をプロットせよ。

4 課題 12(推定値の合成)

$x_i \in \mathbb{R}$, $i = 1, \cdots, 10000$ に対し、次の式で観測データ $y_i \in \mathbb{R}$ が生成されているとする。

$$y_i = \varphi(x_i)\theta + w_i \tag{4.1}$$

ここで

$$\varphi(x) := \begin{bmatrix} 1 & \exp\left(-\frac{(x_i-1)^2}{2}\right) & \exp(-(x_i+1)^2) \end{bmatrix} \tag{4.2}$$

であり、w_i は平均 0、分散有限の同一分布から生成されているもとする。10000 組のデータの組 $\{(x_i, y_i)\}_{i=1}^{10000}$ のうち、最初の 6000 組と残りの 4000 組のデータそれぞれで $\theta \in \mathbb{R}^3$ の推定値を求め、それらから推定値を合成せよ。また、全てデータを使った場合の推定値と比較し、一致することを確かめよ。

5 課題 13(異なる雑音の場合の推定値)

$x_i \in \mathbb{R}$, $i = 1, \cdots, 10000$ に対し、次の式で観測データ $y_i \in \mathbb{R}$ が生成されているとする。

$$y_i = \varphi(x_i)\theta + w_i \tag{5.1}$$

ここで w_i は最初の 6000 組のと残りの 4000 組のデータで、平均ゼロの異なる独立な分布に従うものとする（$i = 1, \cdots, 6000$ では独立同一分布、$i = 6001, \cdots, 10000$ では異なる独立同一分布）。

$$\varphi(x) := \begin{bmatrix} 1 & \exp\left(-\frac{(x_i-1)^2}{2}\right) & \exp(-(x_i+1)^2) \end{bmatrix} \tag{5.2}$$

である。10000 組のデータの組 $\{(x_i, y_i)\}_{i=1}^{10000}$ のうち、最初の 6000 組と残りの 4000 組のデータそれぞれで $\theta \in \mathbb{R}^3$ の推定値と偶然誤差 w_i の分散の推定値を求め、推定された分散を用いて推定値を合成せよ。また、全データを使い、式 (**??**) を用いて得た $\hat{\theta}_{10000}$ と、どちらが優れているか比較考察せよ（比較の仕方を考えること）。

6 課題 14(システム同定)

次のバネ・マス・ダンパ系を表す直線上の運動方程式を考える。

$$M\frac{d^2}{dt^2}y(t) + D\frac{d}{dt}y(t) + Ky(t) = F(t) \tag{6.1}$$

ここで M, K, D は正定数で、$F(t)$ は外力である。$y(t)$ は観測でき、$F(t)$ はこちらで設定できるものとする。このとき、得られるデータから (M, K, D) を決定したい。微少な δt では

$$\frac{d}{dt}y(t) \simeq \frac{y((k+1)\delta t) - y(k\delta t)}{\delta t} \tag{6.2}$$

$$\frac{d^2}{dt^2}y(t) \simeq \frac{y((k+2)\delta t) - 2y((k+1)\delta t) + y(k\delta t)}{\delta t^2} \tag{6.3}$$

と近似できることを利用し、$y_k := y(k\delta t)$, $F_k := F(k\delta t)$ とすると、

$$y_k = \left(2 - \frac{D}{M}\delta t\right) y_{k-1} + \left(1 + \frac{D}{M}\delta t - \frac{K}{M}\delta t^2\right) y_{k-2} + \frac{\delta t^2}{M} F_{k-2} + w_k \tag{6.4}$$

を得る。ここで w_k は近似の際に生じる誤差である。したがって、問題は

$$y_k = x_k^T \theta + w_k, \quad x_k := \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ F_{k-2} \end{bmatrix} \tag{6.5}$$

の θ を求める問題にすることができる。パラメータの真値を $(M, D, K) = (2, 1, 3)$ とし、$\delta t = 0.01$、w_k は $[-1, 1]$ の一様分布に従う各時刻で独立な確率変数であるとして、以下の問いに答えよ。

1. $y_{-2} = 0$, $y_{-1} = 0$ とし、$F_k = 1$, $k = -2, -1, 0, 1, 2, \cdots$ として一定の外力を加えたとき、パラメータ θ を逐次最小二乗法で求めよ。ただし、パラメータの初期値は 0 とし、$k = 10000$ まででよい。また正則化項は工夫せよ。

2. $y_{-2} = 0$, $y_{-1} = 0$ とし、$F_k = \sin(\pi k/5)$, $k = -2, -1, 0, 1, 2, \cdots$ として一定の外力を加えたとき、パラメータ θ を逐次最小二乗法で求めよ。ただし、パラメータの初期値は 0 とし、$k = 10000$ まででよい。また正則化項は工夫せよ。

3. $y_0 = 0$, $y_1 = 0$ とする。このとき、どのような外力を加えればパラメータの推定が上手くいくか考え実装し、逐次最小二乗法で求めよ。ただし、パラメータの初期値は 0 とし、$k = 10000$ まででよい。

7　課題 16(Kalman フィルタ)

次の離散時間ダイナミクスおよび観測方程式が与えられているとする。

$$\theta_k = 0.9\theta_{k-1} + v_k \tag{7.1}$$

$$y_k = 2\theta_k + w_k \tag{7.2}$$

ここで $\theta_k, y_k \in \mathbb{R}$ であり、v_k, w_k は互いに独立な $\mathcal{N}(0, 1)$ に従う確率変数である。また、θ_0 は平均 3、分散 2 をもつ確率分布に従うとする。このとき、$\{y_l\}_{l=1}^k$ までを使った $\theta_k, k = 1, \cdots, 100$ の推定値 $\hat{\theta}_k$ を求め、θ_k を共に時間変化をプロットせよ。

8　課題 18(交互最小二乗法)

$x_i \in [-1, 1]$, $i = 1, \cdots, 10000$ に対して、次の方程式で定められる入出力データが与えられているとする。

$$y_i = \alpha^T \varphi(x_i) \beta + w_i \tag{8.1}$$

ここで $\alpha \in \mathbb{R}^2$, $\beta \in \mathbb{R}^3$, w_i は区間 $[-1, 1]$ の一様分布であるとし、

$$\varphi(x) = \begin{bmatrix} 1 & x & x^2 \\ x & x^2 & x^3 \end{bmatrix} \tag{8.2}$$

の場合を考える。10000 組の $\{(x_i, y_i)\}_{i=1}^{10000}$ から、交互最小二乗法を用いて α, β を推定せよ。このとき、10 組の異なる初期値を用意し、最もよかったパラメータとそのときに使った初期値を示せ。

9　課題 19(K-クラスタリング)

与えられたデータ $x_i \in \mathbb{R}^2$, $i = 1, \cdots, 1000$ を 3 つのクラスターに分類せよ。異なる初期値を 10 組以上選び、最も良さそうな分類をその理由とともに示せ。