Machine Learning (Homework 2)

Due date : 12/01 23:59

1 Sequential Bayesian Learning (50%)

Conjugate prior assures that the posterior distribution has the same functional form as the prior. This means we can use the posterior we computed this time as the prior for the next time. This property plays an important role in sequential Bayesian learning.

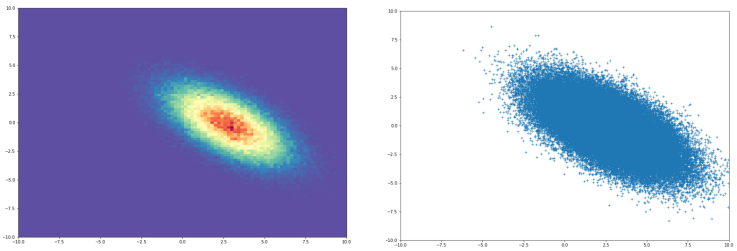
Dataset:
The file 1_data.mat contains two sequences $\mathbf{x} = \{x_1, x_2, ..., x_{100} | 0 \leq x_i \leq 2\}$ and $\mathbf{t} = \{t_1, t_2, ..., t_{100}\}$ which represent the input sequence and the corresponding target sequence, respectively.

Bayesian Learning:
The posterior distribution $p(\mathbf{w}|\mathbf{t}) = \mathcal{N}(\mathbf{w}|\mathbf{m}_N, \mathbf{S}_N)$ with the given prior $p(\mathbf{w}) = \mathcal{N}(\mathbf{w}|\mathbf{m}_0 = \mathbf{0}, \mathbf{S}_0^{-1} = 10^{-6}\mathbf{I})$. The precision of likelihood function $p(\mathbf{t}|\mathbf{w}, \beta)$ or $p(\mathbf{t}|\mathbf{x}, \mathbf{w}, \beta)$ is chosen to be $\beta = 1$.

Basis Function:
Please apply the sigmoidal basis functions $\boldsymbol{\phi} = [\phi_0, \ldots, \phi_{M-1}]^\top$ of the form $\phi_j(x) = \sigma(\frac{x-\mu_j}{s})$ where $\sigma(a)$ is the logistic sigmoid function defined in (3.6). In this exercise, please take the following parameter settings for your basis functions: $M = 3$, $s = 0.1$ and $\mu_j = \frac{2j}{M}$ with $j = 0, \ldots, (M-1)$. please take the data size to be $N = 5, 10, 30$ and 80 for the following two questions:

1. Similar to Fig. 3.9, please generate five curve samples from the parameter posterior distribution.
2. Similar to Fig. 3.8, please plot the predictive distribution of target value t and show the mean curve and the region of variance with one standard deviation on either side of the mean curve.

In addition, similar to the middle column of Fig. 3.7, please arbitrarily select two weights by yourself and carefully plot the corresponding prior distributions when $N = 5, 10, 30$ and 80.



(Both figures are acceptable, but left one will get higher score)

Note : Please train your model by fitting data sequentially, this means when you already compute the result of case $N = 10$, you can continue to compute the result of case $N = 30$.