Day 1 - Afternoon Session

PART-1

Given

$\mathbf{x} = [-1 \quad -0.9 \quad -0.8 \quad -0.7 \quad -0.6 \quad -0.5 \quad -0.4 \quad -0.3 \quad -0.2 \quad -0.1 \; 0 \quad 0.1 \quad 0.2 \quad 0.3 \quad 0.4 \ldots$
$\ldots 0.5 \quad 0.6 \quad 0.7 \quad 0.8 \quad 0.9 \quad 1]$

$\mathbf{t} = [5.12 \quad 4.97 \quad 4.92 \quad 4.83 \; 4.90 \quad 5.06 \quad 5.29 \quad 5.34 \quad 5.36 \quad 5.76 \quad 5.99 \quad 6.30 \ldots$
$\ldots 6.66 \quad 6.70 \quad 7.49 \quad 7.92 \quad 8.48 \; 9.09 \quad 9.70 \quad 10.30 \quad 10.98]$

Order of data is $[\mathbf{A} \quad \mathbf{B} \quad \mathbf{C} \quad \mathbf{A} \quad \mathbf{B} \quad \mathbf{C} \quad \ldots \quad \mathbf{A} \quad \mathbf{B} \quad \mathbf{C}]$

where **A** is Train 1 data, **B** is Train 2 data, and **C** is Train 3 data.

1. Formulate the matrix using Train 1 dataset. $\mathbf{T}^T = \mathbf{w}^T.\phi$

$$[t_1 \; t_2 \; \ldots \; t_7] = [w_1 \; w_2 \; w_3]. \begin{bmatrix} \phi_1(x_1) & \phi_1(x_2) & \ldots & \phi_1(x_7) \\ \phi_2(x_1) & \phi_2(x_2) & \ldots & \phi_2(x_7) \\ \phi_3(x_1) & \phi_3(x_2) & \ldots & \phi_3(x_7) \end{bmatrix}$$

2. Solve for $[w_1 \; w_2 \; w_3]^T = \mathbf{w}$ as follows:
   $\mathbf{T}_{(7\times1)} = \phi^T_{(7\times3)}.\mathbf{w}_{(3\times1)}$ $\qquad$ $\mathbf{w}_{(3\times1)} = (\phi_{(3\times7)}.\phi^T_{(7\times3)} + \lambda\mathbf{I}_{(3\times3)})^{-1}\phi_{(3\times7)}$
   with $\lambda = 0$

   $\boxed{\mathbf{w}_1^{(1)} =}$

3. Plot x_{train1} versus t_{train1} (Scatter Plot)
   Plot x_{train1} versus $\phi^T(x_{train1})\mathbf{w}_1^{(1)}$ (Line Plot)
   Plot x_{out} versus $\phi^T(x_{out})\mathbf{w}^{(1)}$
   where $x_{out} = \left[-0.95 \quad -0.85 \quad \ldots \quad -0.05 \quad 0.05 \quad \ldots \quad 0.95\right]^T$

   Observation: $\boxed{}$

4. Repeat steps 1, 2 and 3 using Train 2, Train 3

5. $\boxed{var =}$ $\qquad$ $\boxed{Bias^2 =}$
   Assume $E[t/x] = g(x) = truevalue$ for calculating $bias^2$.

6. Repeat with $\lambda = 2$, $\lambda = 4$, $\lambda = 5$, $\lambda = 6$, $\lambda = 10$.

7. Plot λ versus var. Also plot λ versus $bias^2$
   $true\,value = [5 \; 4.92 \; 4.88 \; 4.88 \; 4.92 \; 5.00 \; 5.12 \; 5.28 \; 5.48 \; 5.72 \; 6.00 \; 6.32 \ldots$
   $\ldots 6.68 \; 7.08 \; 7.52 \; 8.00 \; 8.52 \; 9.08 \; 9.68 \; 10.32 \; 11.00]$
   NOTE: $\phi_1(x) = 1$, $\phi_2(x) = exp(-\frac{(x-0.5)^2}{0.1}).exp(-\frac{(x+0.5)^2}{0.1})$

PART 2

Estimate $\underline{\mathbf{w}}$ using iterative technique.

1. Initialize $\underline{\mathbf{w}} = [0\ 0\ 0]^T$, $i = 1$
2. Obtain the error corresponding to first data
   $e = [t_i - \phi_1(x_i)w_1 - \phi_2(x_i)w_2 - \phi_3(x_i)w_3]$
3. Update $\underline{w}(t+1) = \underline{w}(t) + \eta e \begin{bmatrix} \phi_1(x_i) \\ \phi_2(x_i) \\ \phi_3(x_i) \end{bmatrix}$
4. Repeat steps 2 and 3 for $i = 2, \ldots, 14$.
5. Compute SSE for the data $i = 15, \ldots, 21$. This is 1 epoch.
6. Repeat steps 2 to 4 for 10 times and plot SSE versus Number of epochs.
7. Identify $\underline{\mathbf{w}}$ vector after 10 epochs.

$w_1 =$	$w_2 =$	$w_3 =$

PART-3 Kernel Trick for Regression Techniques

1. Given
   Train Input: $X_I = \begin{bmatrix} -1 & -0.7 & -0.4 & -0.1 & 0.2 & 0.5 & 0.8 \end{bmatrix}$
   Validation Input: $V_I = \begin{bmatrix} -0.9 & -0.6 & -0.3 & 0 & 0.3 & 0.6 & 0.9 \end{bmatrix}$
   Test Input: $T_I = \begin{bmatrix} -0.8 & -0.5 & -0.2 & 0.1 & 0.4 & 0.7 & 1 \end{bmatrix}$
   Train Output: $X_O = \begin{bmatrix} 5.12 & 4.83 & 5.29 & 5.76 & 6.66 & 7.92 & 9.70 \end{bmatrix}$
   Validation Output: $V_O = \begin{bmatrix} 4.97 & 4.90 & 5.34 & 5.99 & 6.70 & 8.48 & 10.30 \end{bmatrix}$
   Test Output: $T_O = \begin{bmatrix} 4.92 & 5.06 & 5.36 & 6.30 & 7.49 & 9.09 & 10.98 \end{bmatrix}$

   (a) $K(X_1, X_2) = exp(-\frac{(X_1 - X_2)^2}{2\sigma^2})$ be the kernel function
   (b) Choose $\sigma^2 = 0.1$
   (c) Compute the vector for $i = 1$

$$M = \begin{bmatrix} K(X_I(1), X_V(i)) \\ K(X_I(2), X_V(i)) \\ K(X_I(3), X_V(i)) \\ K(X_I(4), X_V(i)) \\ K(X_I(5), X_V(i)) \\ K(X_I(6), X_V(i)) \\ K(X_I(7), X_V(i)) \end{bmatrix} =$$

(d) Normalize the vector: $M = \frac{M}{sum(M)}$

(e) Compute the output: $M^T X_O = \hat{V}_O(i)$

(f) Compute for $i = 2, 3, 4, 5, 6, 7$

(g) Compute $SSE = [V_O - \hat{V}_O]^T [V_O - \hat{V}_O]$

(h) Repeat c,d,e for $\sigma^2 = 0.2, 0.3, 0.4, 0.5$. Tabulate.

$\sigma^2 = 0.2$	$\sigma^2 = 0.3$	$\sigma^2 = 0.4$	$\sigma^2 = 0.5$
$SSE =$	$SSE =$	$SSE =$	$SSE =$

Identify $\boxed{\sigma^2 = \quad}$ corresponding to least SSE.

(i) Compute M using the σ^2 that gives the least SSE.

$$M = \begin{bmatrix} K(X_I(1), X_V(i)) \\ K(X_I(2), X_V(i)) \\ K(X_I(3), X_V(i)) \\ K(X_I(4), X_V(i)) \\ K(X_I(5), X_V(i)) \\ K(X_I(6), X_V(i)) \\ K(X_I(7), X_V(i)) \end{bmatrix} \forall\ i = 1$$

Normalze the matrix vector as $M = \frac{M}{sum(M)}$

(j) Compute the output: $\hat{T}_O(i) = M^T T_I$, with $i = 1$

(k) Repeat steps i and j for $i = 2, 3, 4, 5, 6, 7$. Tabulate.

	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$	$i = 7$
T_O							
$\hat{T}_O$							

(l) $SSE = [T_O(i) - \hat{T}_O(i)]^T [T_O(i) - \hat{T}_O(i)]$

$\boxed{SSE = \quad}$