Assignment 7

FOML (IT-582)

Dhirubhai Ambani University

1 Problem Statement

Consider the true underlying function $f(x) = x^3 - x^2 + x + 2$, with input x in the range $[-3, 3]$.

The task is to perform the following steps to analyze the bias-variance trade-off for two different hypothesis classes:

1. Sample 100 values of x uniformly from the interval $[-3, 3]$ and compute the corresponding true values $f(x_i)$ for $i = 1$ to 100.

2. For each x_i, generate noisy observations $y_i^{(k)} = f(x_i) + \epsilon_i^{(k)}$, where $\epsilon_i^{(k)}$ is independently sampled from a normal distribution $\mathcal{N}(0, 0.1)$. Repeat this process 10 times (for $k = 1$ to 10) to obtain 10 noisy datasets $\mathbf{y}^{(k)}$, each a vector of size 100×1.

3. **Linear Hypothesis Class:** For each dataset $\mathbf{y}^{(k)}$, fit a linear regression model of the form $h(x) = \theta_0 + \theta_1 x$. Obtain the predicted values $\hat{\mathbf{y}}^{(k)}$ (size 100×1) on the same x points. This results in 10 predicted vectors.

4. Compute the bias and variance for the linear models.

5. **Polynomial Hypothesis Class:** Repeat the fitting process using a higher-order polynomial model $h(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4 + \theta_5 x^5$. For each $\mathbf{y}^{(k)}$, obtain predictions $\hat{\mathbf{y}}^{(k)}$ and compute bias and variance in the same manner as above.

6. Plot the true function $f(x)$ over $[-3, 3]$.

7. For the linear case, plot the true $f(x)$ along with the 10 fitted linear model curves on the same graph.

8. For the polynomial case, plot the true $f(x)$ along with the 10 fitted polynomial model curves on the same graph.