**CPT\_S 534 HW3**

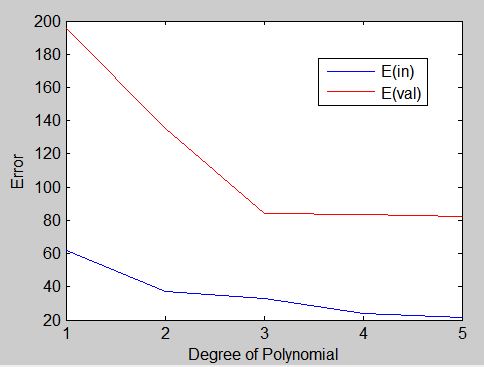
**Yang Zhang**

**11529139**

1. **Calculate Eval at each degree(1-5)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Degree(order)** | **1** | **2** | **3** | **4** | **5** |
| **Eval** | 195.6959 | 135.2558 | 84.3556 | 83.6796 | 82.5139 |
| **Ein** | 67.7956 | 37.0746 | 32.8439 | 23.7083 | 21.8037 |

1. **Plot your result to find the “elbow” in Eval and best complexity for data mining**

****

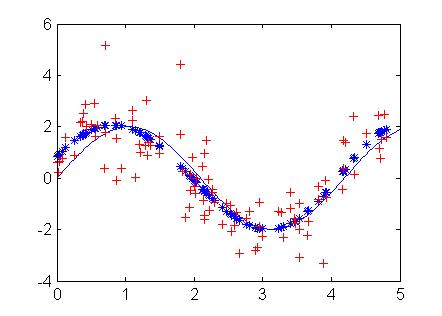
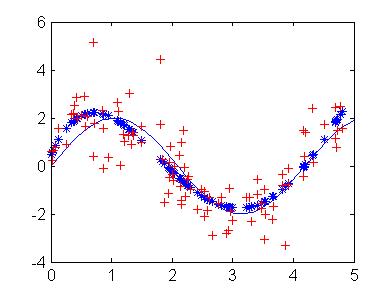
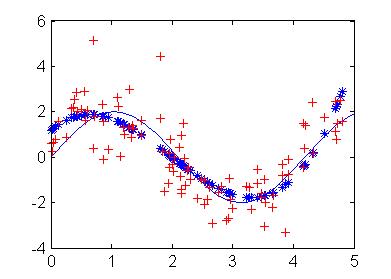
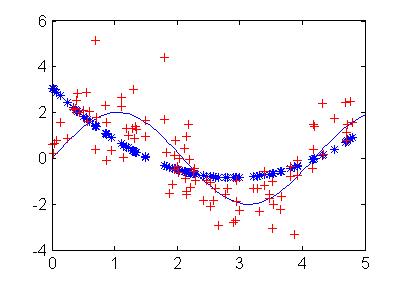
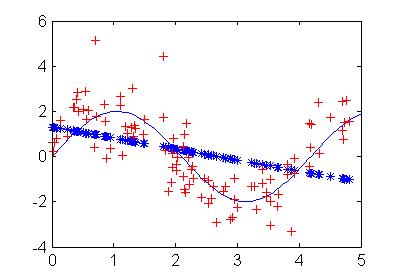
For the graph above, the elbow degree is 3, because after degree of 3 there is no significant decrease of the error.

1. **Use the full data set to find the optimum polynomial of best complexity**

**Show this result as plot of data and fit on the same set of axes.**

**Report the minimum sum of squared residuals and coefficient of determination**

Figures for degree 1 to 5:



Degree of 5

Degree of 4

Degree of 3

Degree of 2

Degree of 1

The optimum polynomial of the complexity is 3, because after degree of 3 there is no significant improve of data fitting.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Degree** | **1** | **2** | **3** | **4** | **5** |
| **Sum of squared residuals(Ein)** | 250.4270 | 175.9094 | 105.7786 | 97.4958 | 94.5210 |
| **Coefficient of determination** | 0.1481 | 0.4016 | 0.6402 | 0.6684 | 0.6785 |

The minimum sum of squared residuals is 94.5210