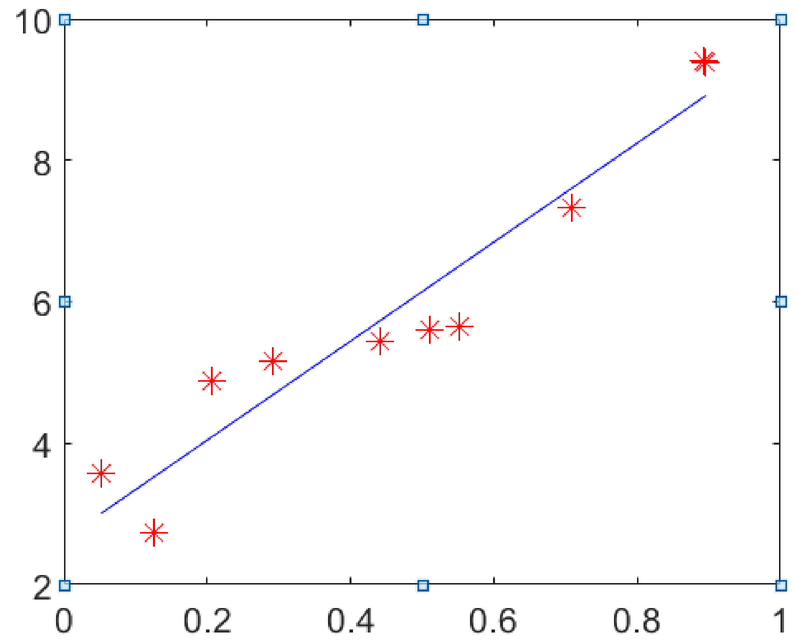
Beyond Linear Regression models

Dr. Pritam Anand.
Assistant Professor,
DA-IICT, Gandhinagar.

Linear Regression model



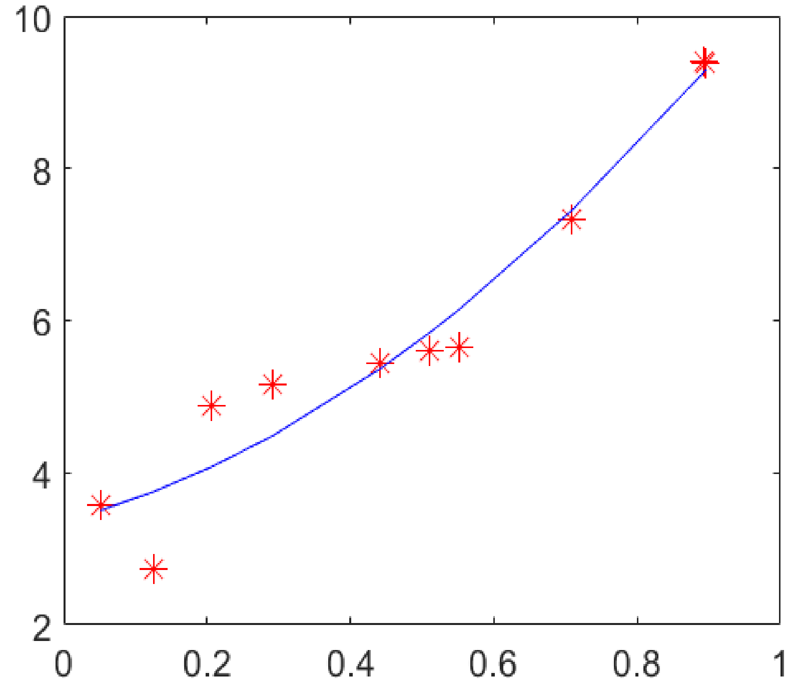
Linear Regression model

Training $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(yi - f(xi))^2} =$ 0.5947

Testing $RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(yi - f(xi))^2} =$ 0.9426

Income (x) (thousand dollar)	Balance (y) (thousand dollar)	Estimated f(x) (thousand dollar)
0.96703	9.675083	9.41205399
0.547232	6.293266	6.47467789
0.972684	9.730614	9.45161938
0.714816	7.474346	7.64728226
0.697729	7.342933	7.5277212
0.216089	4.619033	4.15763074
0.976274	9.765597	9.47673972
0.00623	4.012784	2.68921946
0.252982	4.762698	4.41577473
0.434792	5.626166	5.68791617
0.779383	7.989045	8.09906511
0.197685	4.552625	4.02885271
0.862993	8.705537	8.6840969
0.983401	9.835217	9.52660279
0.163842	4.43522	3.79205019
0.597334	6.622444	6.8252457
0.008986	4.01882	2.70850244
0.386571	5.37153	5.35051307
0.04416	4.096522	2.95461902
0.956653	9.574695	9.33944573

Quadratic Fitting



$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$$

Quadratic Fitting

For given Training Set T = { (x_1,y_1), (x_2,y_2),….., (x_n,y_n)}, we need to solve

$$Min_{(\beta_2,\beta_1,\beta_0)} \quad J(\beta_2,\beta_1,\beta_0) = \quad \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i + \beta_2 x_i^2))^2 \qquad \text{.....(2)}$$

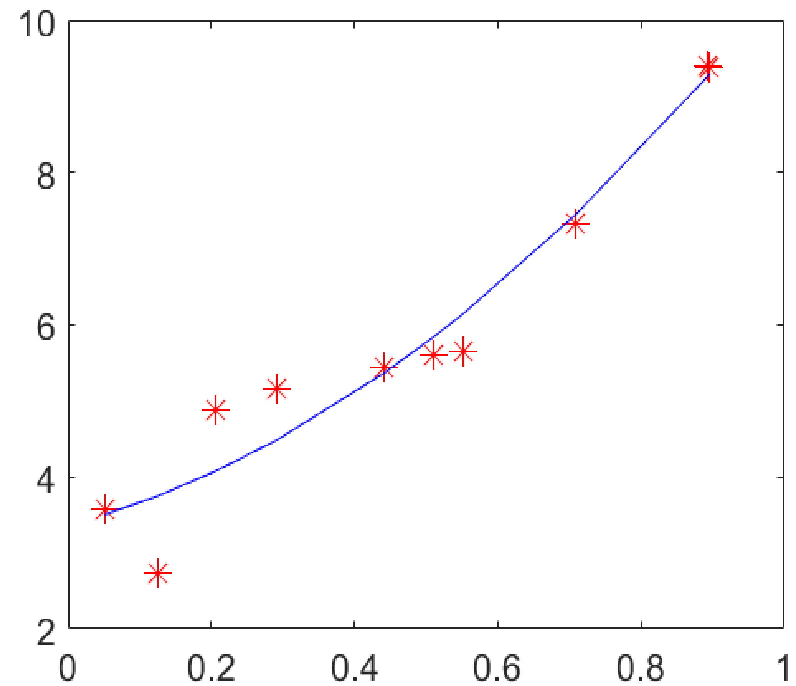
$$u = \begin{bmatrix} \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} \qquad A = \begin{bmatrix} (x_1)^2 & x_1 & 1 \\ (x_2)^2 & x_2 & 1 \\ (x_3)^2 & x_3 & 1 \\ (x_4)^2 & x_4 & 1 \\ (x_5)^2 & x_5 & 1 \\ (x_6)^2 & x_6 & 1 \\ & & \\ & & \\ (x_n)^2 & x_n & 1 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ \\ \\ y_n \end{bmatrix}$$

The Least Square problem (2) reduces to

$$Min_{(u)} \quad J(u) = \quad (Y - Au)^T (Y - Au)$$

$$u = \begin{bmatrix} \beta_2 \\ \beta_1 \\ \beta_0 \end{bmatrix} = (A^T A)^{-1} A^T Y$$

Quadratic Fitting

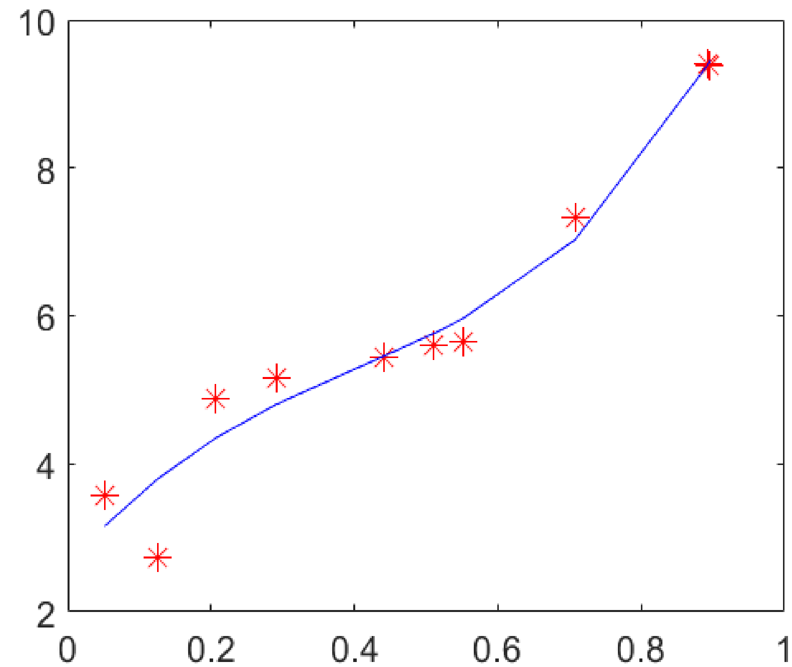


$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$

$= 3.36 + 2.51\,x + 4.58x^2$

Training $RMSE = \ 0.4980$

Fitting with fourth order polynomial



$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4$

For given Training Set $T = \{ (x_1,y_1), (x_2,y_2),\ldots\ldots, (x_n,y_n)\}$, we need to solve

$$Min_{(\beta_4,\beta_3,\beta_2,\beta_1,\beta_0)} \quad J(\beta_4,\beta_3,\beta_2,\beta_1,\beta_0) = \sum_{i=1}^{n} (yi-(\beta_0 + \beta_1 x_i+\beta_2 x_i^2 + \beta_3 x_i^3+\beta_4 x_i^4))^2 \qquad \ldots(3)$$

For given Training Set $T = \{ (x_1,y_1), (x_2,y_2),\ldots, (x_n,y_n)\}$, we need to solve

$$Min_{(\beta_4,\beta_3,\beta_2,\beta_1,\beta_0)} \quad J(\beta_4,\beta_3,\beta_2,\beta_1,\beta_0) = \sum_{i=1}^{n} (y_i-(\beta_0 + \beta_1 x_i+\beta_2 x_i^2 + \beta_3 x_i^3+\beta_4 x_i^4))^2 \qquad \ldots(3)$$

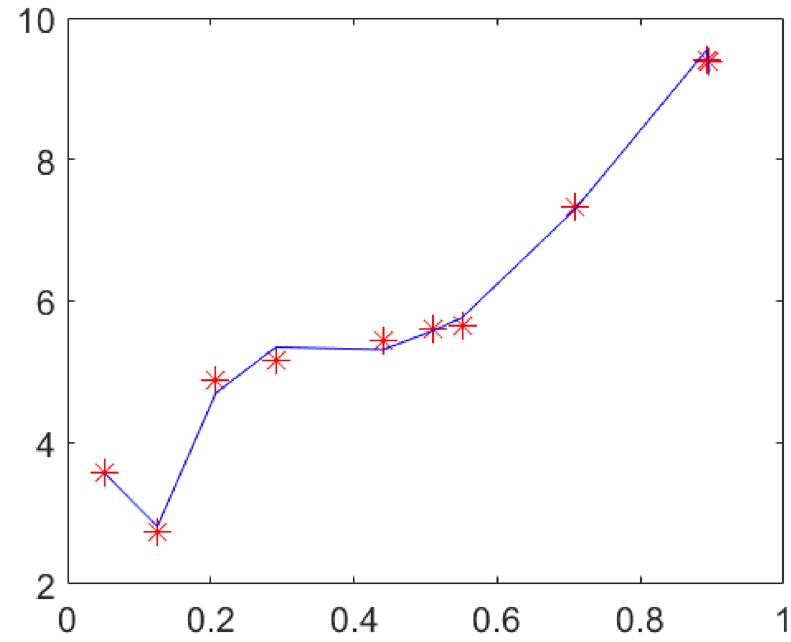
$$u=\begin{bmatrix}\beta_4\\ \beta_3\\ \beta_2\\ \beta_1\\ \beta_0\end{bmatrix} \qquad A = \begin{bmatrix}(x_1)^4 & (x_1)^3 & (x_1)^2 & x_1 & 1\\ (x_2)^4 & (x_2)^3 & (x_2)^2 & x_2 & 1\\ (x_3)^4 & (x_3)^3 & (x_3)^2 & x_3 & 1\\ & & & & \\ & & & & \\ (x_n)^4 & (x_n)^3 & (x_n)^2 & x_n & 1\end{bmatrix} \qquad Y = \begin{bmatrix}y_1\\ y_2\\ y_3\\ y_4\\ y_5\\ y_6\\ \\ \\ y_n\end{bmatrix}$$

The Least Square problem (3) reduces to

$$Min_{(u)} \; J(u) = (Y - Au)^T (Y - Au)$$

$$u=\begin{bmatrix}\beta_4\\ \beta_3\\ \beta_2\\ \beta_1\\ \beta_0\end{bmatrix} = (A^T A)^{-1} A^T Y$$

Fitting with seven order polynomial

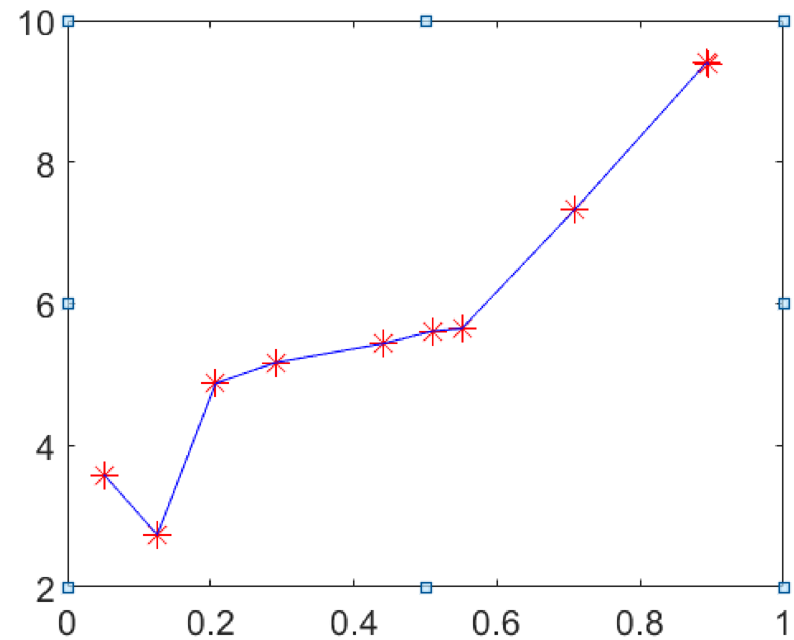


$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \beta_5 x^5 + \beta_6 x^6 + \beta_7 x^7$

$= 11.89 - 283.034\,x + 3015\,x^2 - 14643.7x^3 + 38006.62x^4 - 54565.9x^5 + 40844.5x^6 - 12458.5x^7$

Training $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - f(xi))2} = 0.1186$

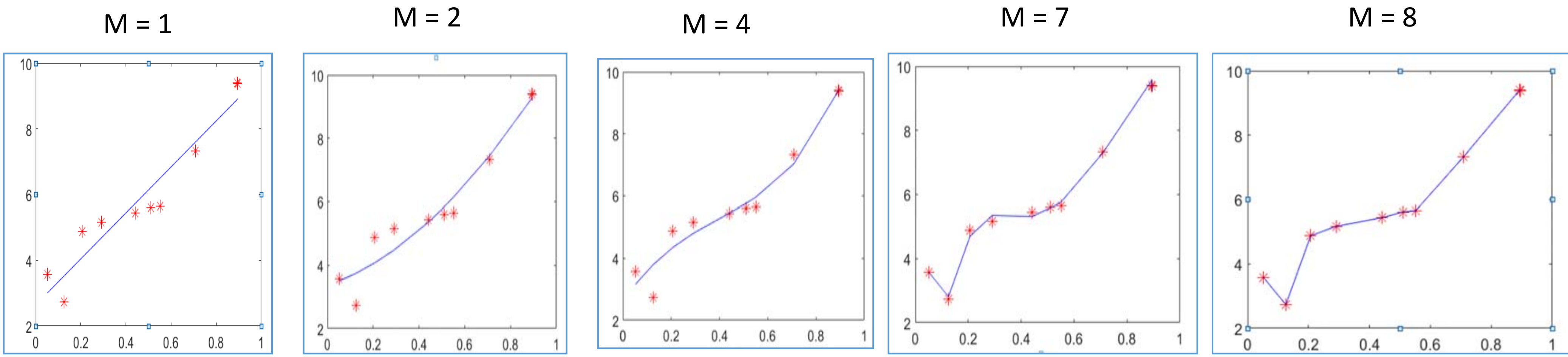
Fitting with eight order polynomial



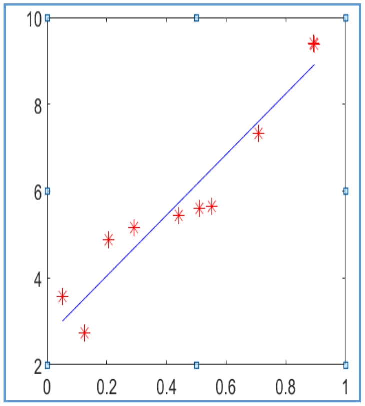
$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \beta_5 x^5 + \beta_6 x^6 + \beta_7 x^7 + \beta_8 x^8$

$= 18.14 - 527.837\,x + 6379\,x^2 - 37080.8x^3 + 120518.8x^4 - 230990x^5 + 256860.2x^6 - 154208x^7 + 38542x^8$

Training $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(yi - f(xi))2} = 0.0026$

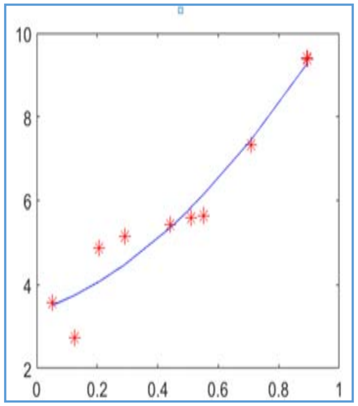

M = 1
M = 2
M = 4
M = 7
M = 8

M = 1



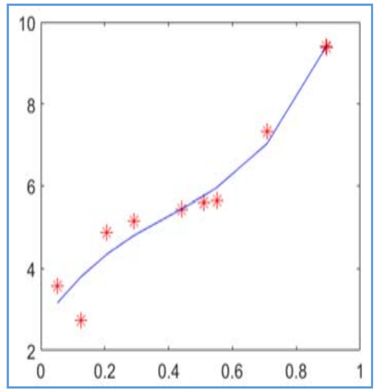
Train RMSE
= 0.5947

M = 2



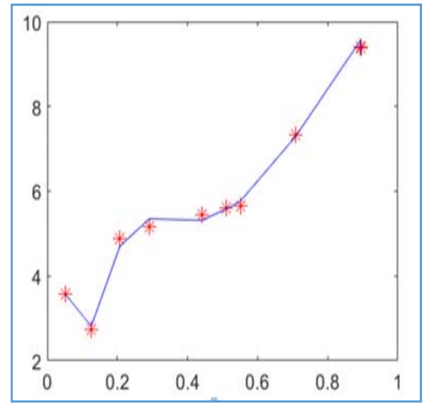
Train RMSE
= 0. 4980

M = 4



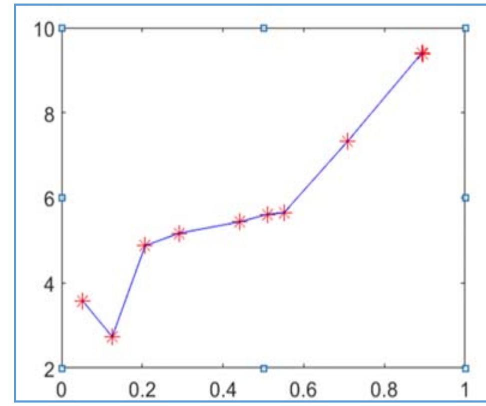
Train RMSE
= 0. 4387

M = 7



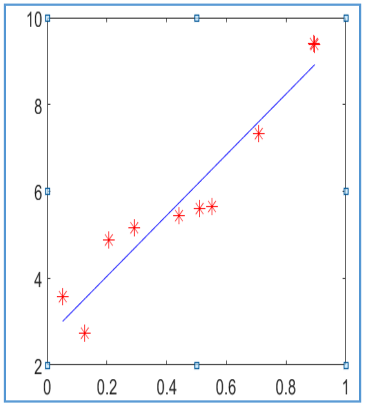
Train RMSE
= 0. 1186

M = 8



Train RMSE
= 0. 0026

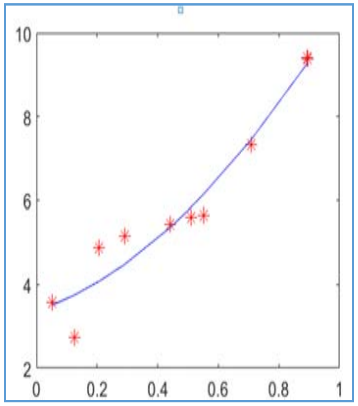
M = 1



Train RMSE
= 0.5947

Test RMSE
= 0.9426

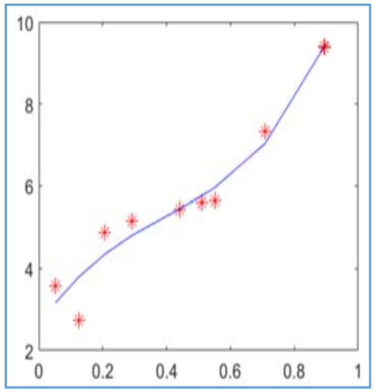
M = 2



Train RMSE
= 0.4980

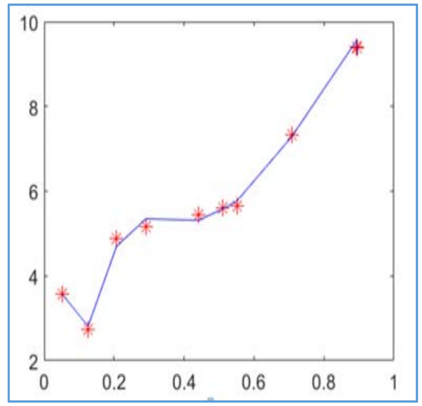
Test RMSE
= 0.7711

M = 4



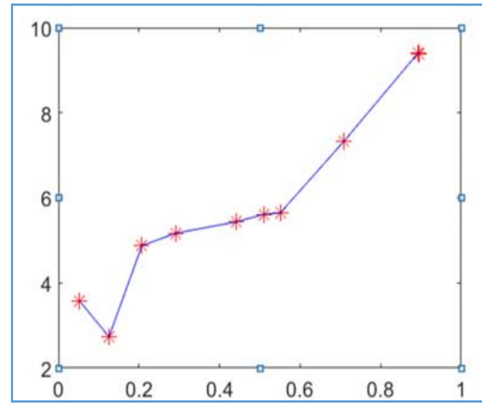
Train RMSE
= 0.4387

M = 7



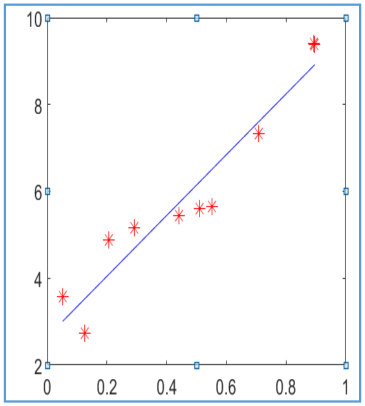
Train RMSE
= 0.1186

M = 8



Train RMSE
= 0.0026

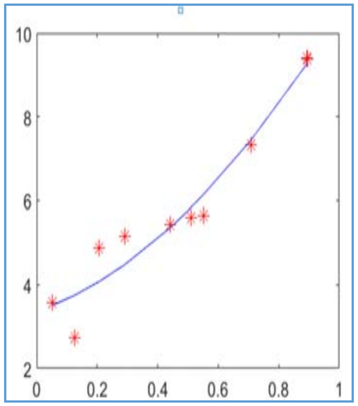
M = 1



Train RMSE
= 0.5947

Test RMSE
= 0.9426

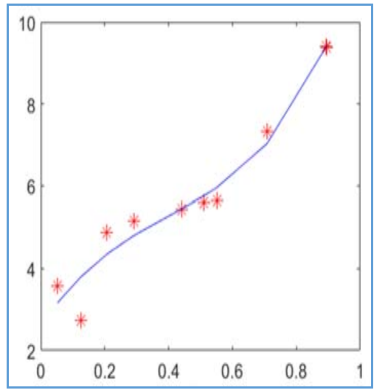
M = 2



Train RMSE
= 0. 4980

Test RMSE
= 0.7711

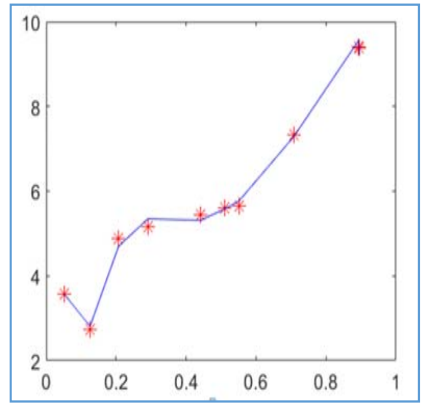
M = 4



Train RMSE
= 0. 4387

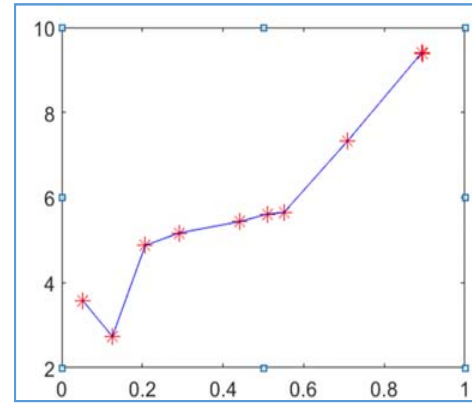
Test RMSE
= 0.9811

M = 7



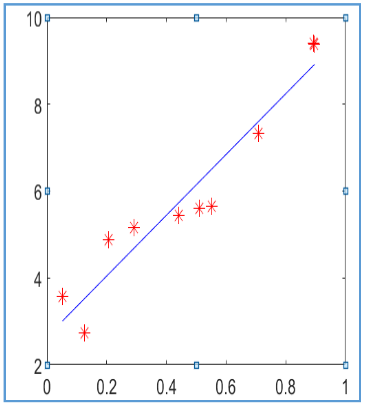
Train RMSE
= 0. 1186

M = 8



Train RMSE
= 0. 0026

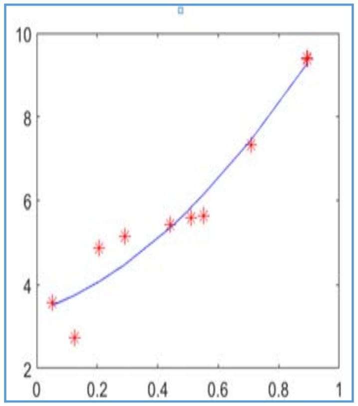
M = 1



Train RMSE
= 0.5947

Test RMSE
= 0.9426

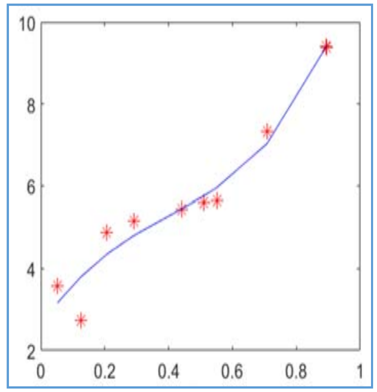
M = 2



Train RMSE
= 0. 4980

Test RMSE
= 0.7711

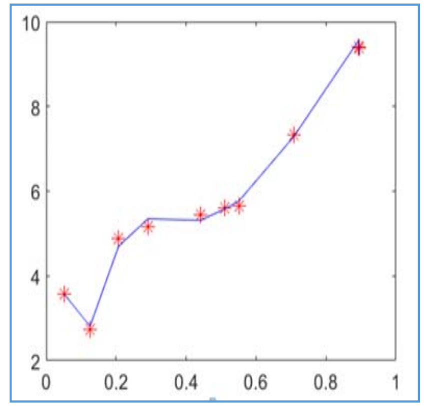
M = 4



Train RMSE
= 0. 4387

Test RMSE
= 0.9811

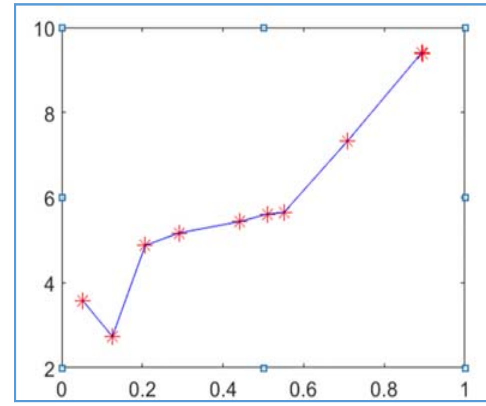
M = 7



Train RMSE
= 0. 1186

Test RMSE
=1.179

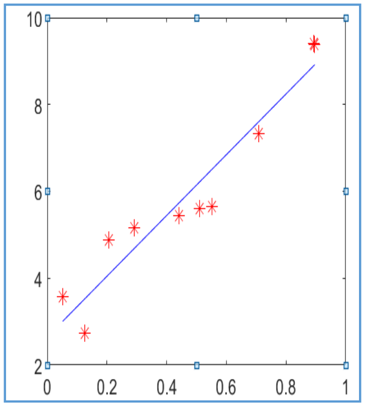
M = 8



Train RMSE
= 0. 0026

Test RMSE
= 3.90

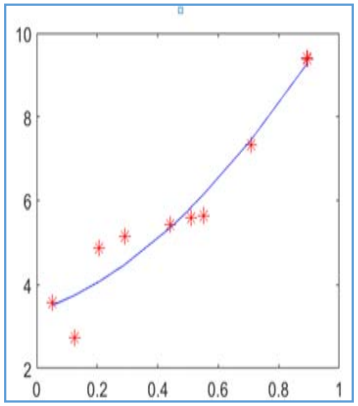
M = 1



Train RMSE
= 0.5947

Test RMSE
= 0.9426

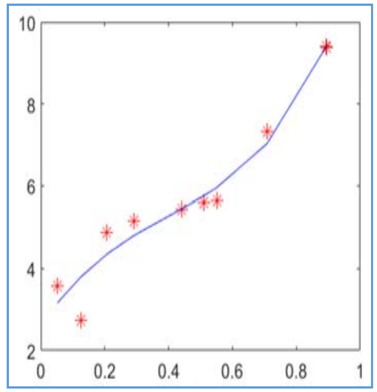
M = 2



Train RMSE
= 0.4980

Test RMSE
= 0.7711

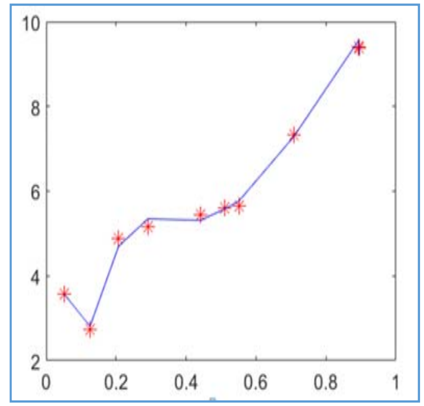
M = 4



Train RMSE
= 0.4387

Test RMSE
= 0.9811

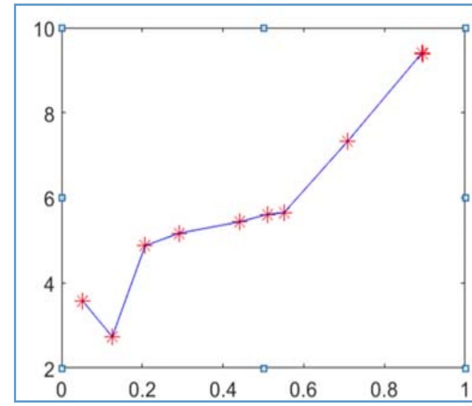
M = 7



Train RMSE
= 0.1186

Test RMSE
=1.179

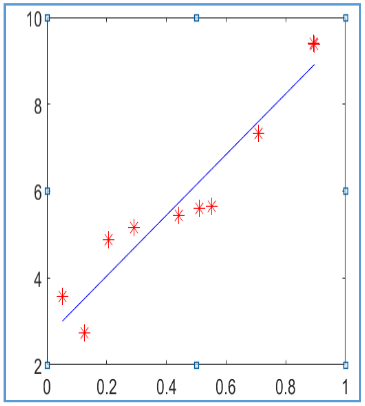
M = 8



Train RMSE
= 0.0026

Test RMSE
= 3.90

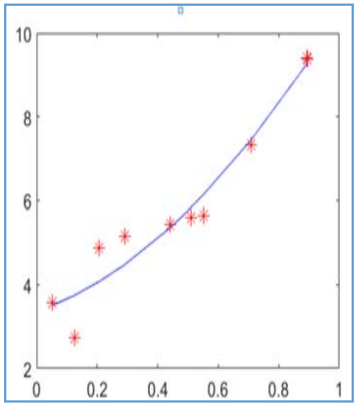
M = 1



Train RMSE
= 0.5947

Test RMSE
= 0.9426

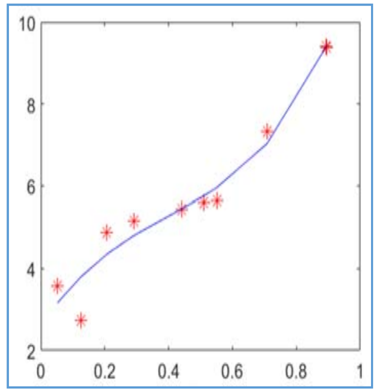
M = 2



Train RMSE
= 0. 4980

Test RMSE
= 0.7711

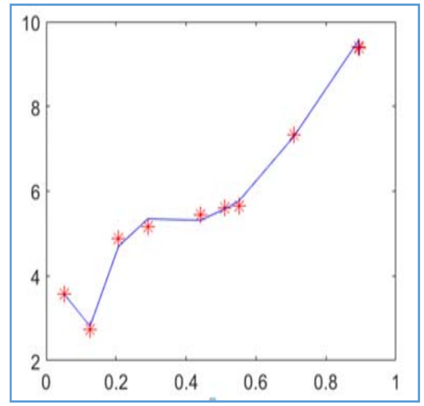
M = 4



Train RMSE
= 0. 4387

Test RMSE
= 0.9811

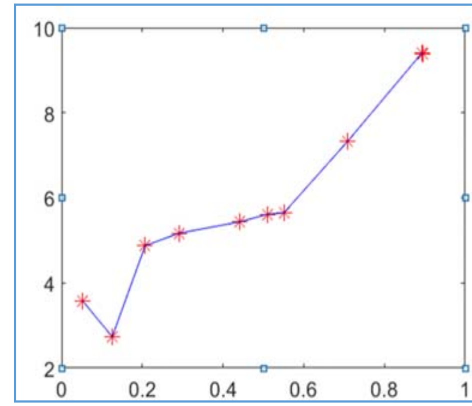
M = 7



Train RMSE
= 0. 1186

Test RMSE
=1.179

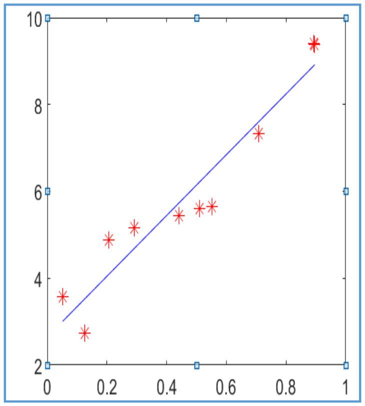
M = 8



Train RMSE
= 0. 0026

Test RMSE
= 3.90

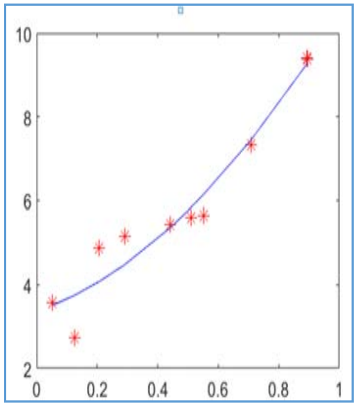
M = 1



Test RMSE
= 0.9426

$$\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} = \begin{bmatrix}2.6\\ 6.9\end{bmatrix}$$

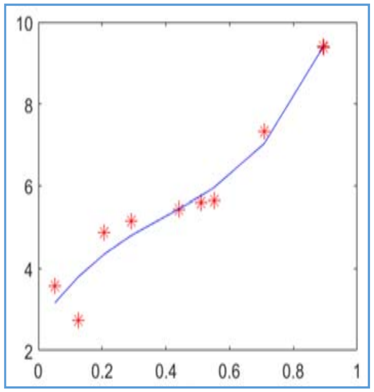
M = 2



Test RMSE
= 0.7711

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix} = \begin{bmatrix}3.36\\ 2.51\\ 4.58\end{bmatrix}$$

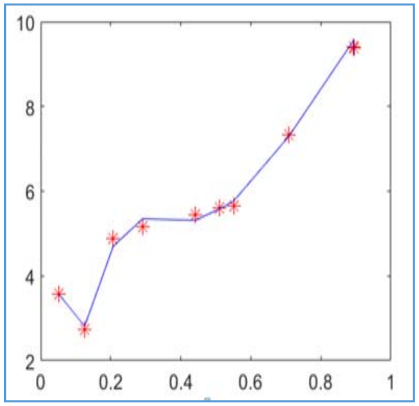
M = 4



Test RMSE
= 0.9811

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\end{bmatrix} = \begin{bmatrix}2.62\\ 11.14\\ 15.58\\ 9.17\\ 4.26\end{bmatrix}$$

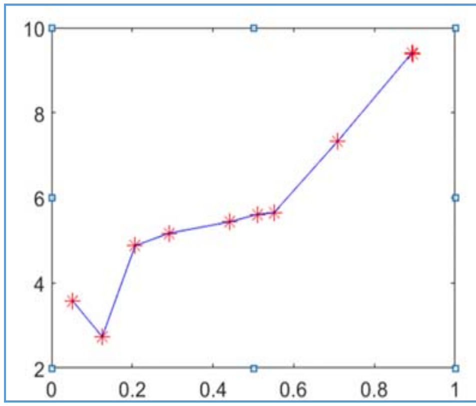
M = 7



Test RMSE
=1.179

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\\ \beta_5\\ \beta_6\\ \beta_7\end{bmatrix} = \begin{bmatrix}11.89\\ -283.034\\ 3015.61\\ -14643.7\\ 38006.62\\ -54565.9\\ 40844.45\\ -12458.5\end{bmatrix}$$

M = 8



Test RMSE
= 3.90

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\\ \beta_5\\ \beta_6\\ \beta_7\\ \beta_8\end{bmatrix} = \begin{bmatrix}18.14\\ -527.83\\ 6379.38\\ 37080\\ 120518.80\\ -230390\\ 256860.2\\ 154208\\ 38542.75\end{bmatrix}$$

Improving the prediction for M=7

$$u = \begin{bmatrix} \beta_7 \\ \beta_6 \\ \vdots \\ \beta_0 \end{bmatrix}$$

$$\underset{u}{Min} \ \cancel{\frac{1}{n}} \left\{ \sum_{i=1}^{n} (yi-(\beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \beta_4 x_i^4 + \beta_5 x_i^5 + \beta_6 x_i^6 + \beta_7 x_i^7))^2 \right\}$$

$$(Y - Au)^T (Y - Au)$$

$$+ \left\{ \frac{\lambda}{2} \underbrace{(\beta_0^2+\beta_1^2+\beta_2^2+\beta_3^2+\beta_4^2+\beta_5^2+\beta_6^2+\beta_7^2)}_{u^T u} \right\}$$

User defined parameter

$$\min_{u} J(u) = \underbrace{(y - Au)^T (y - Au)}_{} + \lambda u^T u, \quad \text{where}$$

$$u = \begin{bmatrix} \beta_7 \\ \beta_6 \\ - \\ - \\ \beta_0 \end{bmatrix}$$

$$\nabla_u J(u) = -2A^T(y - Au) + 2\lambda u = 0$$

$$\Rightarrow -A^T y + A^T A u + \lambda u = 0$$

$$\Rightarrow -A^T y + A^T A u + \lambda I u = 0$$

$$\Rightarrow -A^T y + (A^T A + \lambda I) u = 0$$

$$\Rightarrow (A^T A + \lambda I) u = A^T y$$

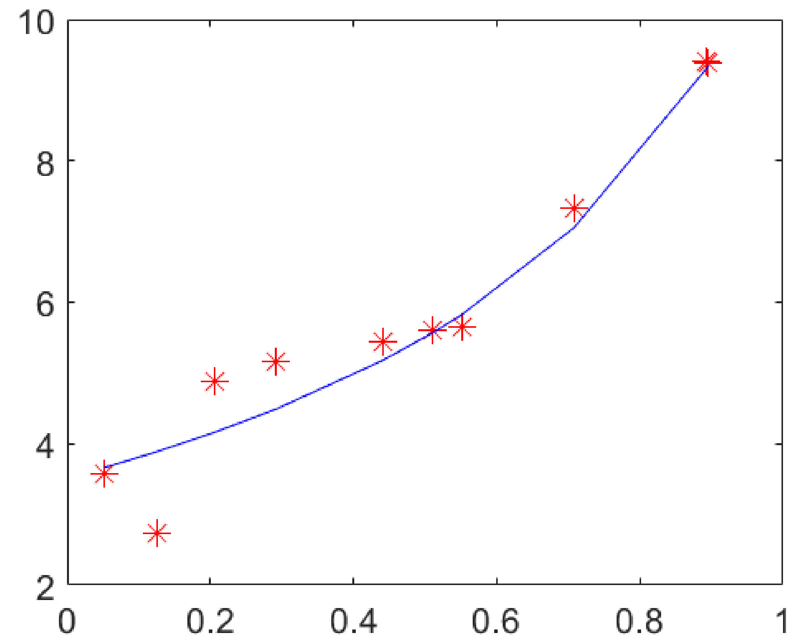
$$\Rightarrow u = (A^T A + \lambda I)^{-1} A^T y$$

$$A = \begin{bmatrix} x_1^7 & x_1^6, & \dots & 1 \\ - & - & & - \\ - & - & & - \\ x_n^7 & x_n^6, & \dots & 1 \end{bmatrix}$$

$$y = \begin{bmatrix} y_n \\ - \\ - \\ - \\ y_1 \end{bmatrix}$$

Estimation with regularization

$$\begin{bmatrix}\beta_0\\\beta_1\\\beta_2\\\beta_3\\\beta_4\\\beta_5\\\beta_6\\\beta_7\end{bmatrix} = \begin{bmatrix}3.51\\2.75\\1.60\\1.05\\0.78\\0.61\\0.50\\0.41\end{bmatrix}$$



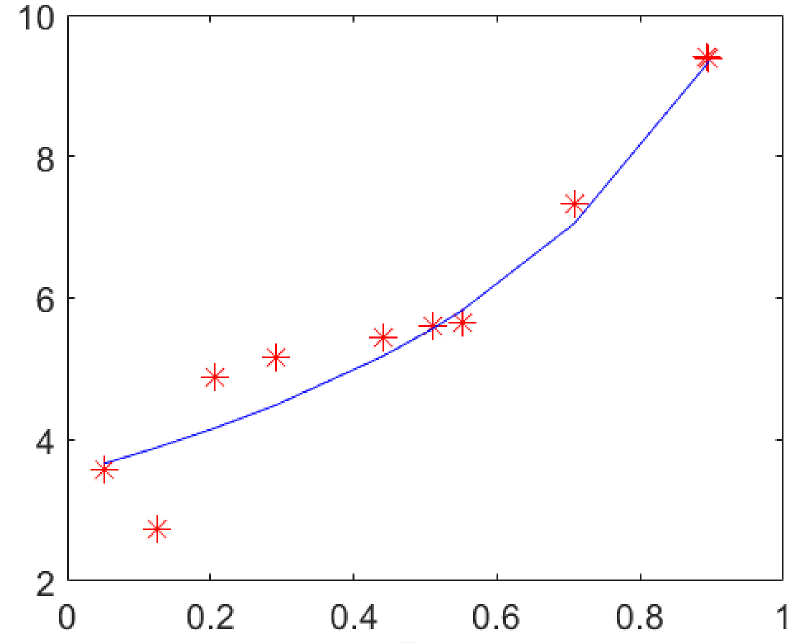
M = 7 order polynomial estimation

Estimation with regularization

Train RMSE
= 0.4989

Test RMSE
= 0.8646

$$\begin{bmatrix}\beta_0\\\beta_1\\\beta_2\\\beta_3\\\beta_4\\\beta_5\\\beta_6\\\beta_7\end{bmatrix} = \begin{bmatrix}3.51\\2.75\\1.60\\1.05\\0.78\\0.61\\0.50\\0.41\end{bmatrix}$$



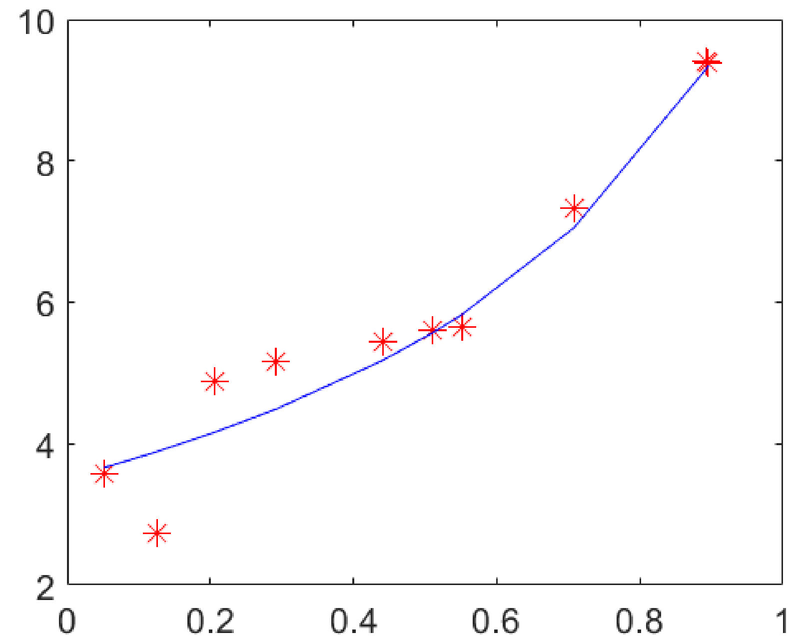
M = 7 order polynomial estimation

Estimation with regularization

Train RMSE
= 0.4989

Test RMSE
= 0.8646
which was 3.90

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \end{bmatrix} = \begin{bmatrix} 3.51 \\ 2.75 \\ 1.60 \\ 1.05 \\ 0.78 \\ 0.61 \\ 0.50 \\ 0.41 \end{bmatrix}$$



M = 7 order polynomial estimation
with λ = 0.25