Math 156
HW1

Osman Akar
004-649-066

October 10, 2017

Task 1

Let's define $t = (t_1, t_2, ..., t_N)^T$ and let $\mathbf{Z}$ be N x $(M+1)$ matrix such that i^{th} row is $\mathbf{z}_i$. We have

$$\begin{aligned}
\frac{\partial}{\partial w_i} J(\mathbf{w}) &= \frac{1}{2}\sum_{n=1}^{N} \frac{\partial}{\partial w_i}(y(x_n, \mathbf{w}) - t_n)^2 \\
&= \frac{1}{2}\sum_{n=1}^{N} 2(y(x_n, \mathbf{w}) - t_n)\frac{\partial}{\partial w_i}(y(x_n, \mathbf{w}) - t_n) \\
&= \sum_{n=1}^{N}(y(x_n, \mathbf{w}) - t_n)x_n^i \\
&= \sum_{n=1}^{N}(\mathbf{z}_n \cdot \mathbf{w} - t_n)x_n^i \\
&= (\mathbf{Z}\mathbf{w}^T - t)^T \cdot (x_1^i, x_2^i, ..., x_N^i)^T
\end{aligned}$$

Thus we can conclude that

$$\triangledown J(\mathbf{w}) = ((\mathbf{Z}\mathbf{w}^T - t)^T\mathbf{Z})^T = \mathbf{Z}^T(\mathbf{Z}\mathbf{w}^T - t)$$

so that the equation $J(\mathbf{w}) = 0$ can be written as $\mathbf{Z}^T\mathbf{Z}\mathbf{w}^T = \mathbf{Z}^T t$. Here we choose $A = \mathbf{Z}^T\mathbf{Z}$ and $b = \mathbf{Z}^T t$

Task 2

We will consider two cases.

Case 1 $M + 1 \leq N$

Let Z_0 be the upper $(M+1)$x$(M+1)$ part of $\mathbf{Z}$, so that the rows of Z_0 are $\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_{M+1}$. Z_0 is *Vardenmonde matrix*, thus it is invertible as $x_1, x_2, ..., x_{M+1}$ are pairwise different. Therefore $rank(\mathbf{Z}) \geq rank(Z_0) = M + 1$. As an linear algebra result [1], we have $rank(A) = rank(\mathbf{Z}^T\mathbf{Z}) = rank(\mathbf{Z}) \geq M + 1$. Moreover, we know that A is $(M+1)$x$(M+1)$ matrix, which concludes A is an invertible matrix, so the equation $A\mathbf{w}^T = b$ has a unique solution.

[1] see https://math.stackexchange.com/questions/349738/prove-rank-ata-rank-a-for-any-a-m-times-n

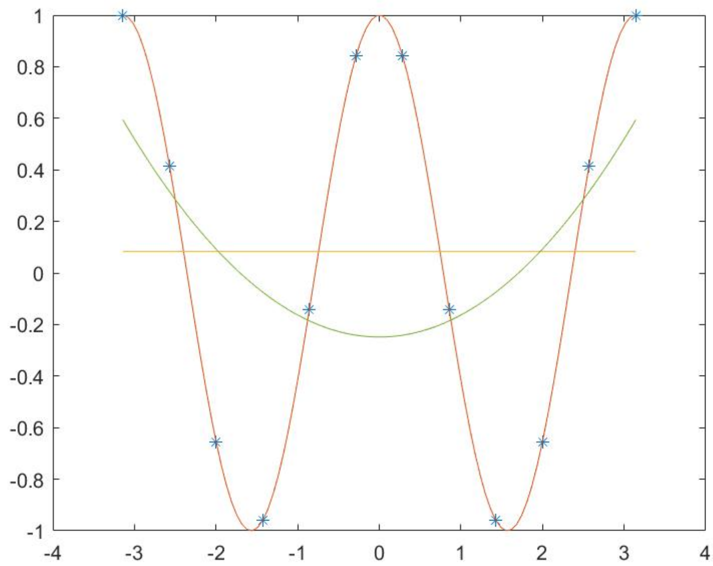

Figure 1: M=1 orange, M=2 and M=3 green, original function red

Case 2 $M + 1 > N$

Let Z_1 be the leftmost NxN part of $\mathbf{Z}$, so that the rows of Z_1 are $\hat{z}_1, \hat{z}_2, ..., \hat{z}_N$, where $\hat{z}_i = (1, x_i, x_i^2, ..., x_i^{N-1})$. Again Z_1 is an invertible *Vardenmonde matrix*. Thus $\mathbf{Z}$ is full column rank, thus t must be in the column space of $\mathbf{Z}$, which means the equation $\mathbf{Z}\mathbf{w}^T = t$ has at least one solution. By multiplying the previous equation with $\mathbf{Z}^T$ from left, we see that $\mathbf{w}$ is also a solution of the equation $A\mathbf{w}^T = b$. As $M + 1 > N$, there are more columns in $\mathbf{Z}$ then its rank, so the solution is not unique.

Notes on Task 4, 6 and 8

Note that the $\mathbf{w}$ are the same for $N = 2k$ and $N = 2k + 1$, so their graphics coincide. It appears that for the data set D_1, the higher values of M gives better approximation to function. On the other hand, for the data set D_2, the values $M = 7, 8$ and 9 seems to give better approximation to the function. As M gets closer to N, the approximate function passes thorough the data points, but is not the best fit for the original function.

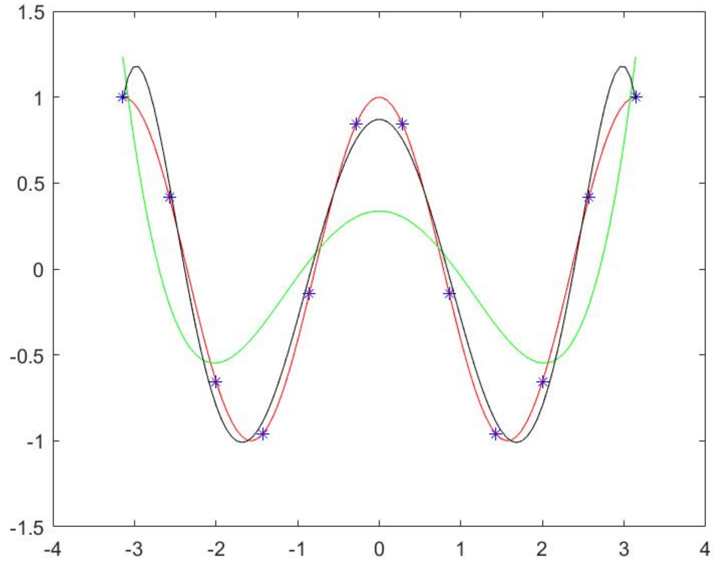

Figure 2: M=4,5 is green, M=6,7 is black, original function is red

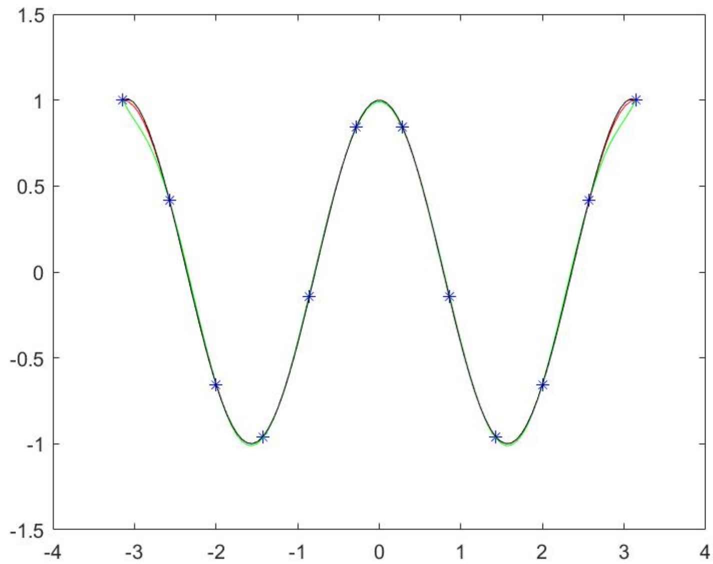

Figure 3: M=8,9 is green, M=10,11 is black, original function is red

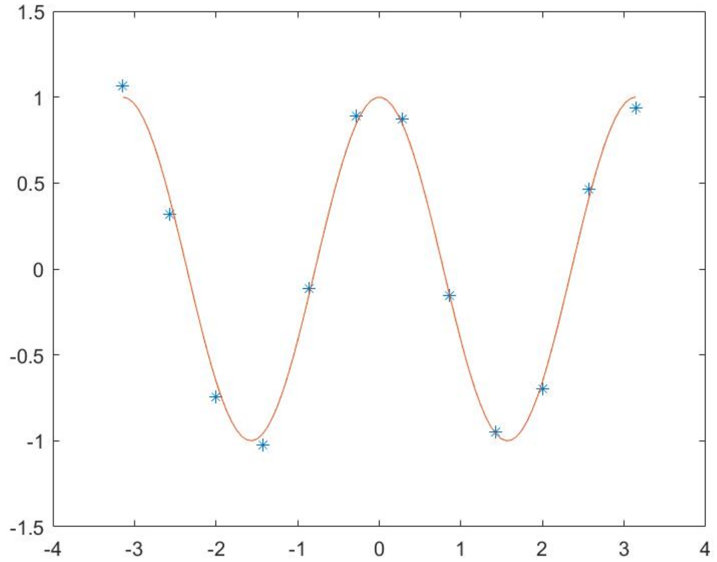

Figure 4: Task 7: New Data Set D_2

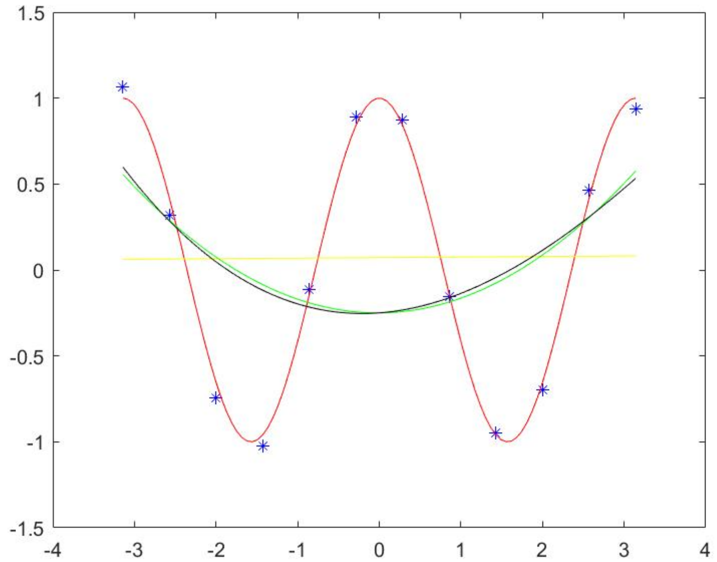

Figure 5: M=1 is yellow, M=2 is green, M=3 is black, original function is red

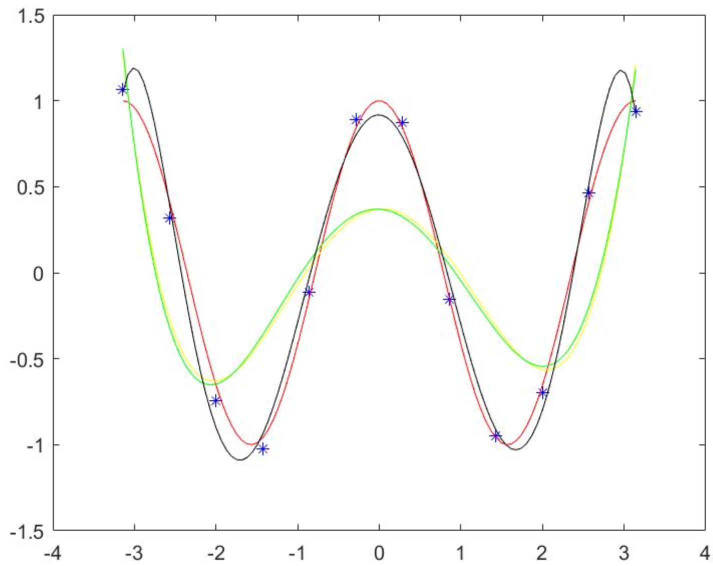

Figure 6: M=4 is yellow, M=5 is green, M=6 is black, original function is red

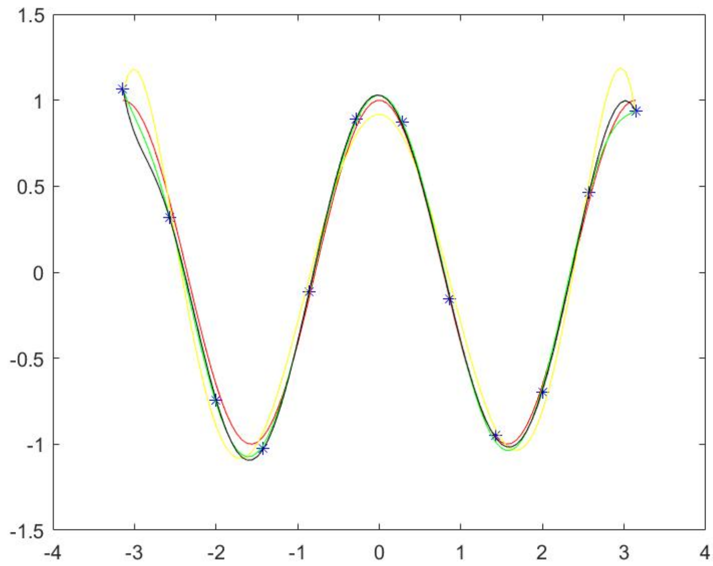

Figure 7: M=7 is yellow, M=8 is green, M=9 is black, original function is red

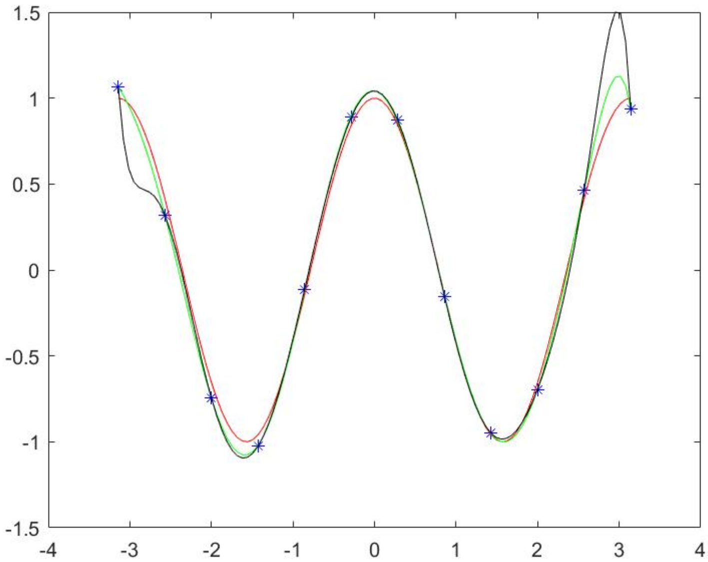

Figure 8: M=10 is green, M=11 is black, original function is red