Project SI

By Sima Razvan and Deniz Eker

Index: 24

Fitting an unknown function

- The main problem of this is how do you fit the provided data on a polynomial ?
- One could try to use x1 and x2 and use them to have phi =[phi_x1,phi_x2] but that is not a viable option as you cannot fit it on the polynomial, since both X's need to fit on the polynomial so that we can get the correct output.
- Then how?
- One solution is to think phi as a matrixes of vectors to calculate the correct output.

Fitting an unknown function using linear regression

IV steps to achieve this :

1. Defining variables and reshaping existing matrixes

2.Polynomial regression for data identification + optimal theta

3.Polynomial regression for validation data

4.Finding MSE and finding the optimal degree

Part I: Defining variables and reshaping existing matrixes

1. We'll have to introduce the given data into our code .
2. We will have to create a space or container for our values that will be calculated in the following parts.
3. We took a max degree and a starting degree as a sampling rate to deduce the most optimal degree.
4. As for the last step we will need to reshape the matrices of Y so
5. One solution could look like this :

```
load("proj_fit_24.mat");
n1=length(id.X{1});
idx1=id.X{1};
idx2=id.X{2};
valx1=val.X{1};
valx2=val.X{2};
idY=id.Y;
valY=val.Y;
degree=15;
mseid_array=zeros(degree, 1);
mseval_array=zeros(degree, 1);
idYhat=[]; valYhat=[];
currentdegree=1;
idY=reshape(idY, [], 1);
valY=reshape(valY,[],1);
idPHI=cell(degree, 1);
valPHI=cell(degree, 1);
theta_id=cell(degree, 1);
```

Part II: Polynomial regression for data identification + optimal theta

1. For the linear regression we would start from degree 1 and go till the max degree we gave , which in our case is 15.
2. We will go through all pairs of data points , and we will use an auxiliary “index” variable to store the current degree and this will be used to generate different powers of x2, and k will generate for x1.
3. This is important so that we can create the polynomial , for example if we had degree 1 we would have 3 terms[1,x1,x2] and for 3 we would have 10[1,x1,x2,…,x1*x2^2].
4. And we will compute the values of the polynomial term and add it to the line array , making a matrix of vectors.
5. One solution for linear regression for the id part could look like this:

```
for currentdegree=1:degree
idPHI{currentdegree}=[];
for i=1:n1
line=[];
for j=1:n1
aux=currentdegree;
for index=aux:-1:0
for k=0:index
line=[line;idx1(i)^k*idx2(j)^(index-k)];
end
end
end
idPHI{currentdegree}=[idPHI{currentdegree};
line];
end
totalcoeffs=((currentdegree+1)*(currentdegree+2))
/2;
idPHI{currentdegree}=reshape(idPHI{currentdegr
ee}, totalcoeffs, []);
idPHI{currentdegree}=idPHI{currentdegree}';
theta_id{currentdegree}=idPHI{currentdegree}\id
```

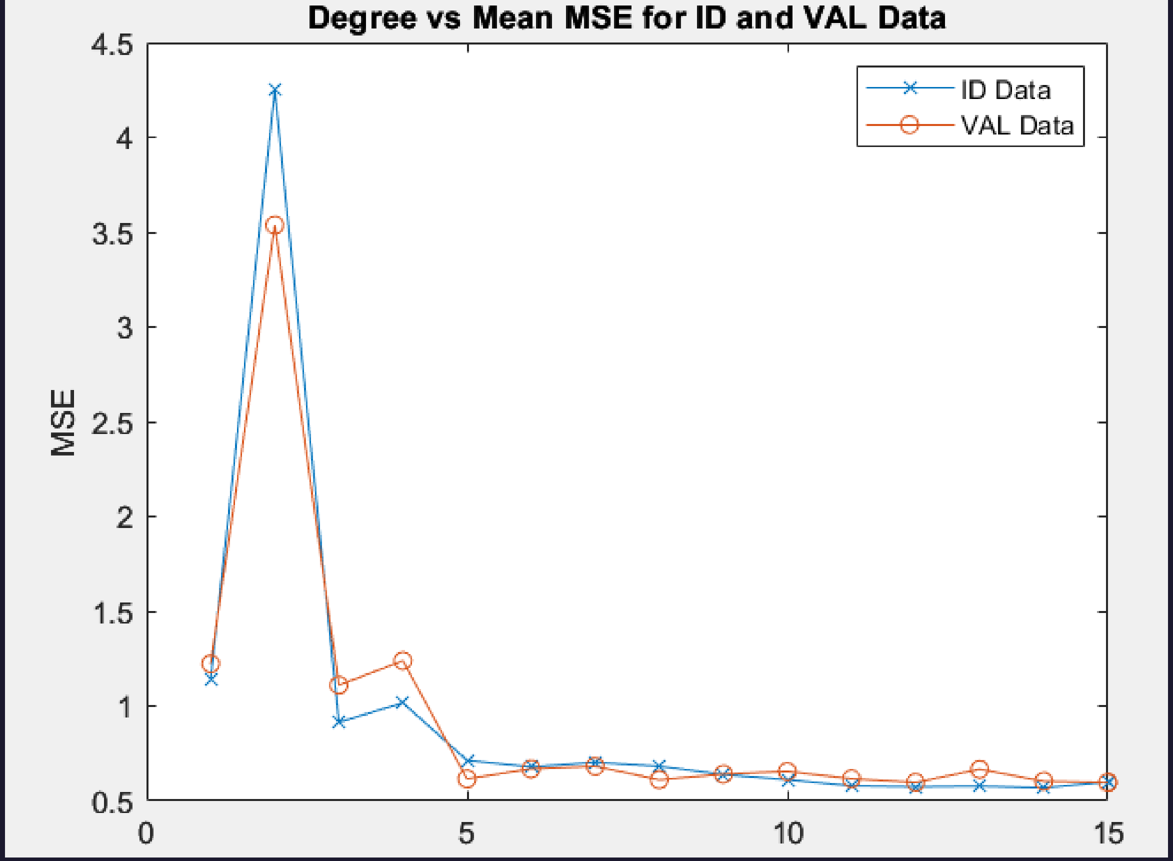
Part III: Polynomial regression for validation data

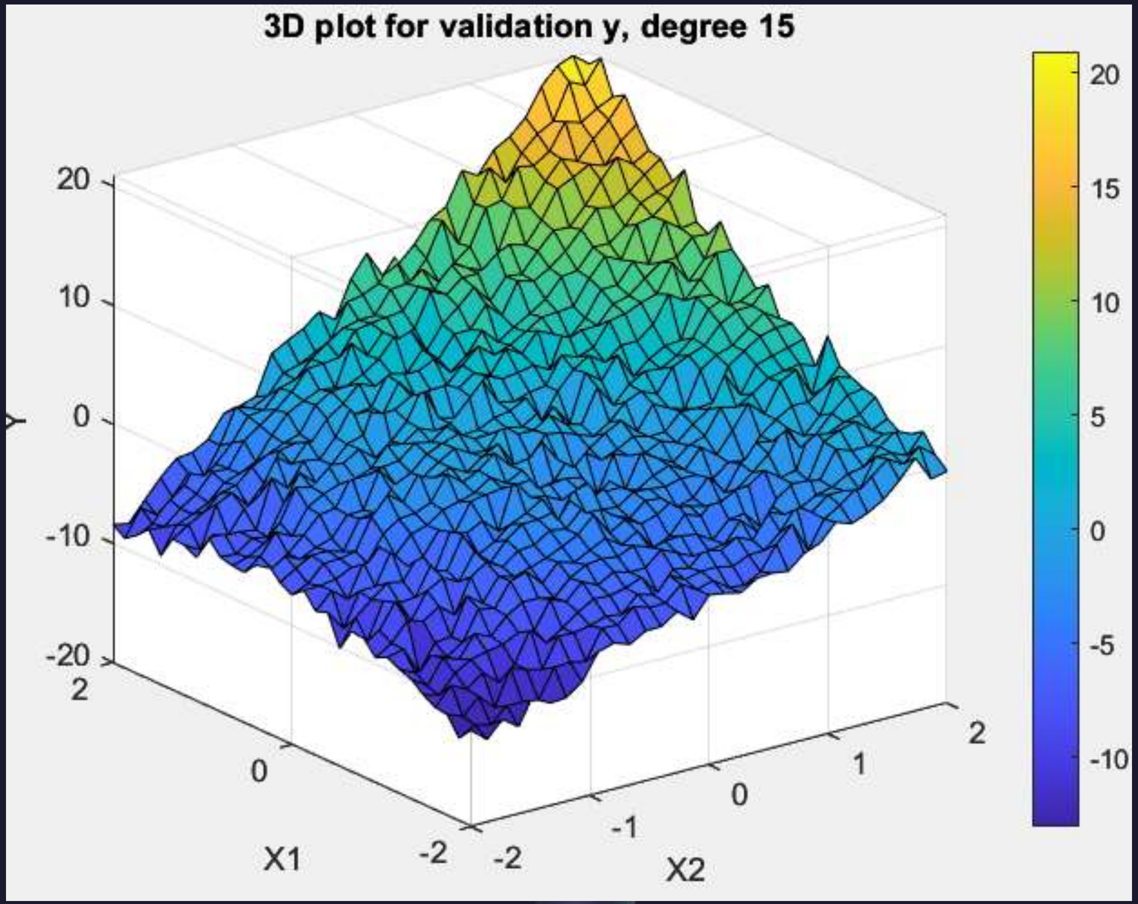
1. We will proceed to do the same thing as the identification however we will use the theta we calculated for identification.
2. And so we will have a similar solution like the identification part :

```
for currentdegree=1:degree
valPHI{currentdegree}=[];
for i=1:31
line=[];
for j=1:31
aux=currentdegree; for index=aux:-1:0
for k=0:index
line=[line; valx1(i)^k*valx2(j)^(index-k)];
end
end
end
valPHI{currentdegree}=[valPHI{currentdegree};
line];
end
totalcoeffs=((currentdegree+1)*(currentdegree+2)
)/2;
valPHI{currentdegree}=reshape(valPHI{currentde
gree}, totalcoeffs, []);
valPHI{currentdegree}=valPHI{currentdegree}';
```

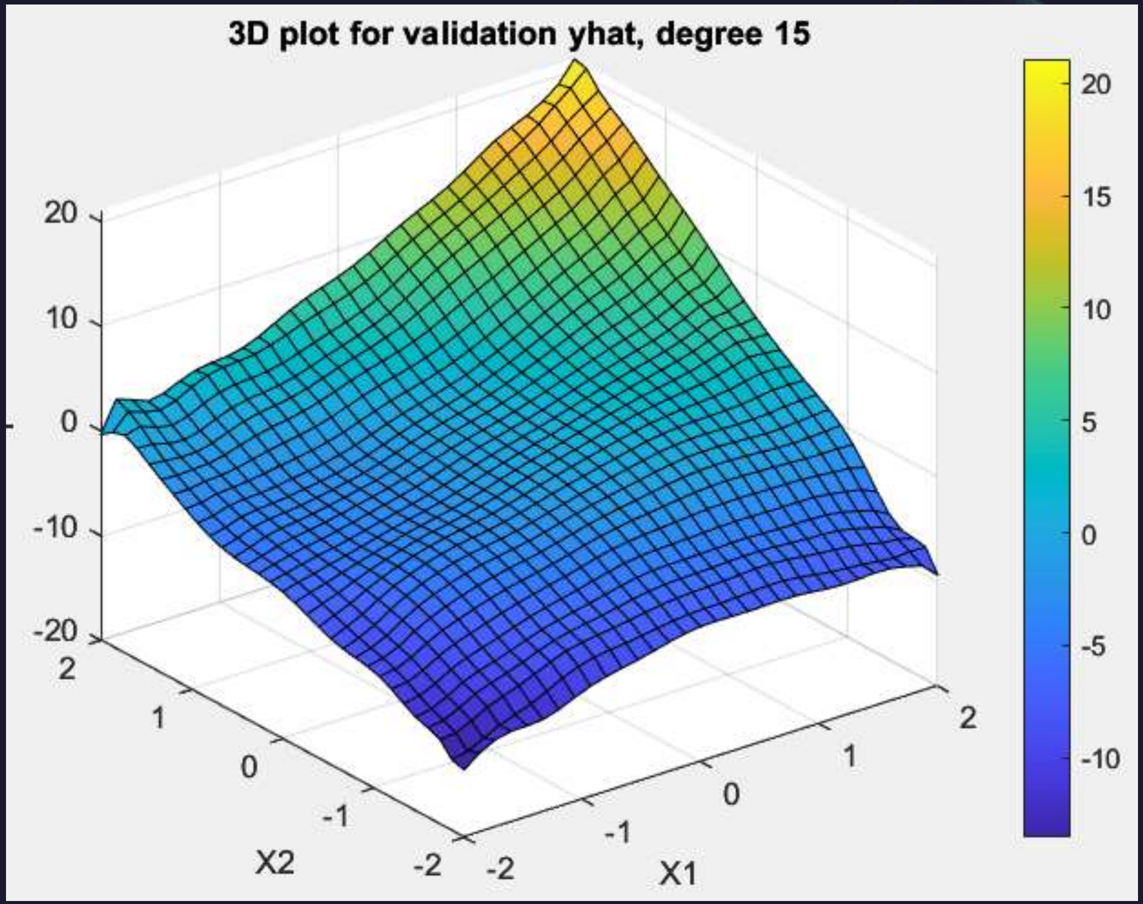
Part IV: Finding MSE and finding the optimal degree

1. **In order to find the most optimal degree we will need to look for the lowest value of MSE.**
2. **In our case we found the most optimal degree at 15.**
3. **After finding the optimal degree we will also plot the Yhat and Y.**





The original output



The approximated output

Conclusion:

- Overall ,the purpose of this project targets creating a polynomial in order to evaluate its performance and choose the most suitable degree for the polynomial in order to combat disruptions from noises.

Team Bober

Deniz Eker

Sigma Razvan

Thank You