ELL 409

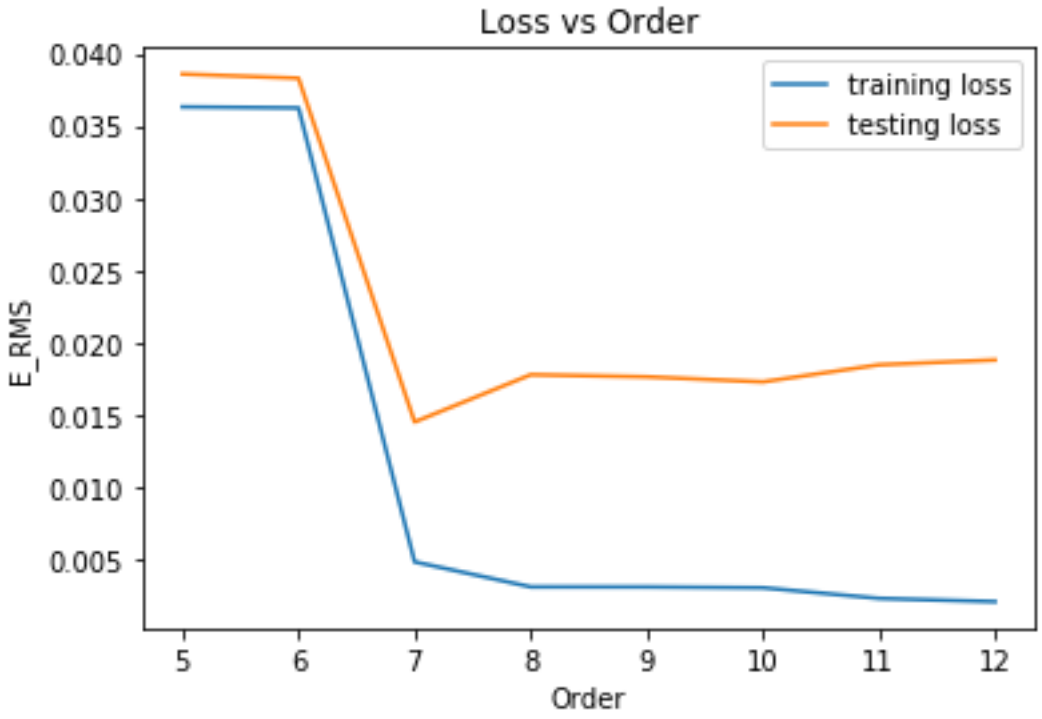
Assignment 1

Part 1(a)

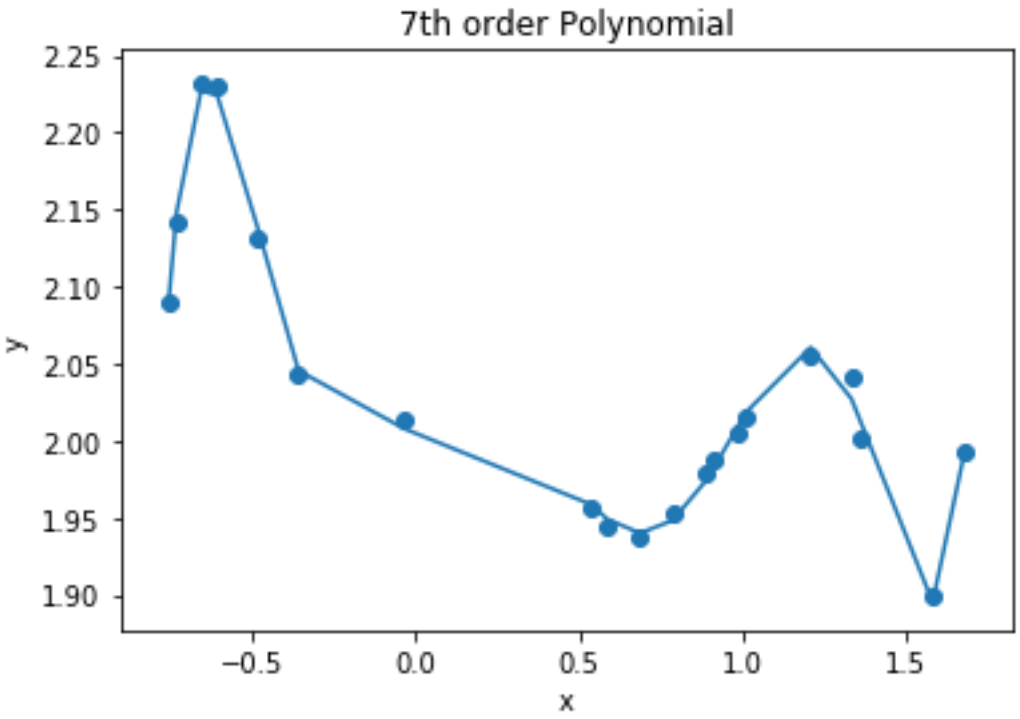
20 data points

1. Least-square error function

Moore-Penrose pseudo inverse



By analysing above graph we can conclude that order 7 will be perfect choice for given data set. As, after that, it started over fitting.

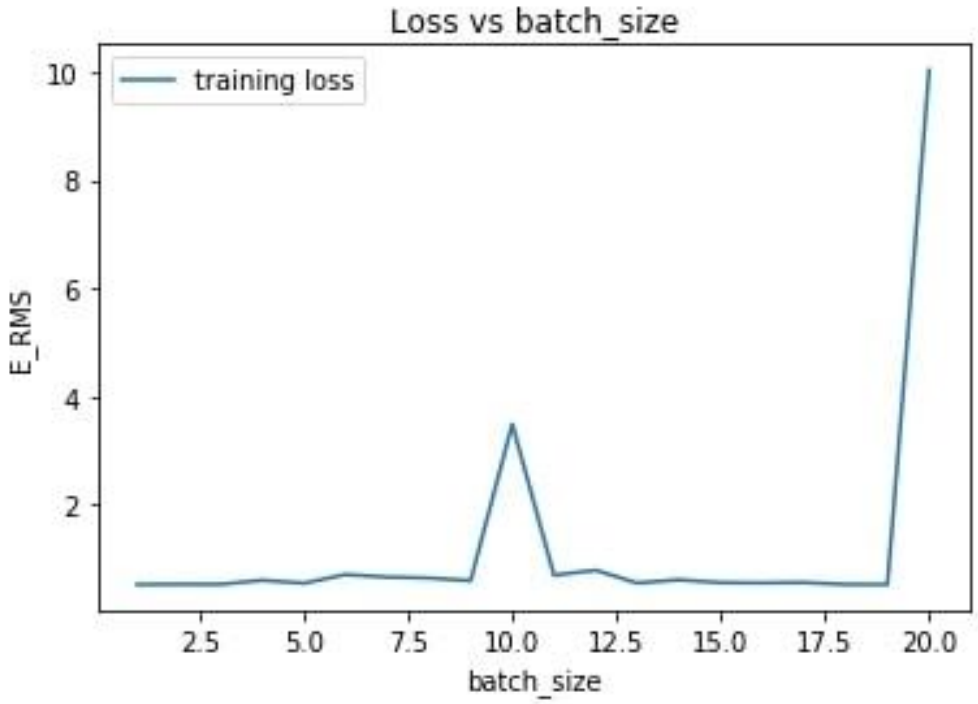


Root mean square error = 0.0059

Underlying guessed polynomial :

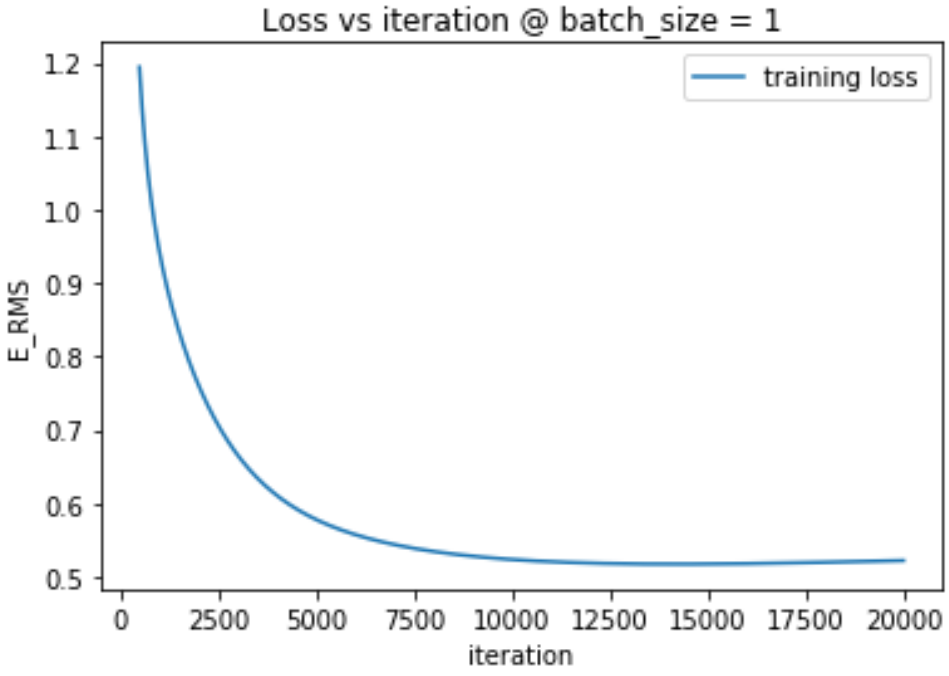
$2.01315277 + 0.15147253\ X - 0.11893576\ X^2 - 1.92457524\ X^3 + 2.00873072\ X^4 + 2.06697449\ X^5 - 3.13578898\ X^6 + 0.95378557\ X^7$

Gradient descent

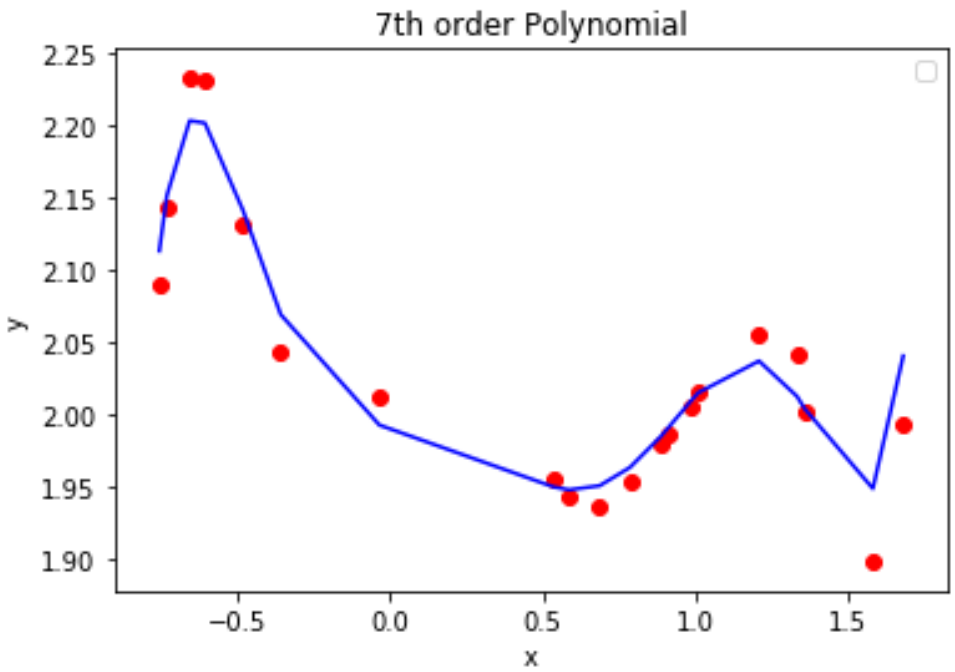


lr = 0.0005

Batch size selection for small data set has not any extra effect.



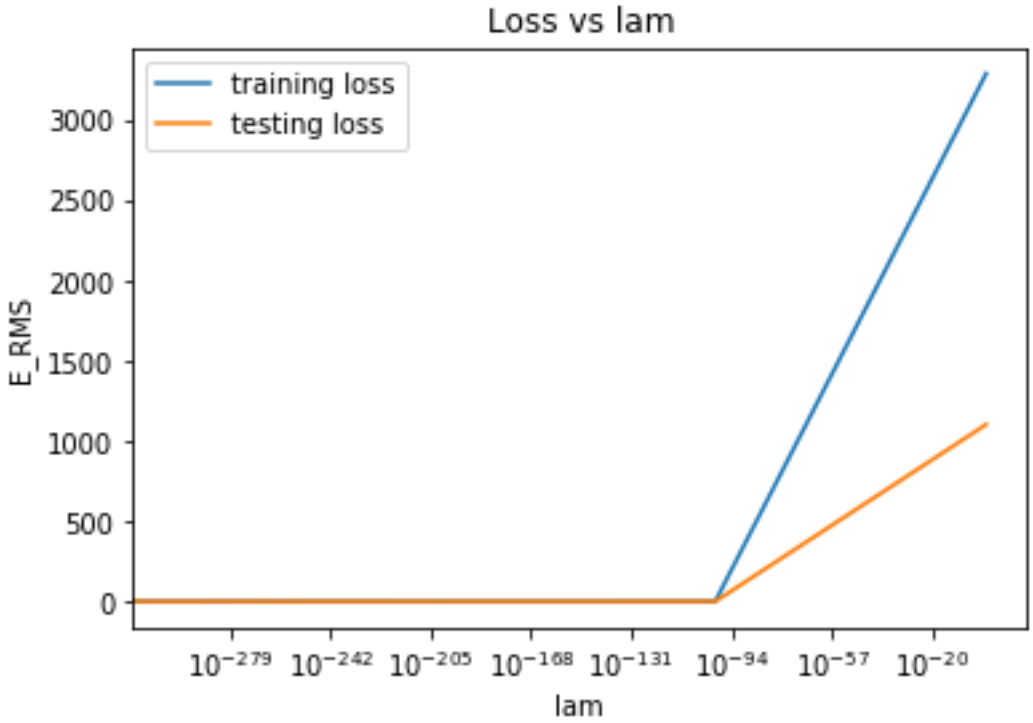
With iteration loss is decreasing, it shows SGD cause convergence.



Root mean square error = 0.522

Underlying guessed polynomial :

1.99132929 -0.03630172 X 0.11420844 X^2 -0.99157933 X^3 +0.97156647 X^4 + 1.21337188 X^5 -1.81259696 X^6 + 0.56295446X^7



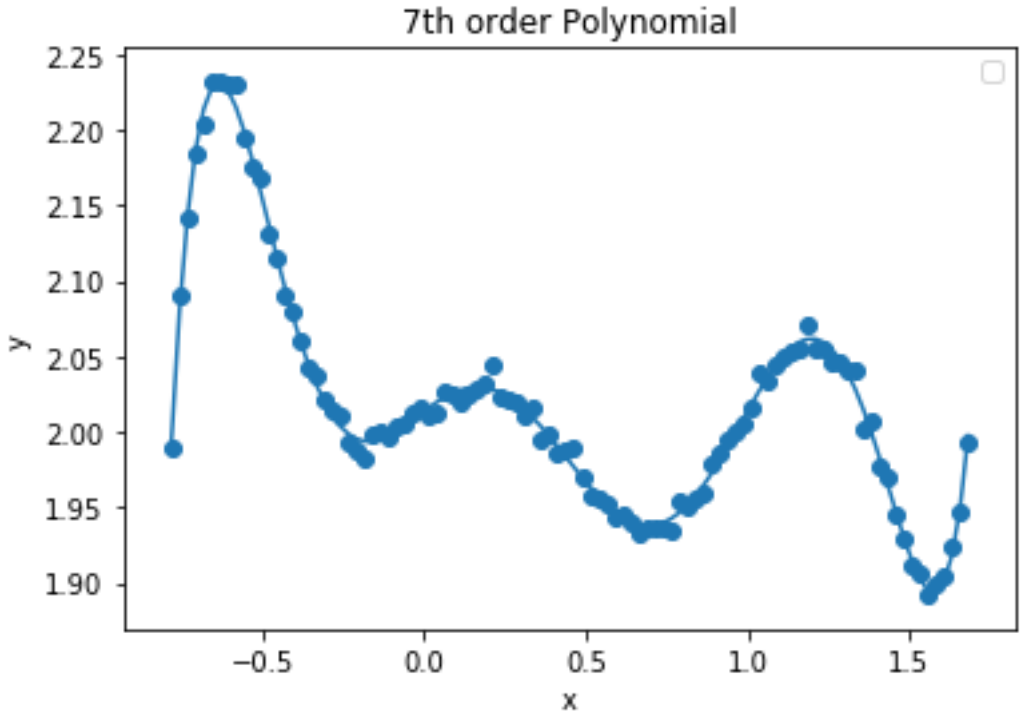
100 Best estimate for λ is 10^{-100} , beyond it, model started to under fitting.

101 data points

1. Least-square error function

Moore-Penrose pseudo inverse

Here again best estimate for order is 7 .

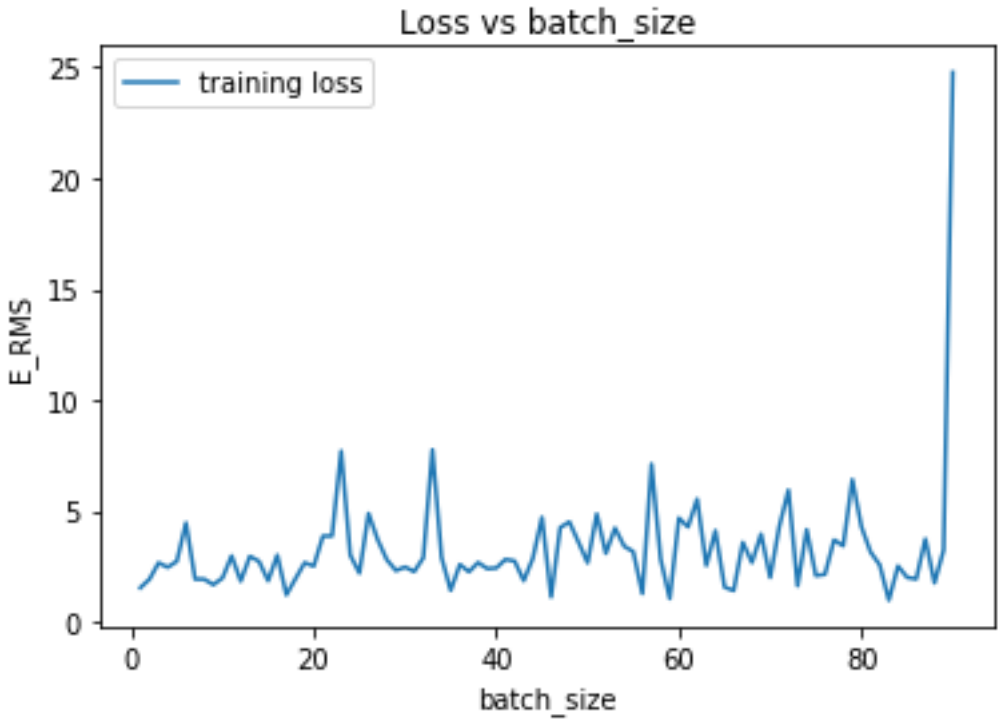


Root mean square error = 0.00587

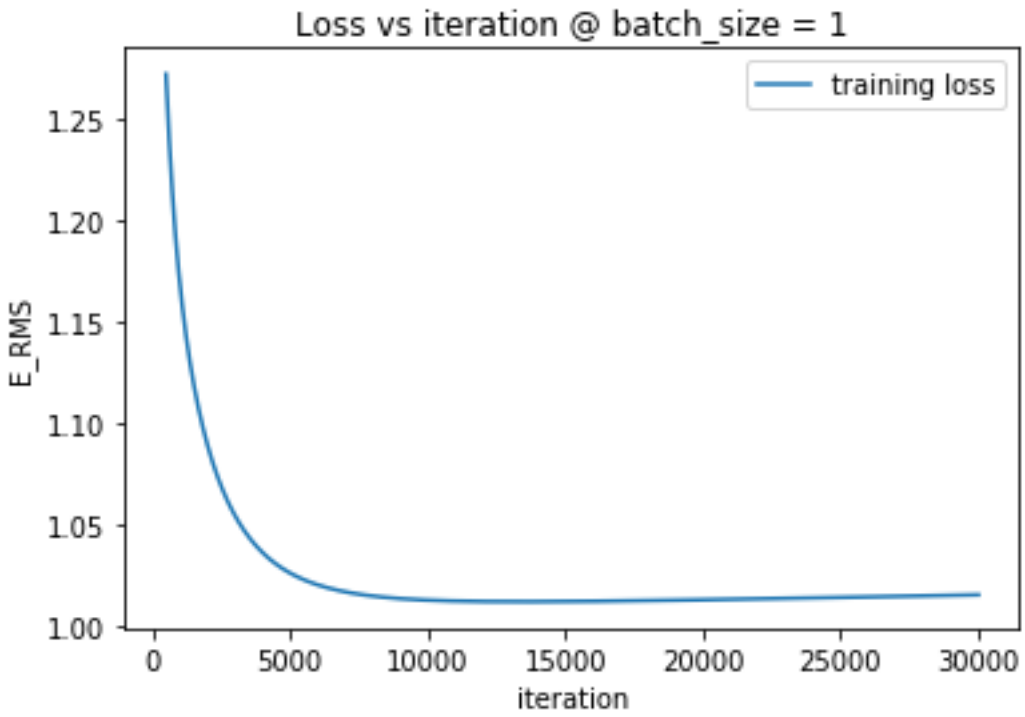
Underlying guessed polynomial :

$2.01363553 - 0.16161487X - 0.12645421X^2 - 1.99733251X^3 + 2.07061393X^4 + 2.15673593\,X^5 - 3.25234915X^6 + 0.98762133\,X^7$

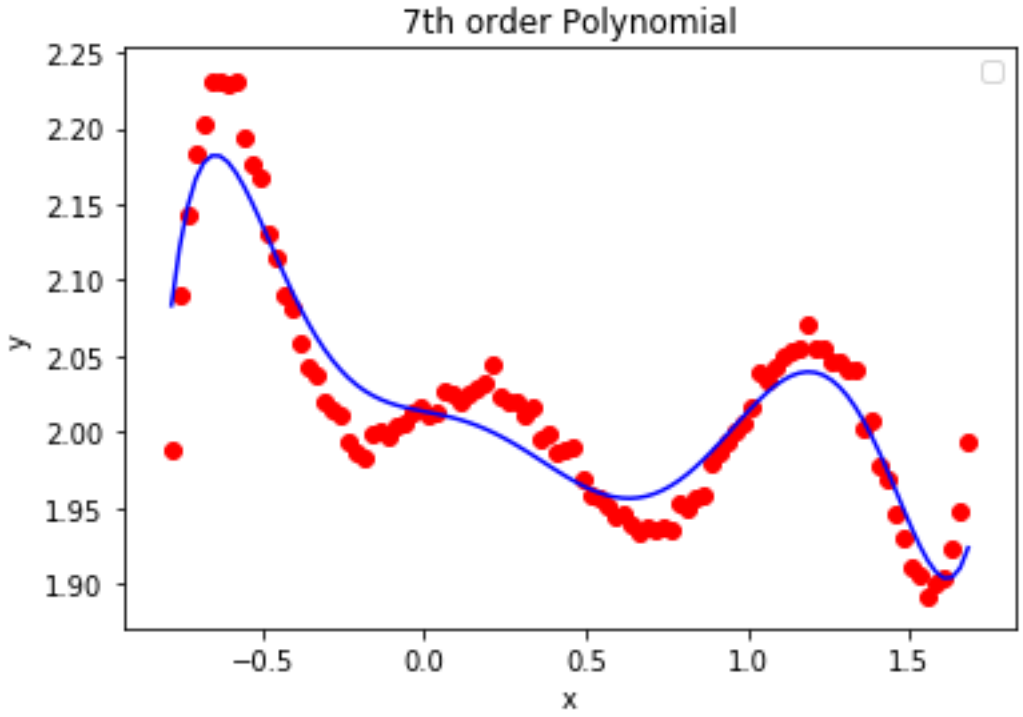
Gradient Descent



For 100 data set also batch size selection doesn't show extra effect but batch size = 1 (SGD) give best result .



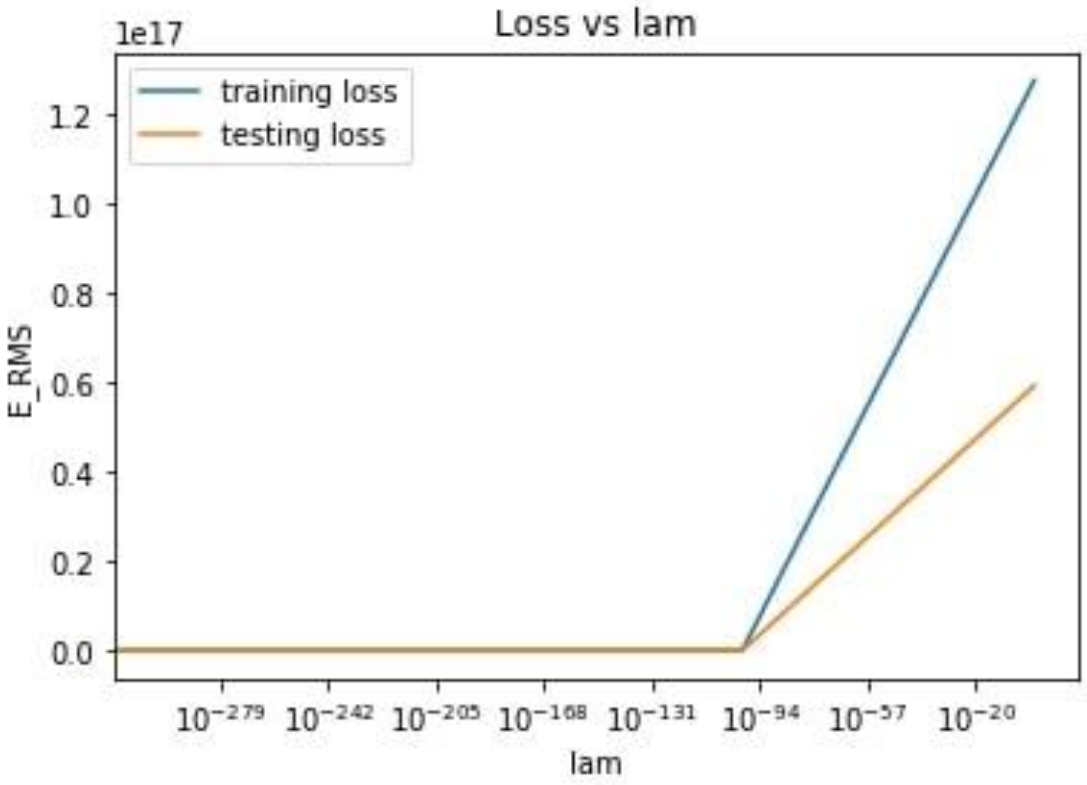
With iteration loss is decreasing, it shows SGD cause convergence.



Root mean square error = 1.015

Underlying guessed polynomial :

2.0133439 -0.0418903 X +0.0104605 X^2 -0.78206063 X^3 + 0.90955236 X^4 + 0.91907707 X^5 -1.46908119 X^6 + 0.45386308X^7



Best estimate for λ is 10^{-100} , beyond it, model started to under fitting.

Final estimate of underlying polynomial:

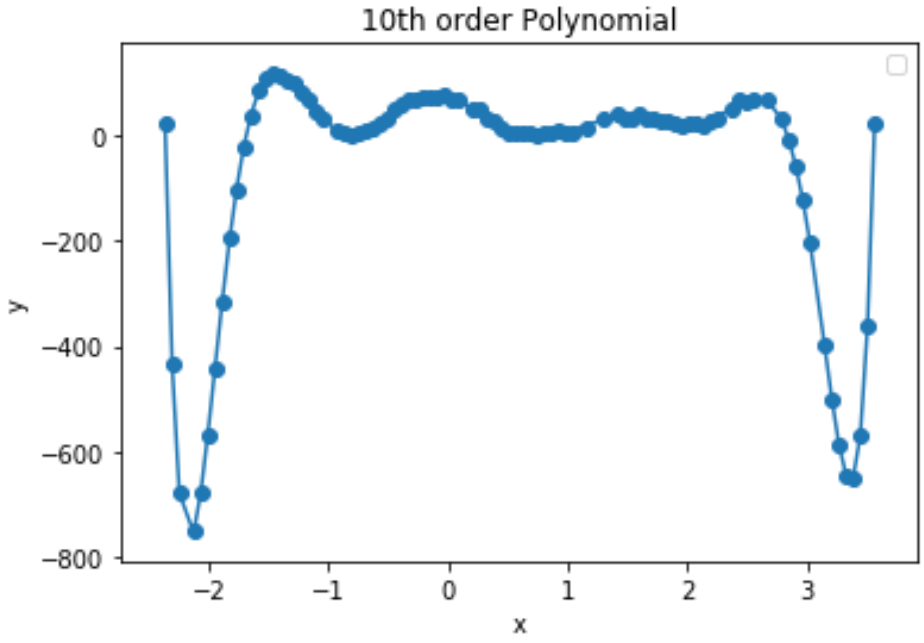
2.01363553 -0.16161487X -0.12645421X^2 -1.99733251X^3 +2.07061393X^4 + 2.15673593 X^5 -3.25234915X^6 + 0.98762133 X^7

Part 1(b)

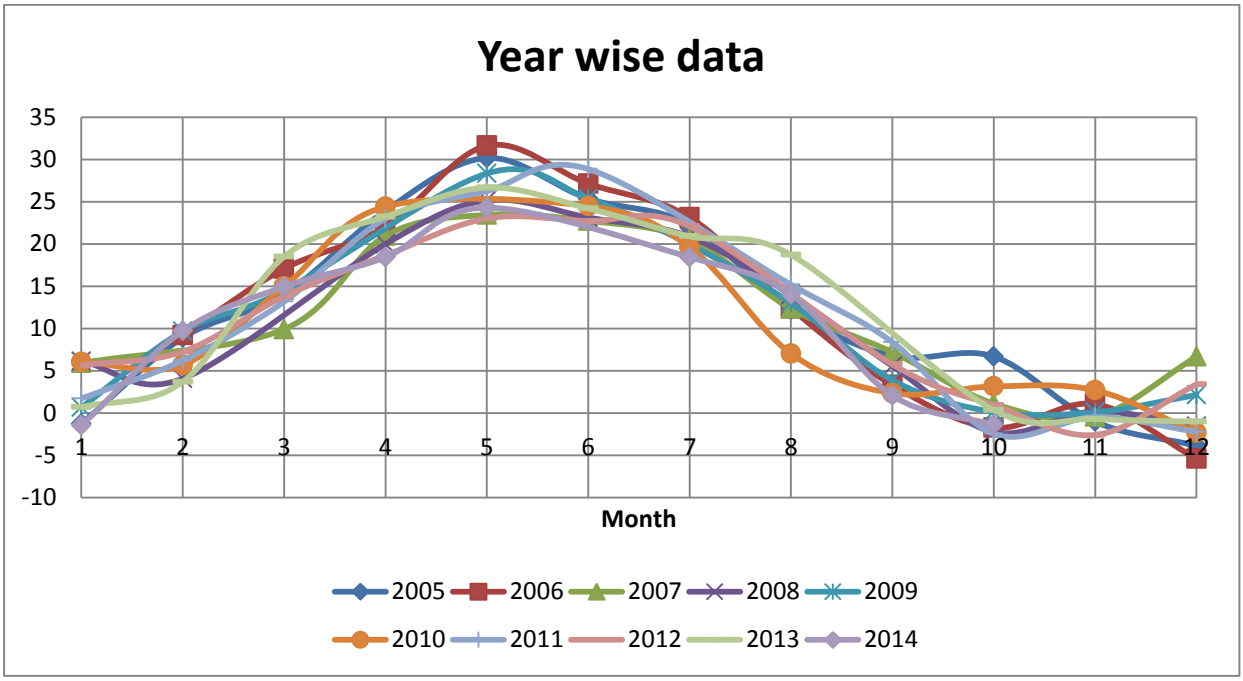
Moore-Penrose pseudo inverse

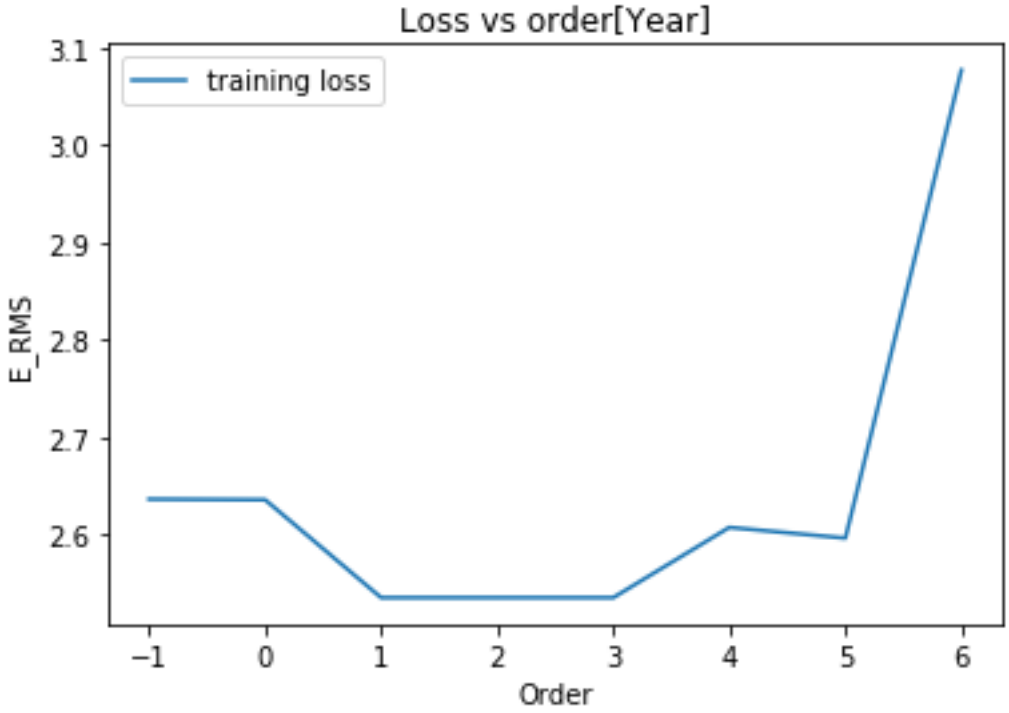
By analysing above graph we can conclude that order 10 will be perfect choice for given data set. As, after that, it started over fitting.

Variance of noise is 3.527690



Part 2


Year wise data
35
30
25
20
15
10
5
0
-5
-10
1
2
3
4
5
6
7
8
9
10
11
12
Month
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014


Loss vs order[Year]
training loss
E_RMS
3.1
3.0
2.9
2.8
2.7
2.6
−1
0
1
2
3
4
5
6
Order


Loss vs order[Month] @ order[year] = 3
training loss
testing loss
E_RMS
4.00
3.75
3.50
3.25
3.00
2.75
2.50
3
4
5
6
7
8
9
10
11
Order