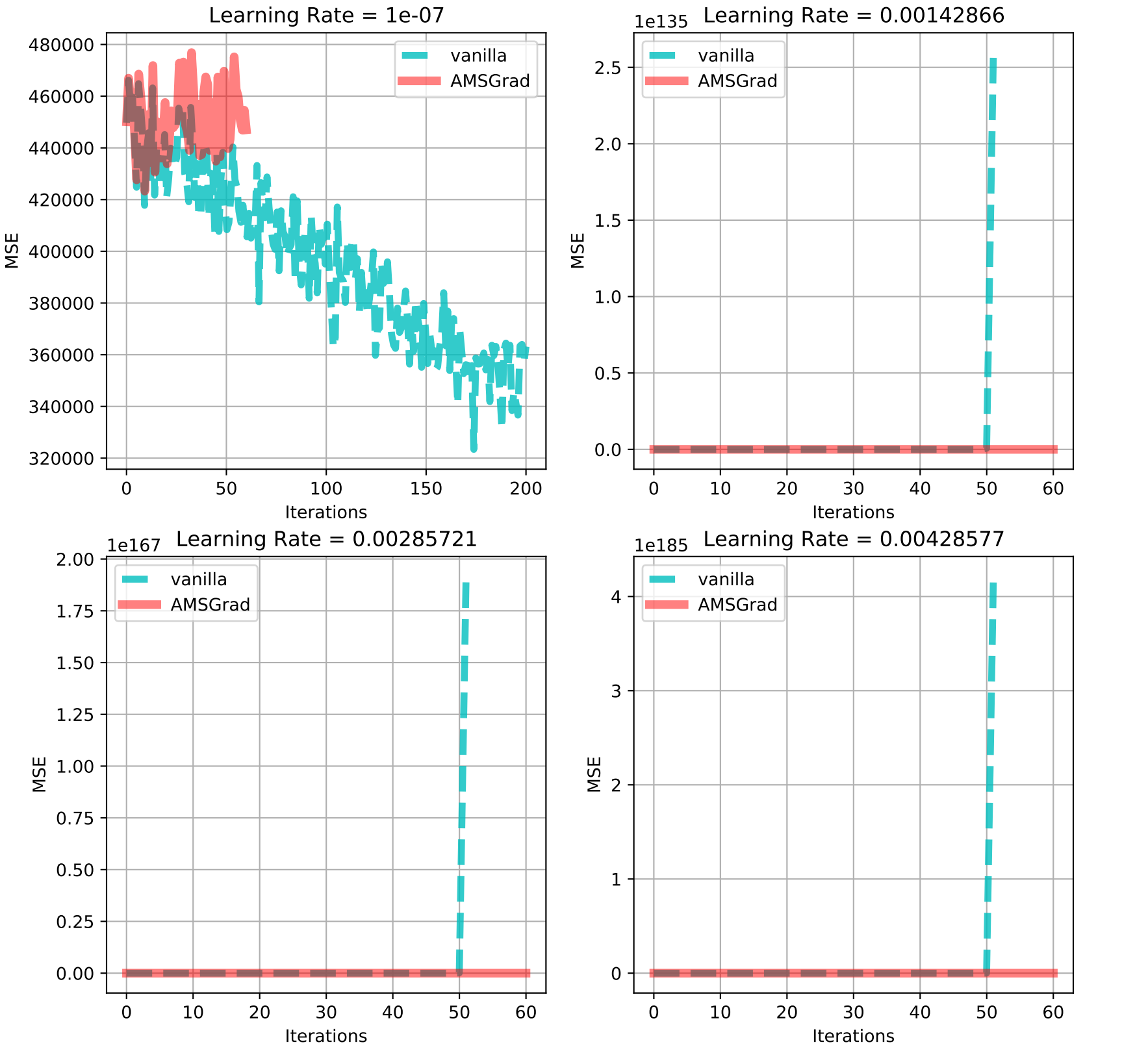
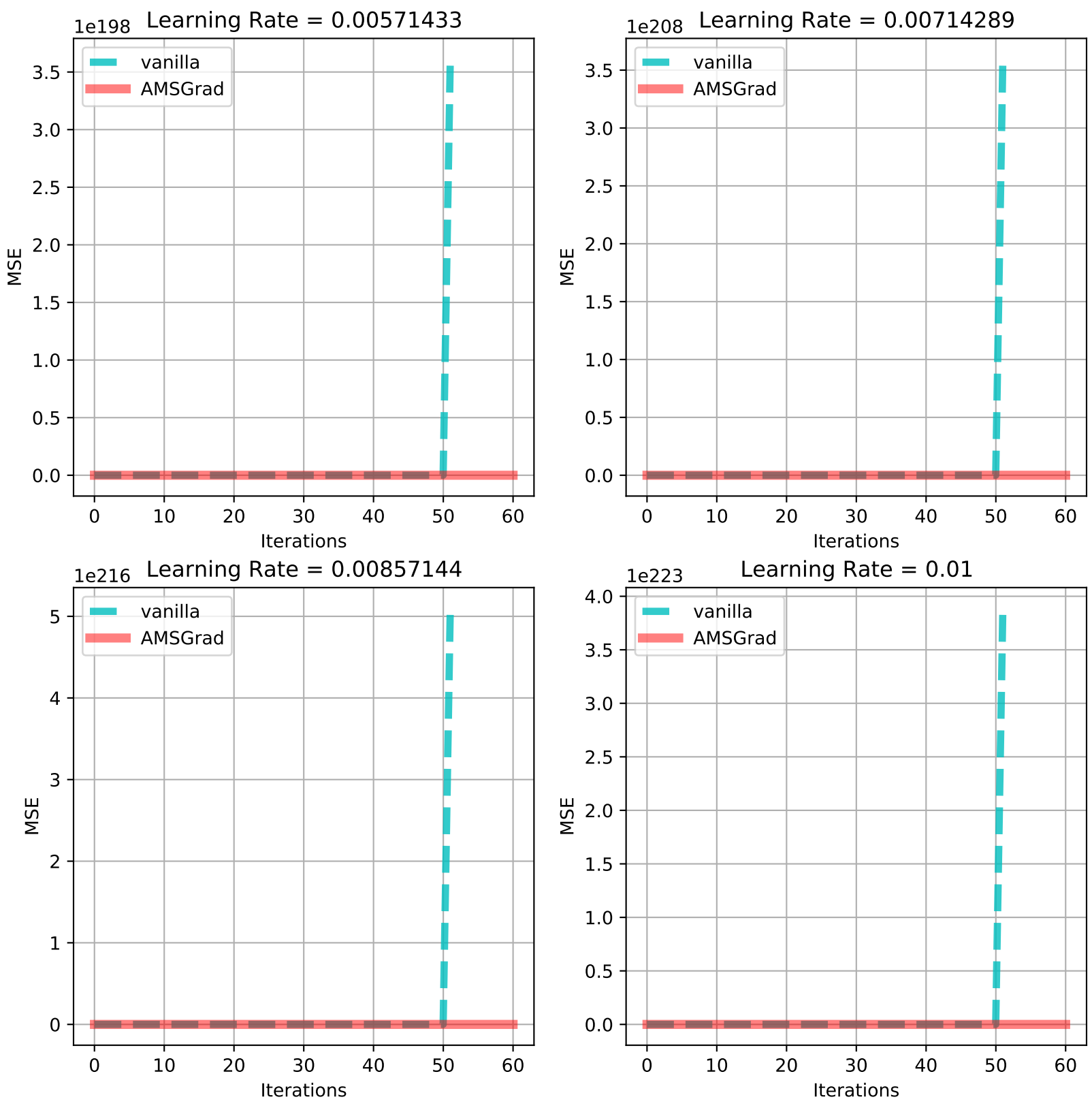
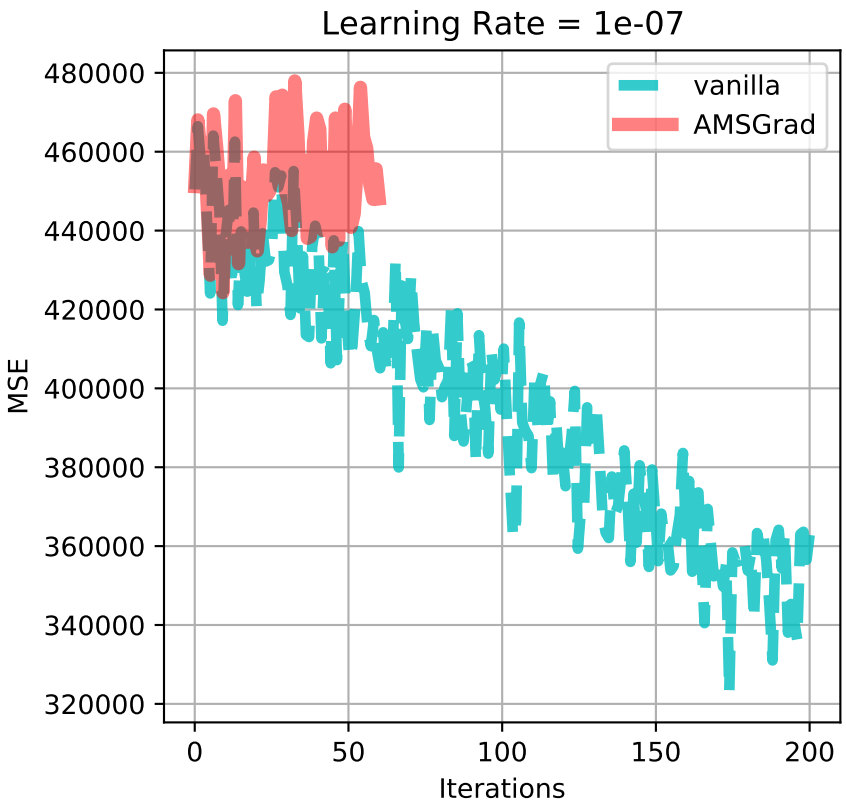
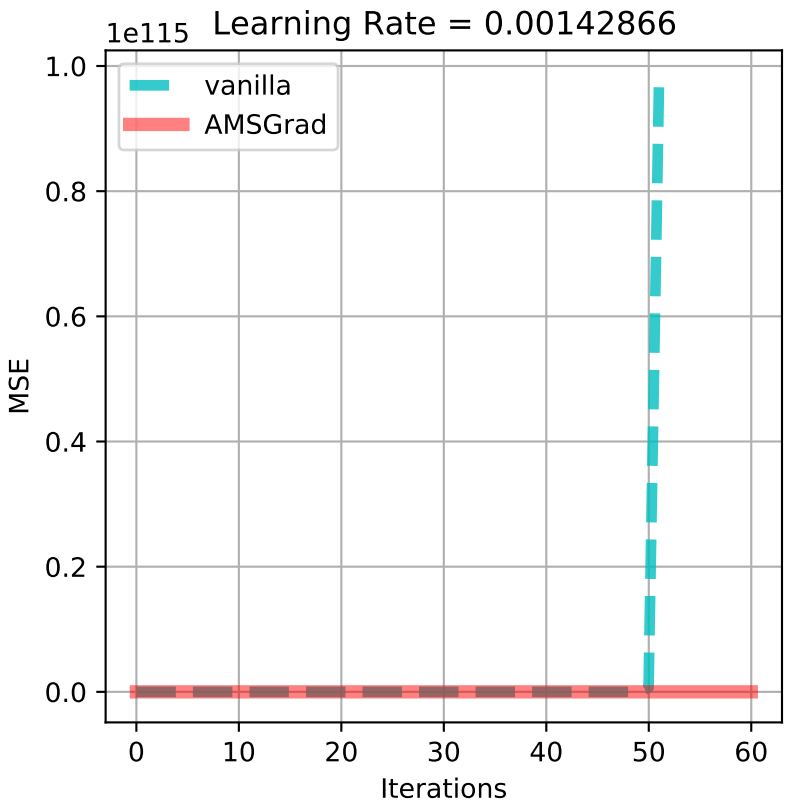
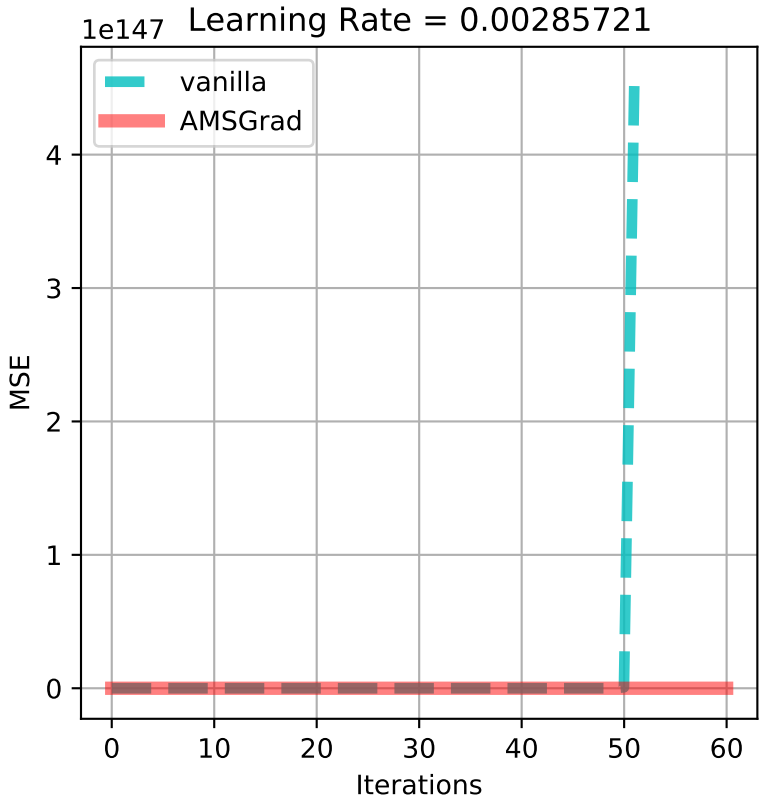

Historical MSE vs Iterations, Activation sigmoid
Learning Rate = 1e-07
Learning Rate = 0.00142866
Learning Rate = 0.00285721
Learning Rate = 0.00428577
vanilla
AMSGrad
MSE
Iterations
1e135
1e167
1e185

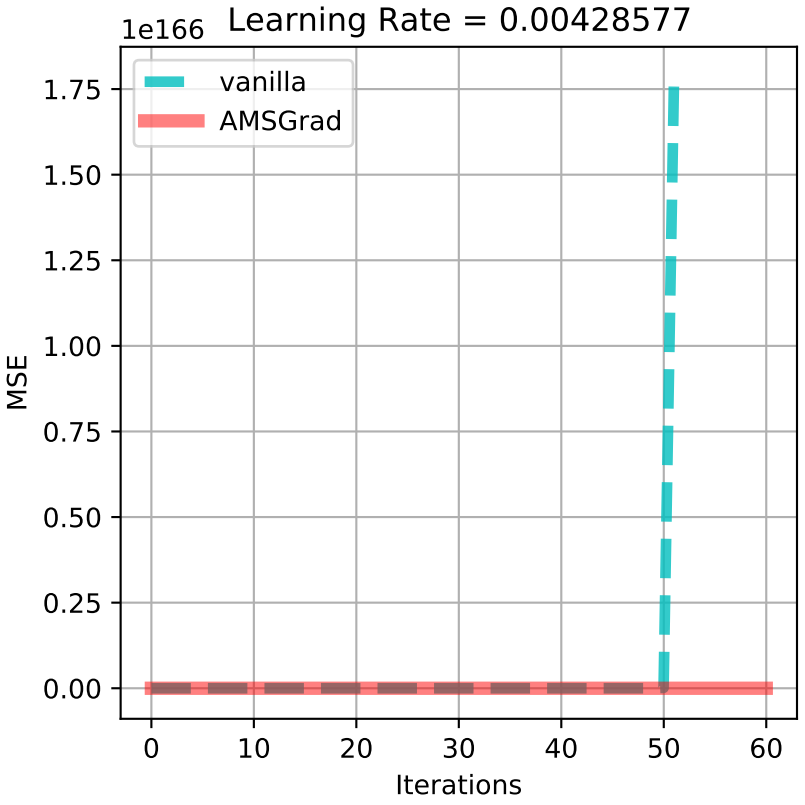

Historical MSE vs Iterations, Activation sigmoid
Learning Rate = 0.00571433
1e198
Learning Rate = 0.00714289
1e208
Learning Rate = 0.00857144
1e216
Learning Rate = 0.01
1e223
vanilla
AMSGrad
MSE
Iterations

Historical MSE vs Iterations, Activation tanh

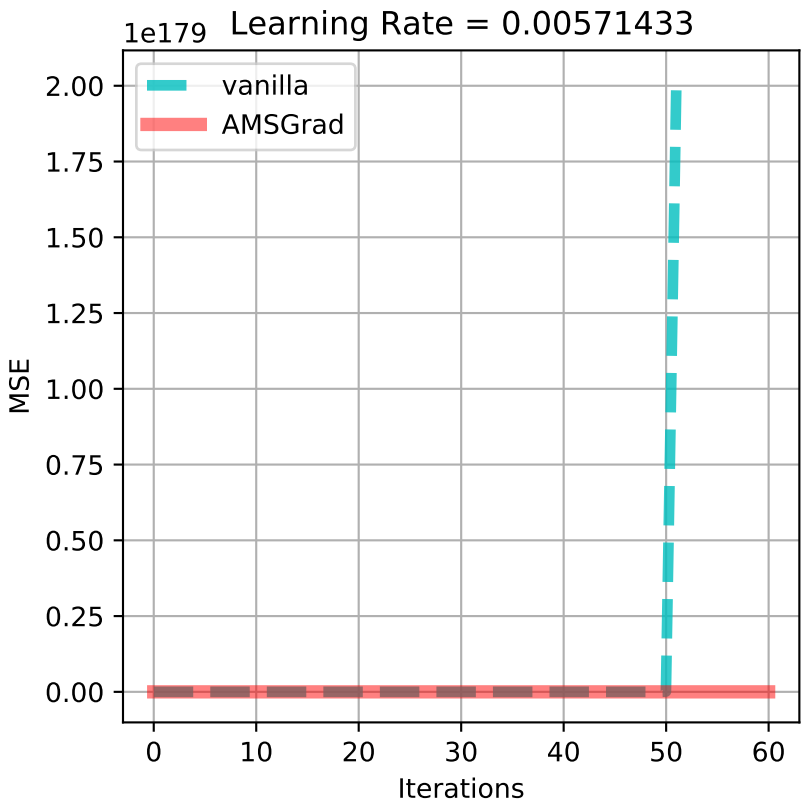


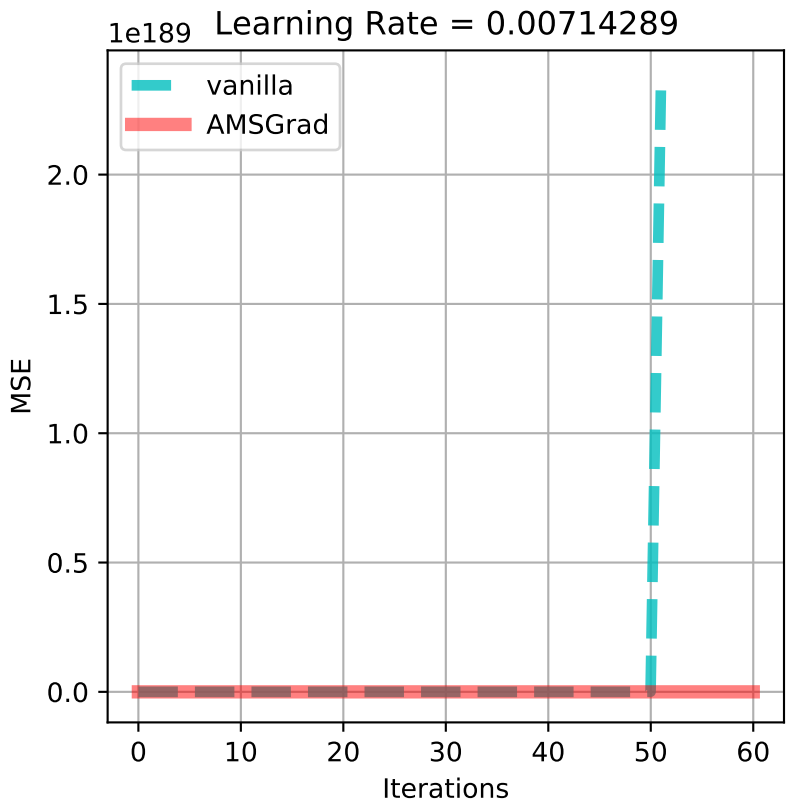


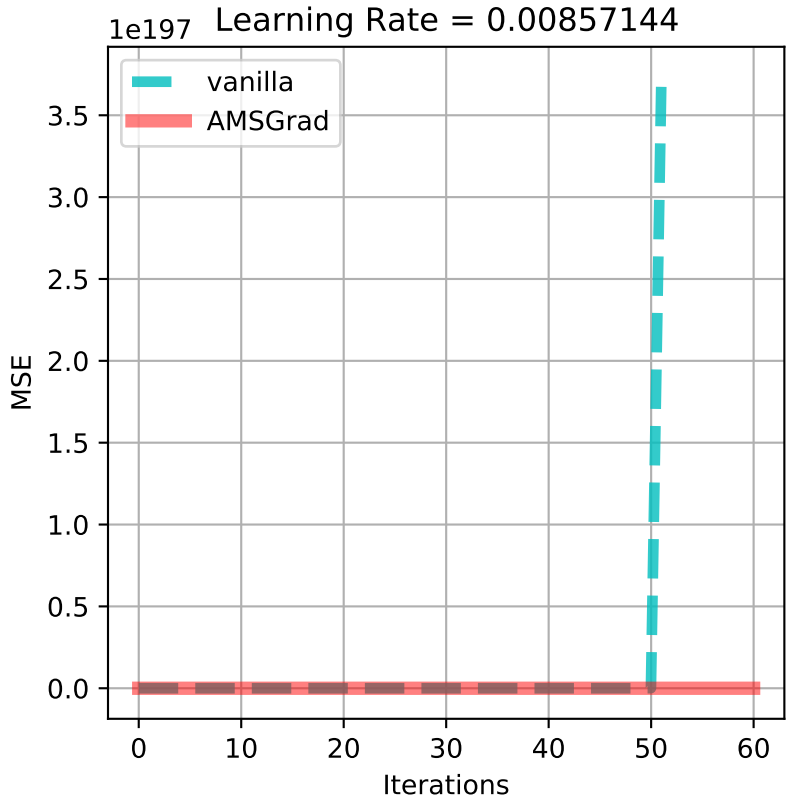


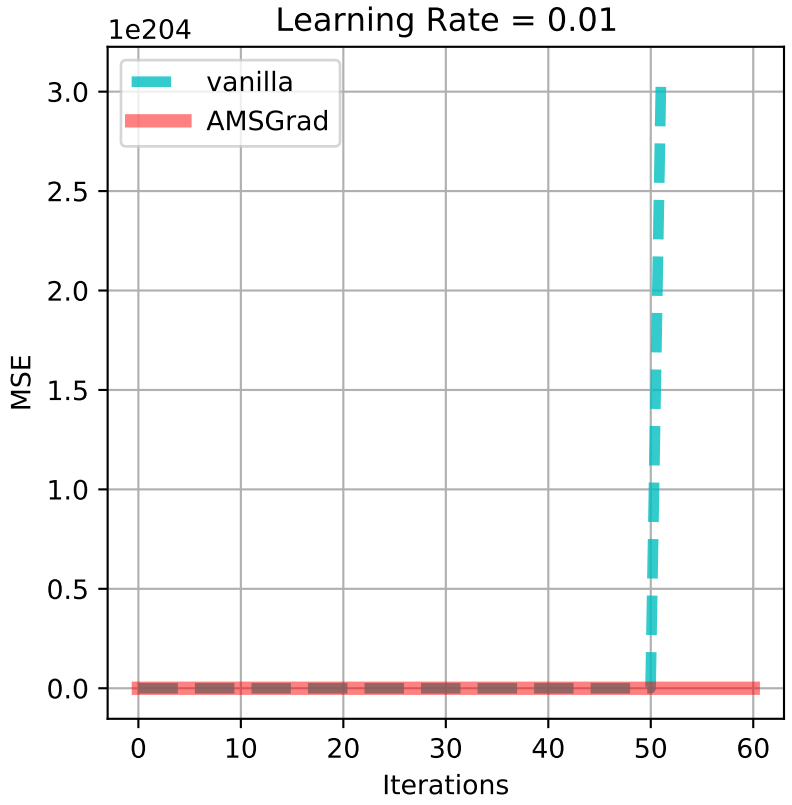


Historical MSE vs Iterations, Activation tanh

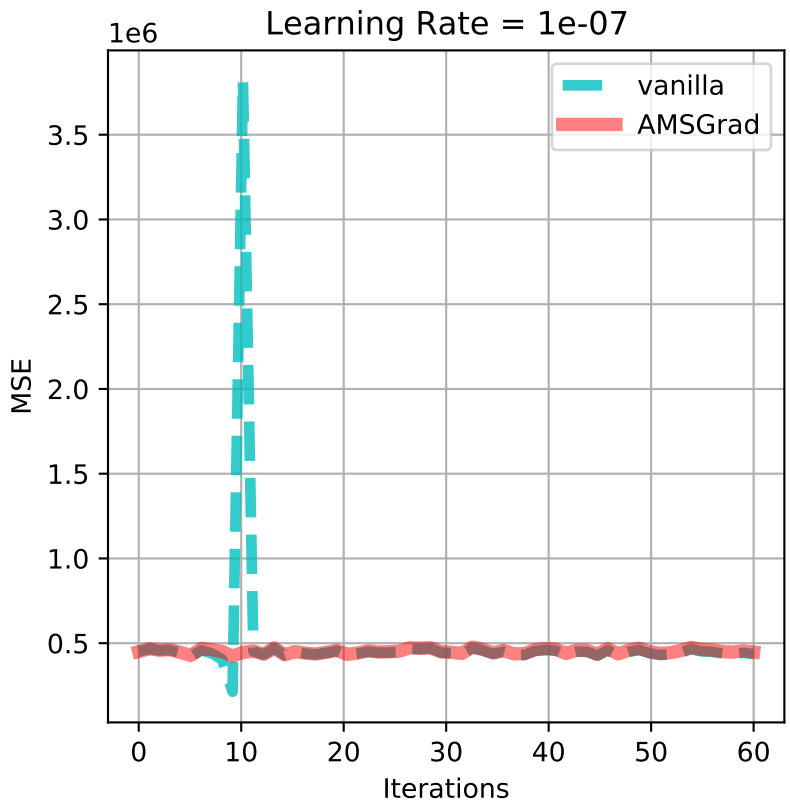


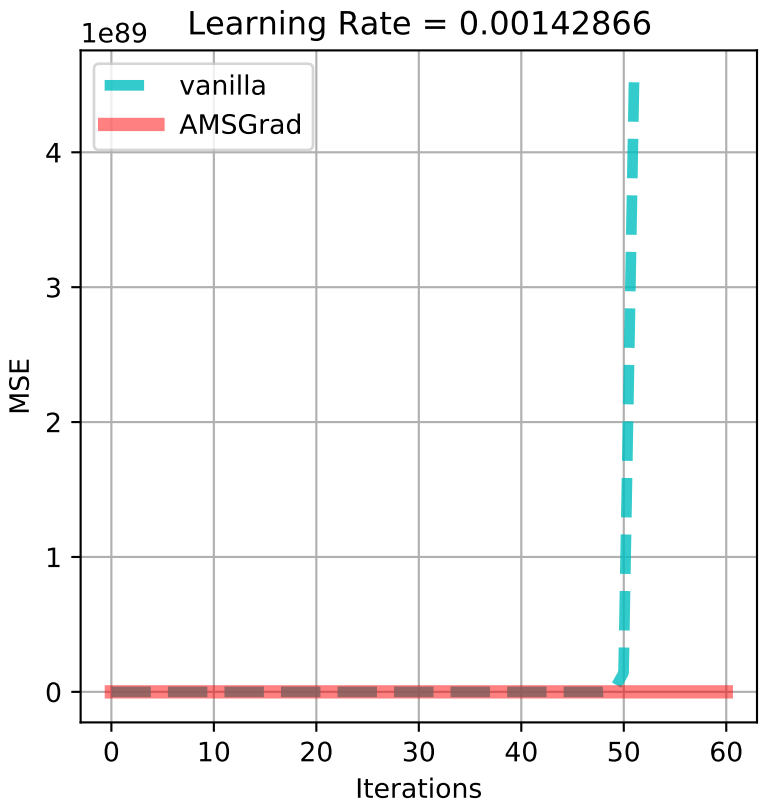


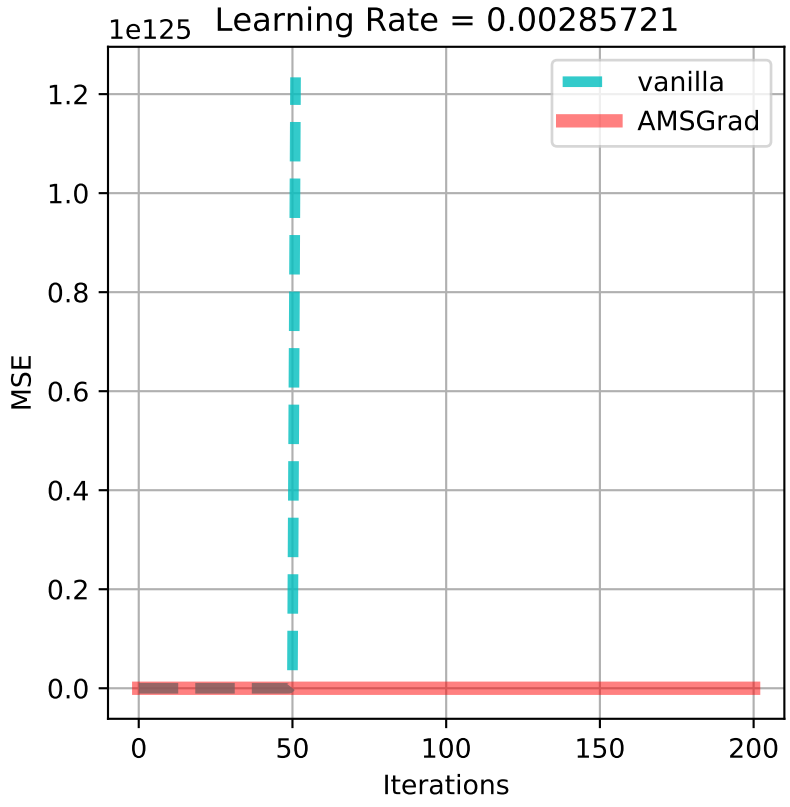


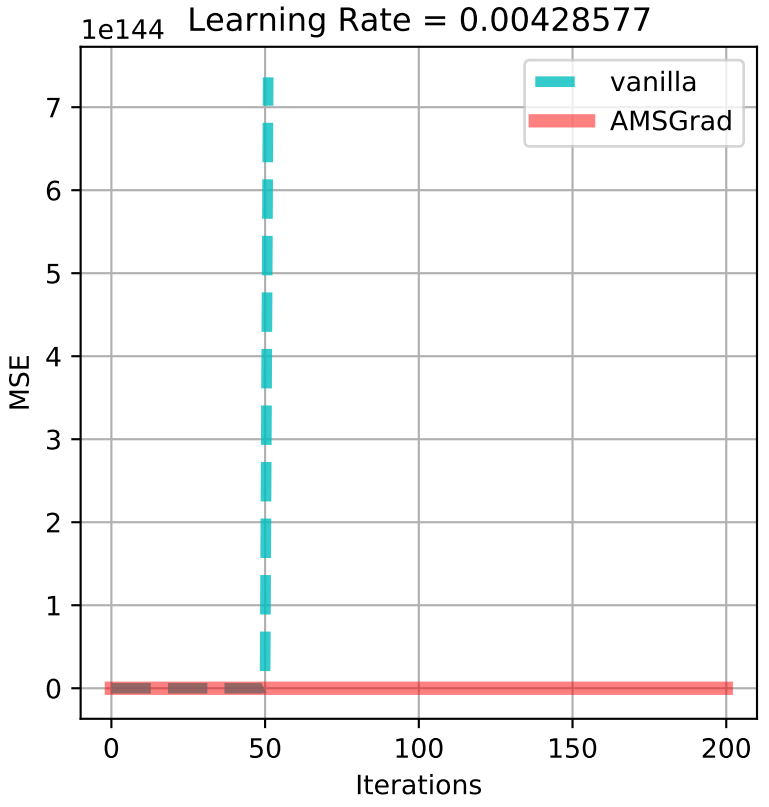


Historical MSE vs Iterations, Activation relu

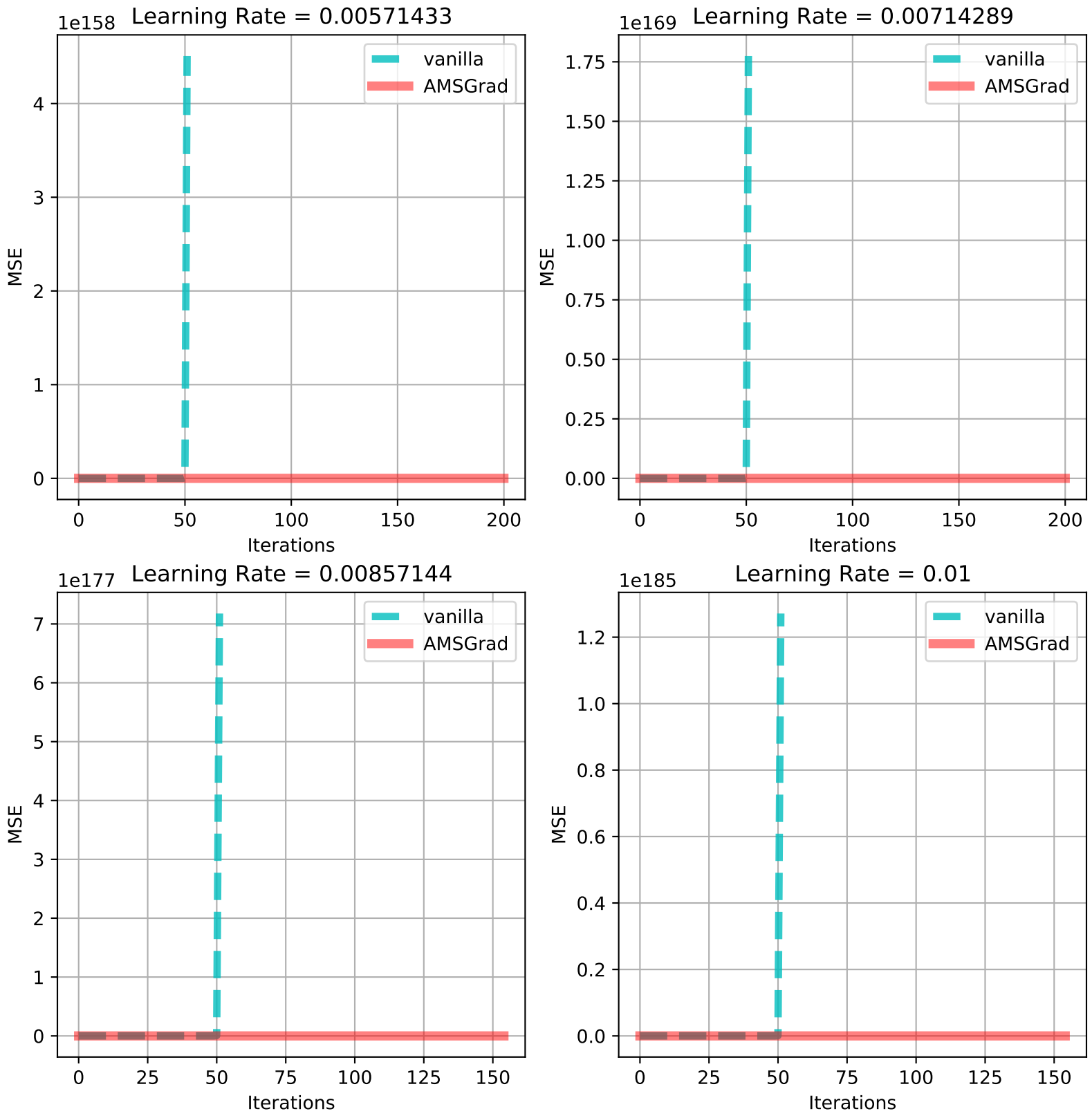


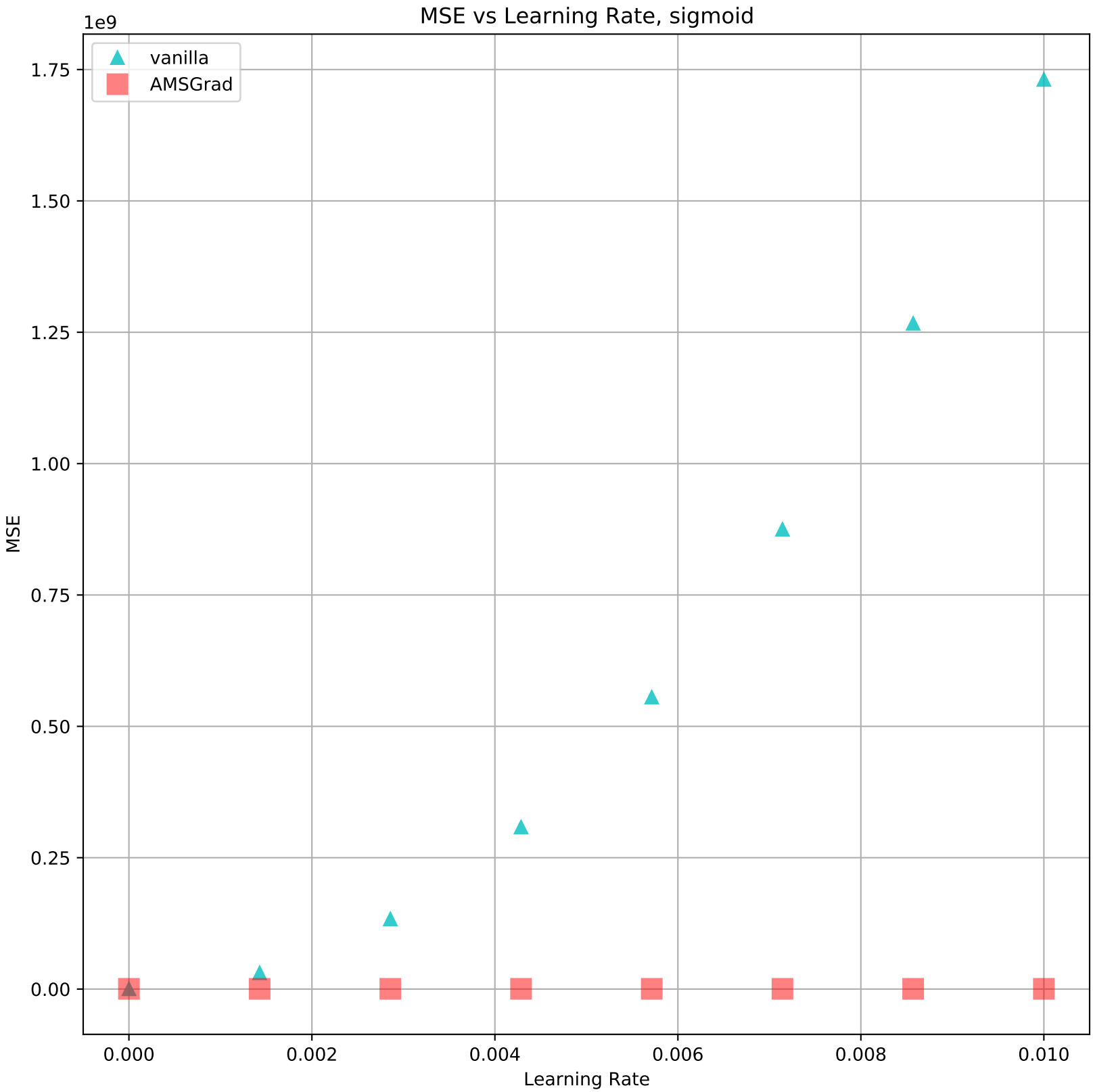


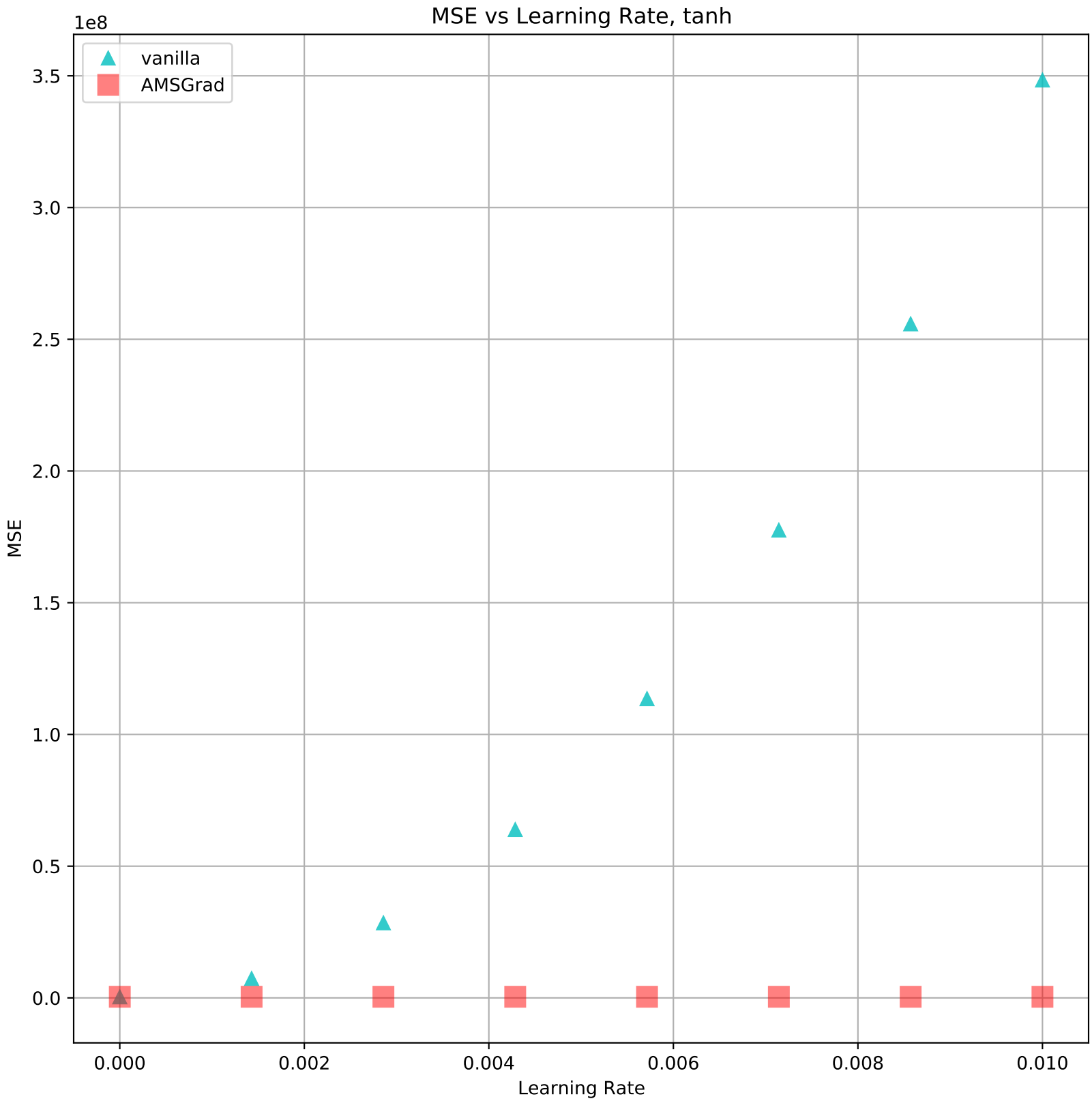


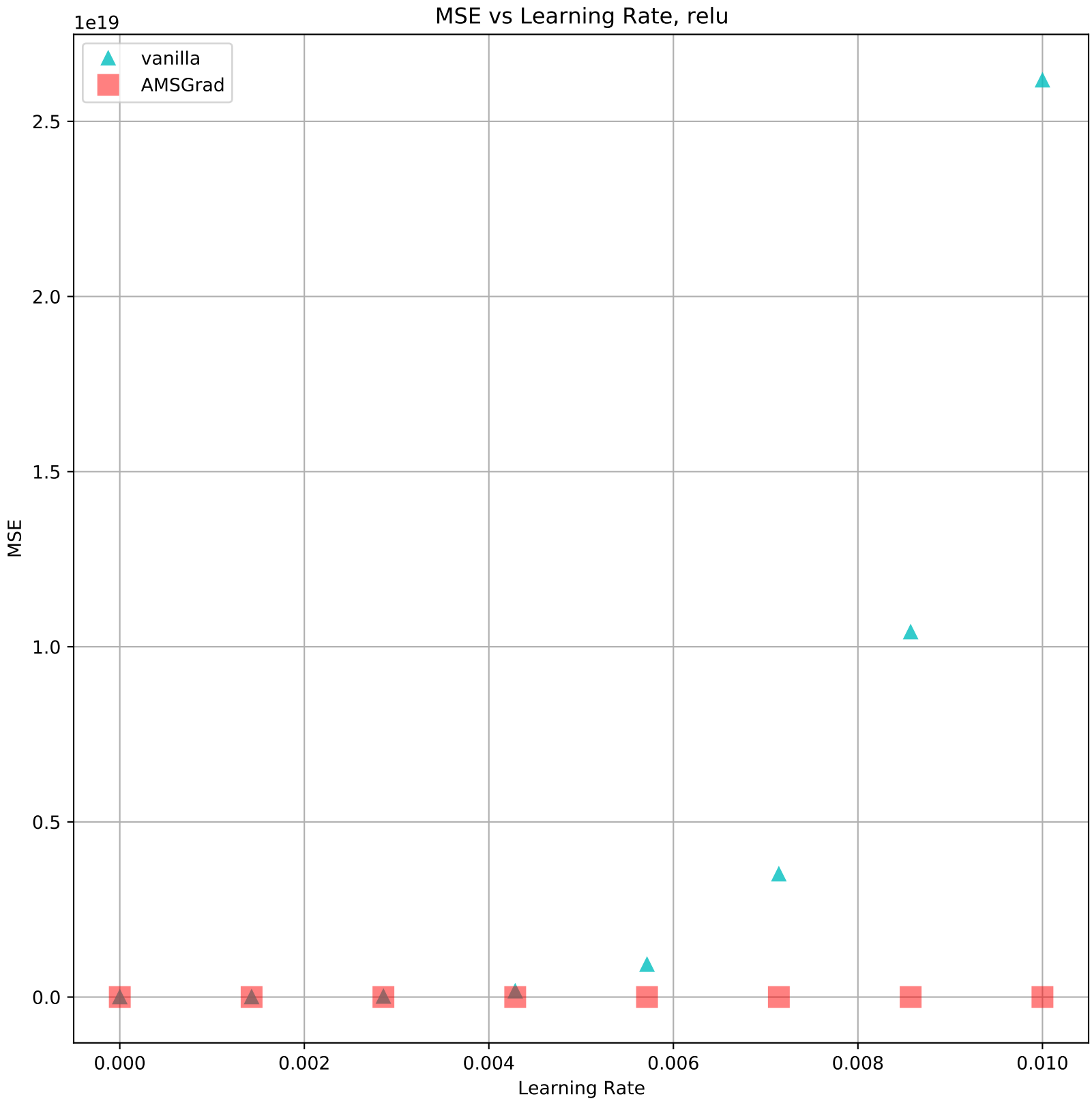


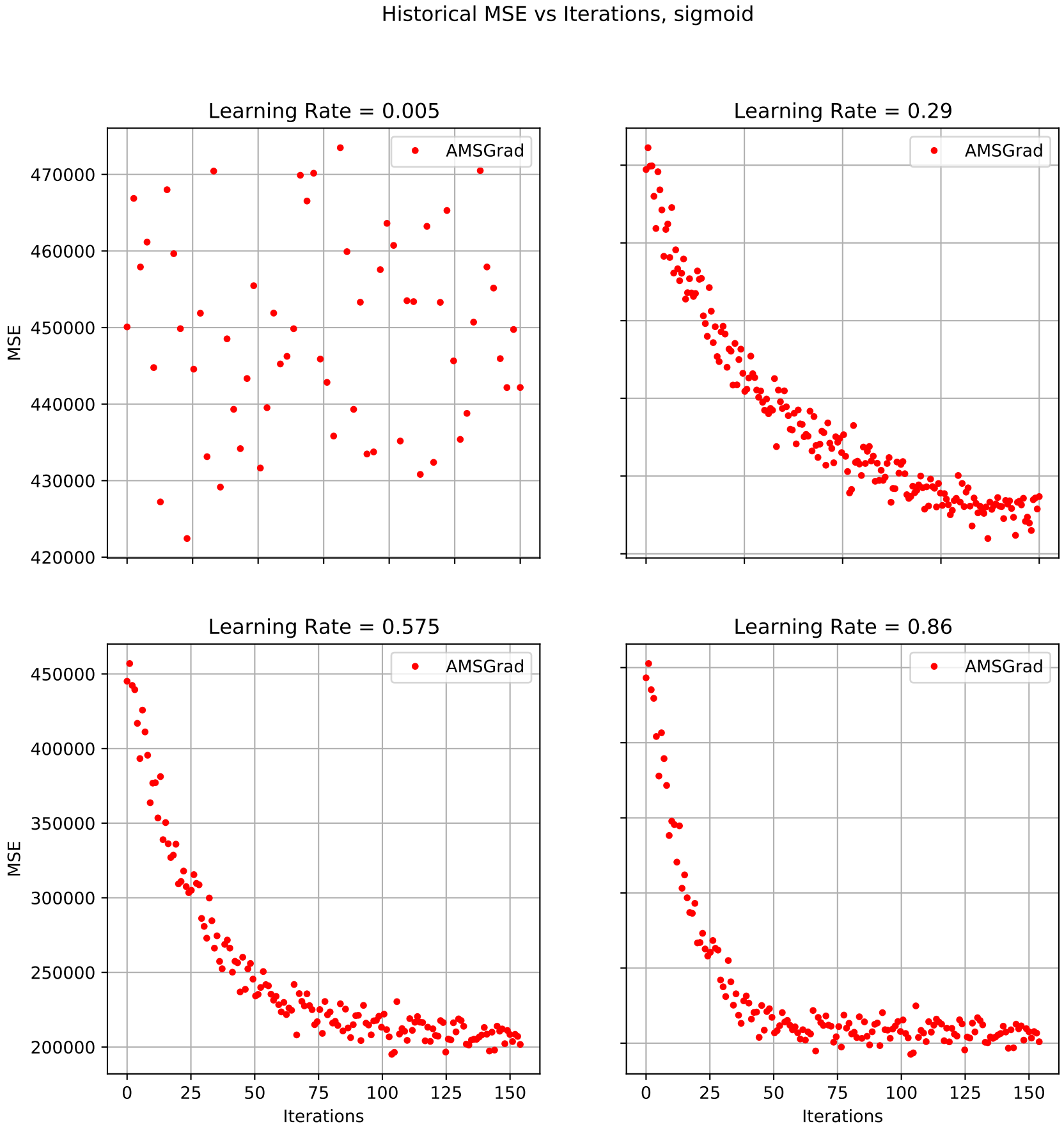
Historical MSE vs Iterations, Activation relu

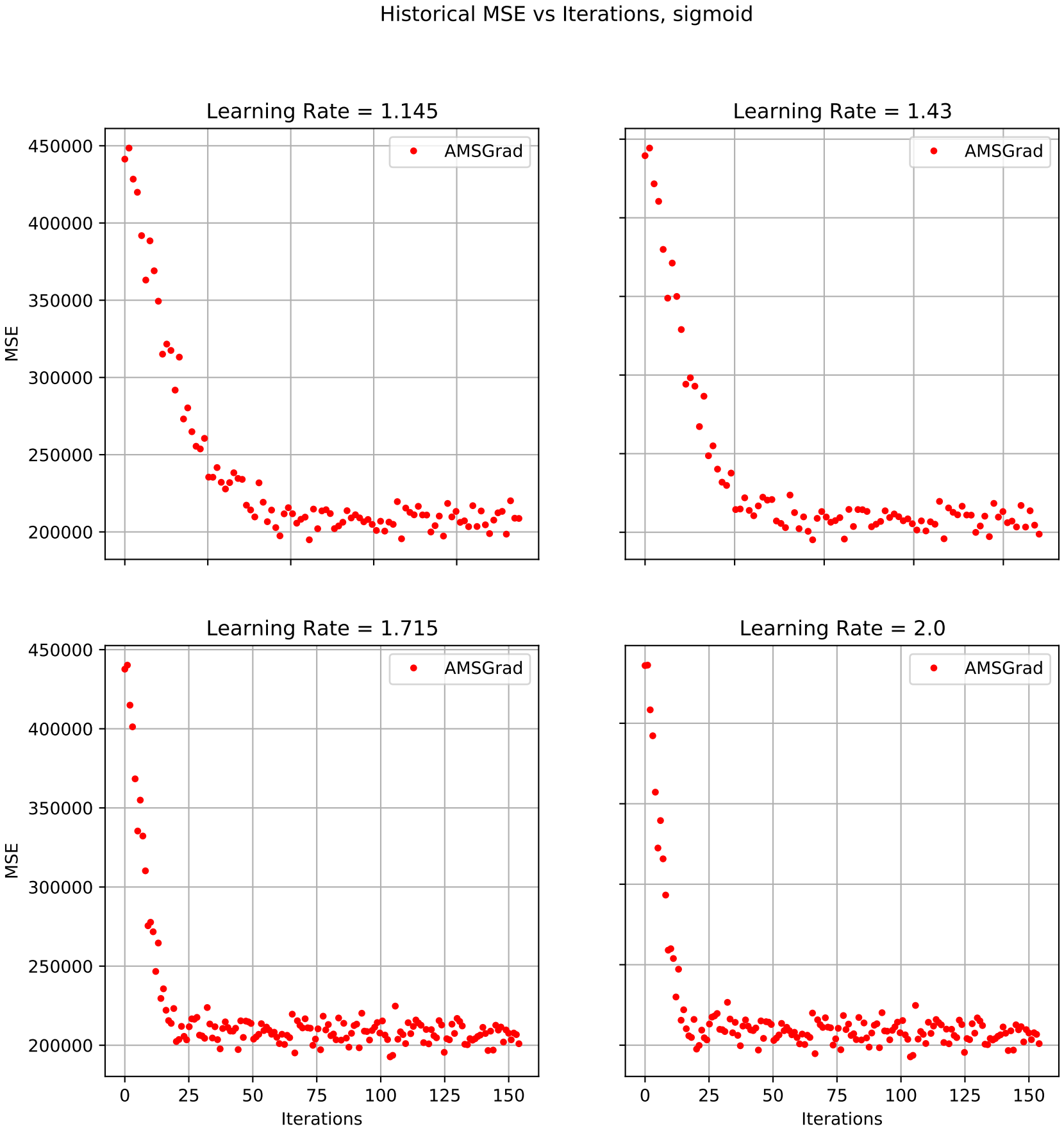


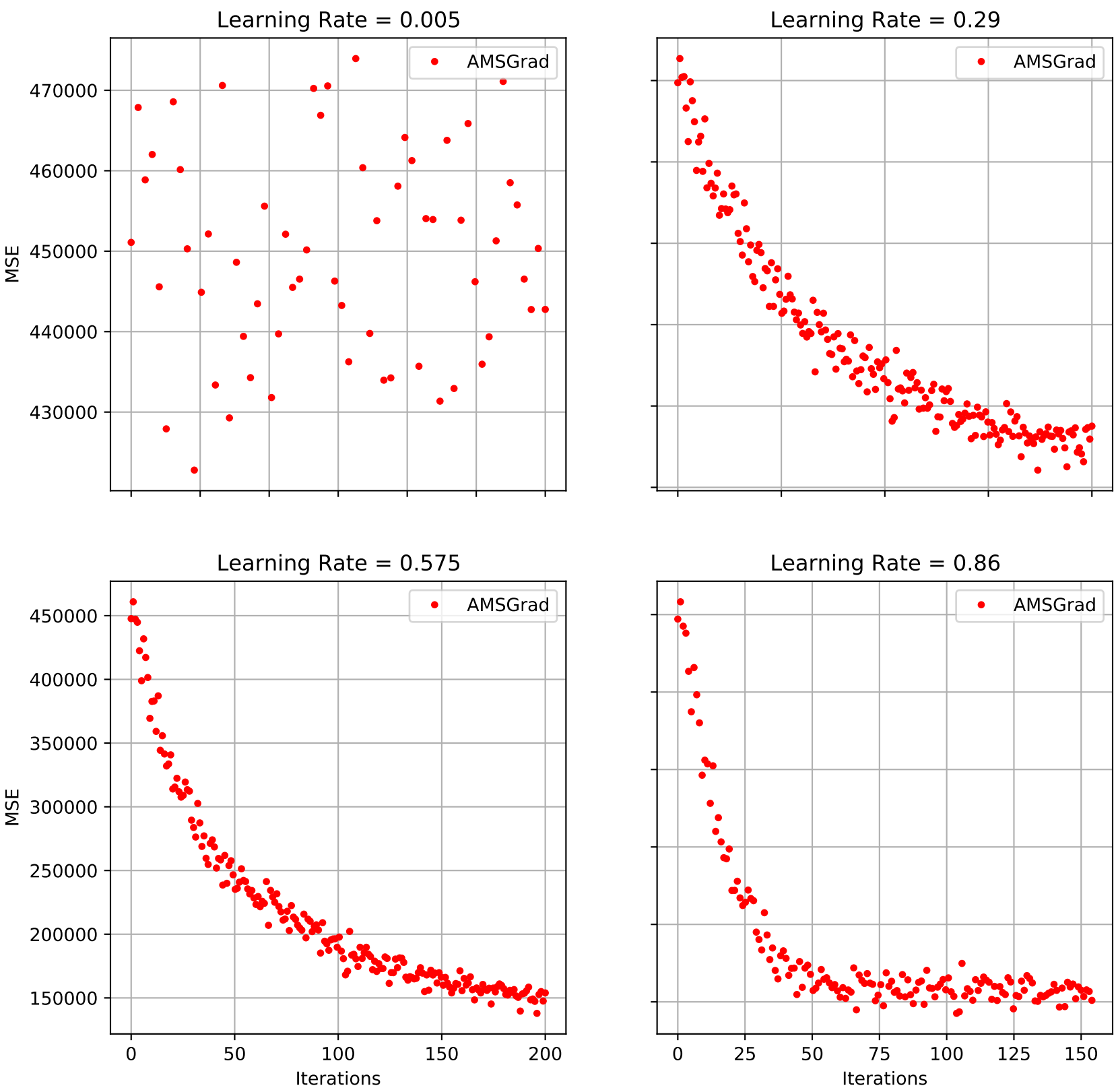

MSE vs Learning Rate, sigmoid
1e9
vanilla
AMSGrad
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
MSE
0.000
0.002
0.004
0.006
0.008
0.010
Learning Rate

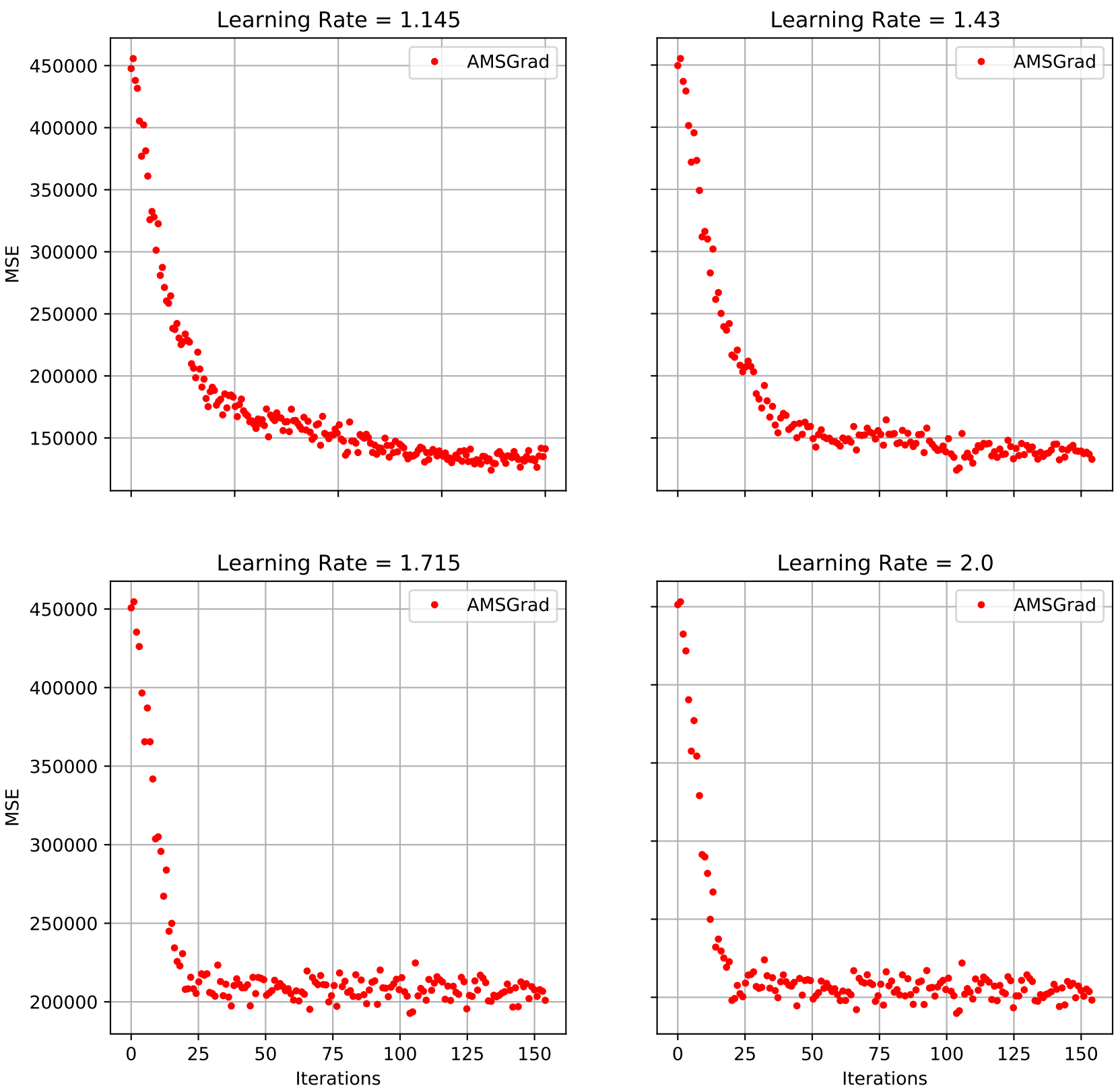

MSE vs Learning Rate, tanh
1e8
vanilla
AMSGrad
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
MSE
0.000
0.002
0.004
0.006
0.008
0.010
Learning Rate

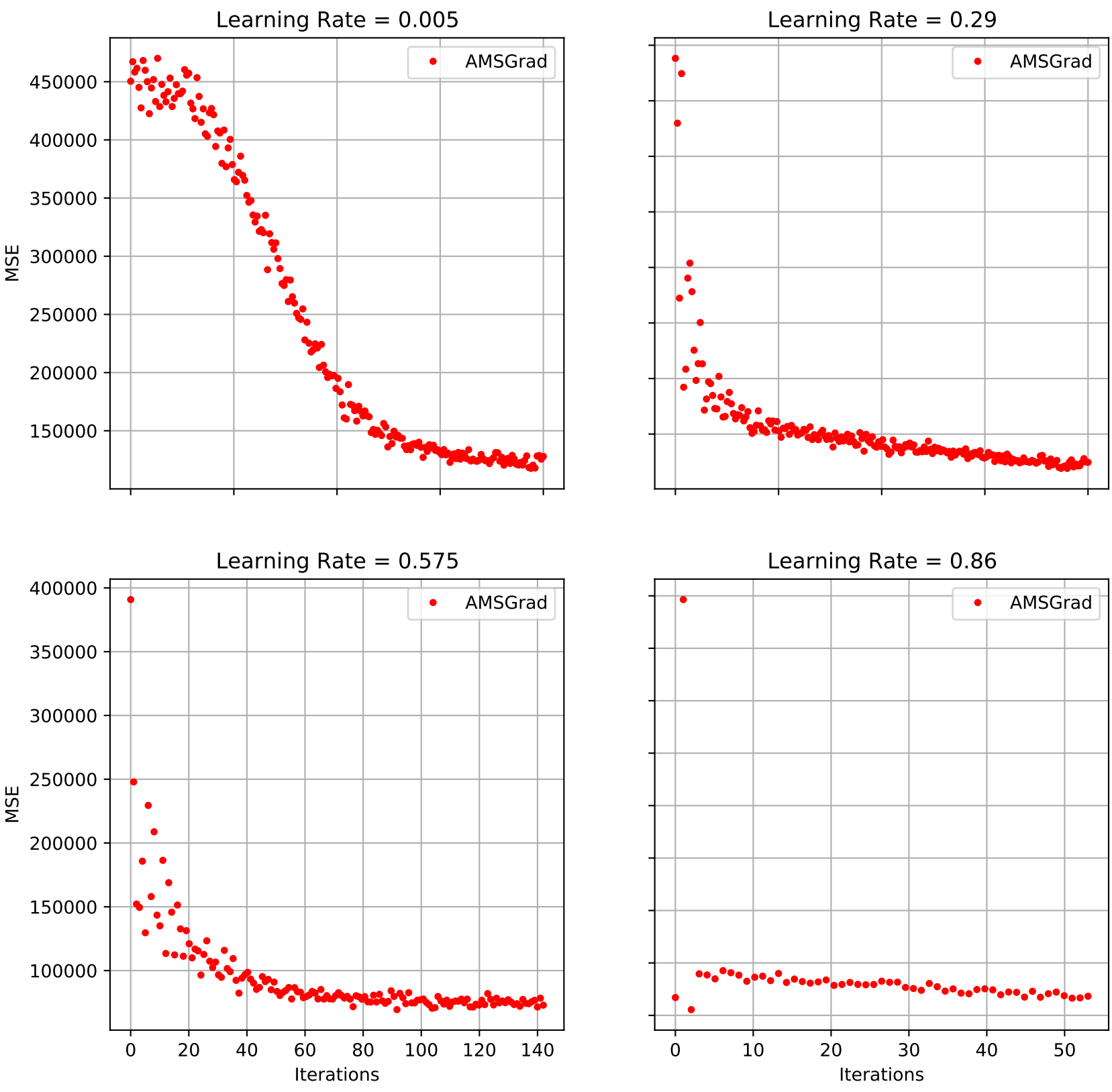

MSE vs Learning Rate, relu
1e19
vanilla
AMSGrad
2.5
2.0
1.5
1.0
0.5
0.0
MSE
0.000
0.002
0.004
0.006
0.008
0.010
Learning Rate

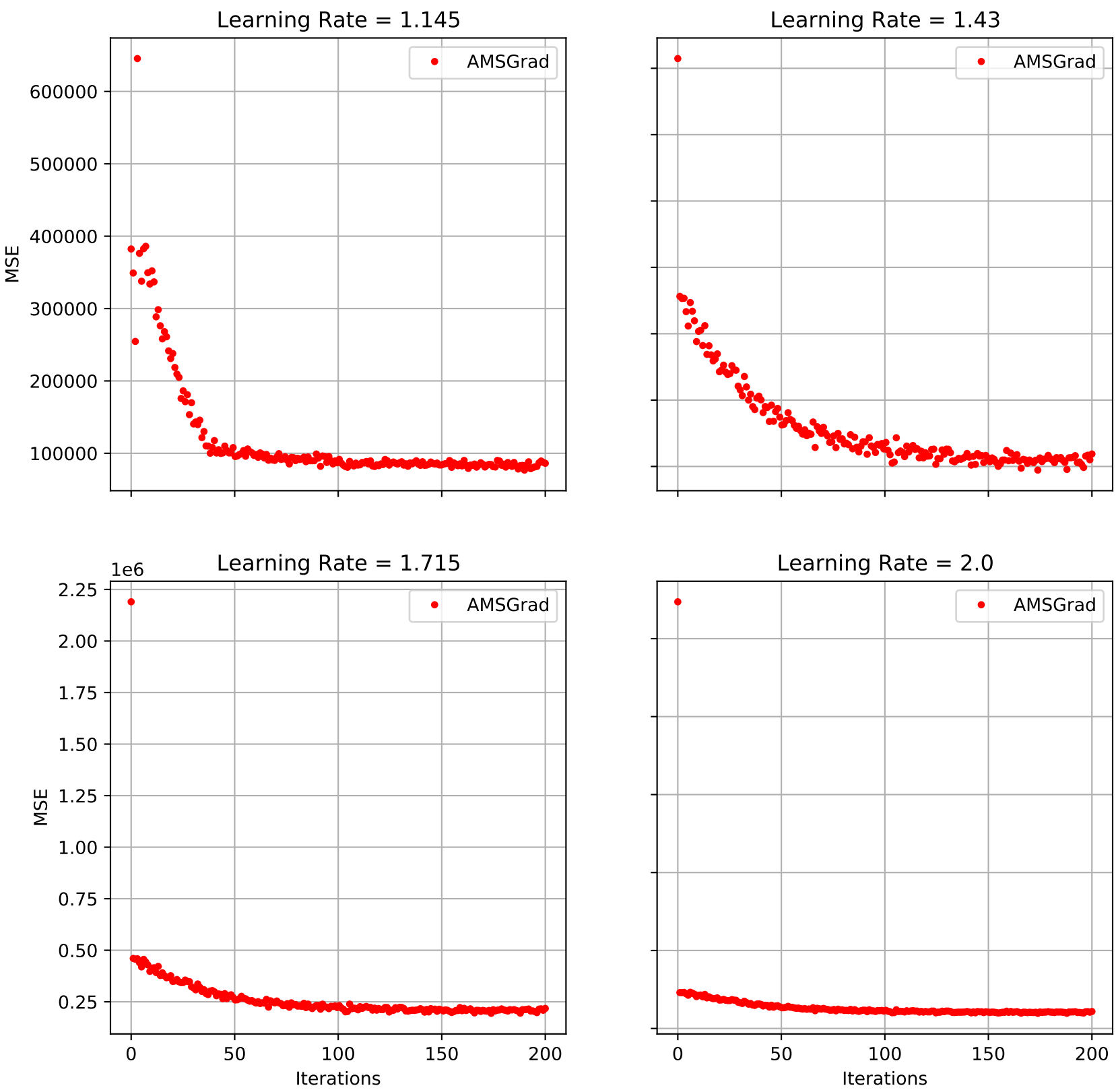

Historical MSE vs Iterations, sigmoid
Learning Rate = 0.005
Learning Rate = 0.29
Learning Rate = 0.575
Learning Rate = 0.86
AMSGrad
MSE
Iterations
470000
460000
450000
440000
430000
420000
400000
350000
300000
250000
200000
0
25
50
75
100
125
150

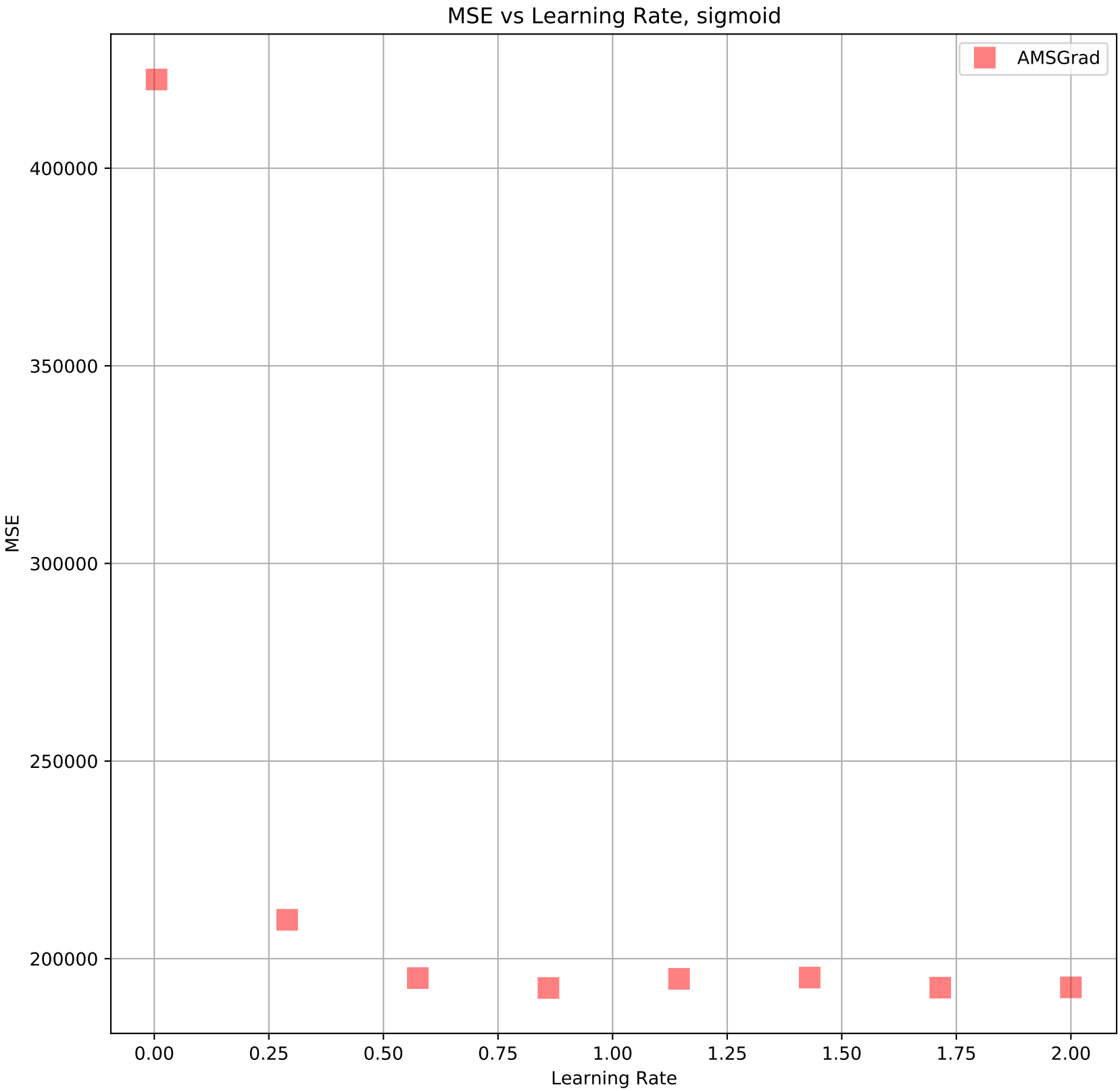

Historical MSE vs Iterations, sigmoid
Learning Rate = 1.145
Learning Rate = 1.43
Learning Rate = 1.715
Learning Rate = 2.0
AMSGrad
MSE
Iterations
450000
400000
350000
300000
250000
200000
0
25
50
75
100
125
150

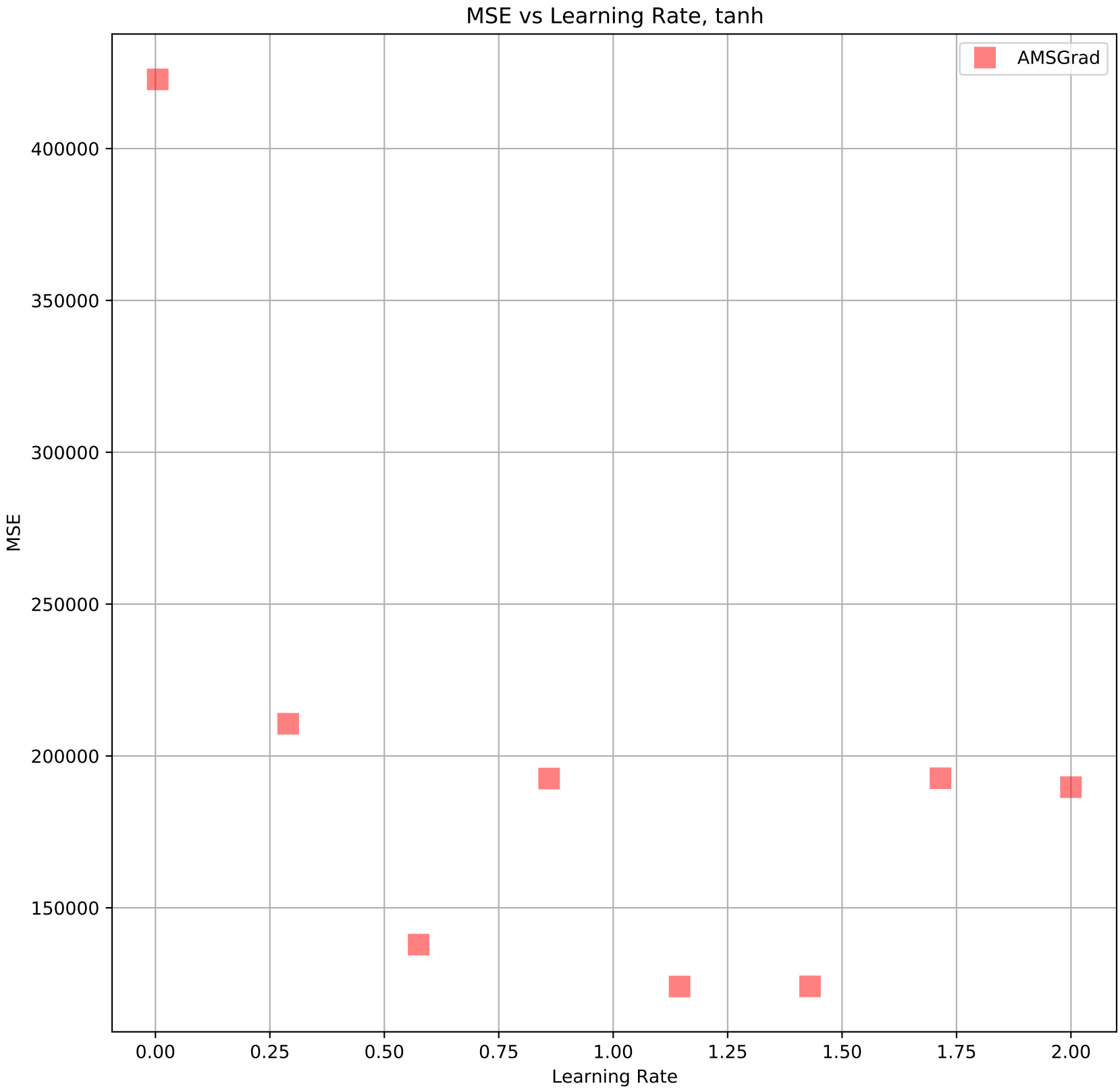

Historical MSE vs Iterations, tanh
Learning Rate = 0.005
Learning Rate = 0.29
Learning Rate = 0.575
Learning Rate = 0.86
AMSGrad
MSE
Iterations
470000
460000
450000
440000
430000
400000
350000
300000
250000
200000
150000
0
25
50
75
100
125
150
200

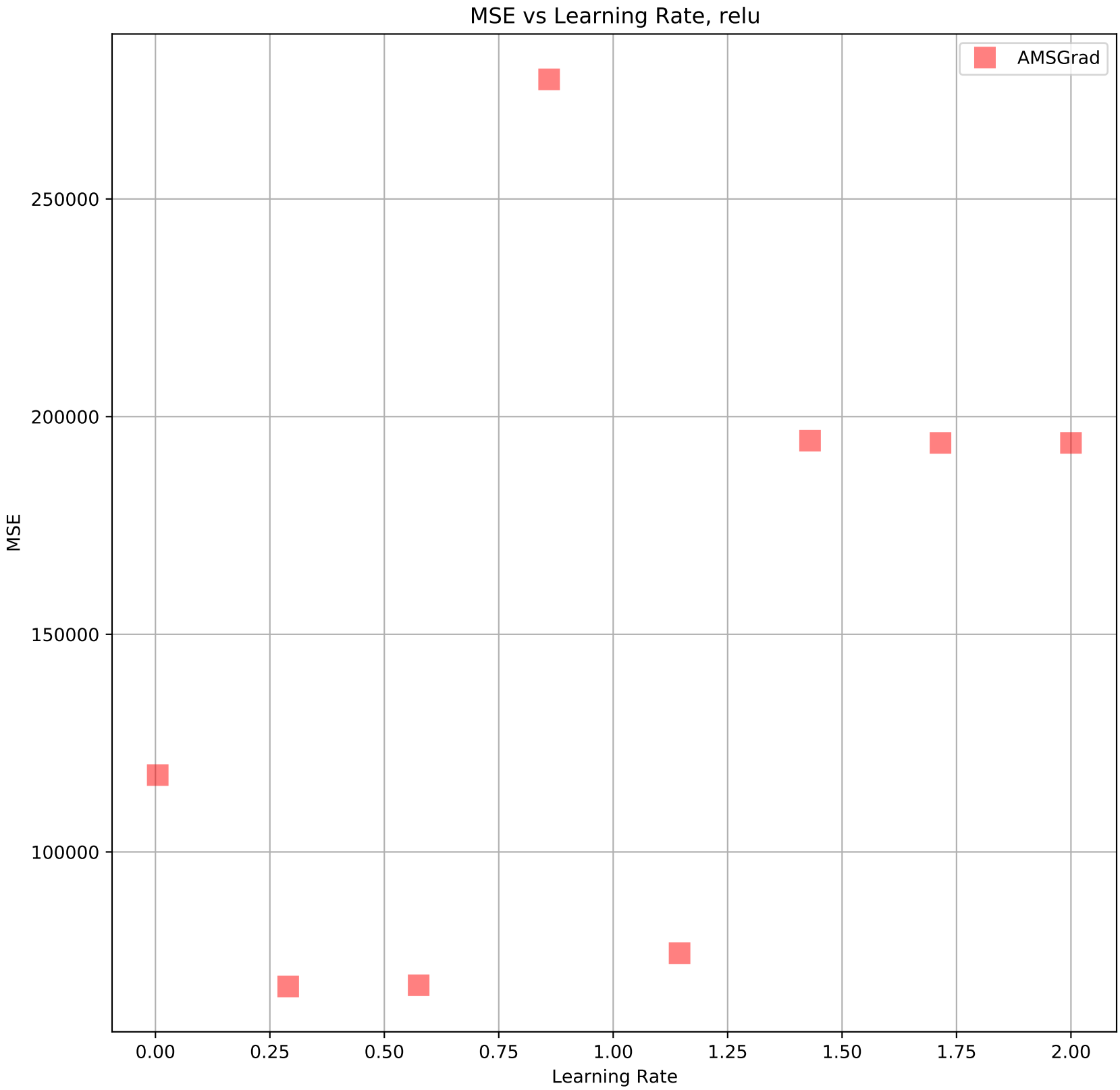

Historical MSE vs Iterations, tanh
Learning Rate = 1.145
Learning Rate = 1.43
Learning Rate = 1.715
Learning Rate = 2.0
AMSGrad
MSE
Iterations
450000
400000
350000
300000
250000
200000
150000
0
25
50
75
100
125
150


Historical MSE vs Iterations, relu
Learning Rate = 0.005
Learning Rate = 0.29
Learning Rate = 0.575
Learning Rate = 0.86
AMSGrad
MSE
Iterations
450000
400000
350000
300000
250000
200000
150000
100000
0
20
40
60
80
100
120
140
10
30
50


Historical MSE vs Iterations, relu
Learning Rate = 1.145
Learning Rate = 1.43
Learning Rate = 1.715
Learning Rate = 2.0
AMSGrad
MSE
1e6
Iterations
600000
500000
400000
300000
200000
100000
2.25
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0
50
100
150
200


MSE vs Learning Rate, sigmoid
AMSGrad
400000
350000
300000
250000
200000
MSE
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
Learning Rate


MSE vs Learning Rate, tanh
AMSGrad
400000
350000
300000
250000
200000
150000
MSE
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
Learning Rate


MSE vs Learning Rate, relu
AMSGrad
250000
200000
150000
100000
MSE
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
Learning Rate

Performance Comparison Sorted by MSE Vanilla

Activation	Learning Rate	Iterations	MSE	R2
sigmoid	8.000000000000001e-07	400	111641.51734863684	0.463108815390054
tanh	1.15e-06	400	114722.22262100408	0.4482935070496076
sigmoid	4.65e-06	395	121251.86472980437	0.4168920412676166
sigmoid	8.000000000000001e-07	300	122626.75801779158	0.4102800916665148
sigmoid	4.65e-06	300	124406.60743346185	0.40172068219322354
sigmoid	4.3e-06	300	125413.61487275657	0.3968779191259175
sigmoid	4.3e-06	342	125413.61487275657	0.3968779191259175
tanh	1.15e-06	300	126582.72123813306	0.3912556119739835
tanh	2.55e-06	300	134194.78333150822	0.35464871937976916
tanh	2.55e-06	342	134194.78333150822	0.35464871937976916
sigmoid	3.95e-06	259	134396.44174509982	0.35367893119405647
sigmoid	4.3e-06	200	134621.70097592365	0.3525956451714951
sigmoid	4.65e-06	200	135800.2559294307	0.3469278991560021
sigmoid	3.95e-06	200	136037.66257370653	0.34578619544660616
tanh	1.5e-06	259	137904.7664033713	0.3368071739257139
sigmoid	4.5000000000000003e-07	400	139783.52096616433	0.3277721232853743
tanh	1.5e-06	200	141987.64054472075	0.3171723715115098
tanh	2.55e-06	200	143553.71224512128	0.30964103271945054
tanh	5e-06	155	148665.00006664486	0.28506052325888376
sigmoid	2.2e-06	194	152648.2423086749	0.2659048570094408
tanh	1.15e-06	200	153569.02356825254	0.26147676114544893
sigmoid	8.000000000000001e-07	200	156032.77785743418	0.24962840947197784
tanh	4.65e-06	154	157451.88135427903	0.2428038501539076
tanh	3.95e-06	154	173809.09379462985	0.16414097121262872
sigmoid	4.5000000000000003e-07	300	174574.49236390833	0.16046012062659287
tanh	3.2500000000000002e-06	154	175410.64638857933	0.15643900253823784
tanh	4.3e-06	154	176373.6349693102	0.15180792897212103
tanh	8.000000000000001e-07	259	182659.8985229804	0.12157688620009677
tanh	8.000000000000001e-07	200	188410.29673863764	0.09392285405055234
tanh	1.85e-06	154	206938.415974528	0.004819945729385822

Performance Comparison Sorted by MSE AMSGrad

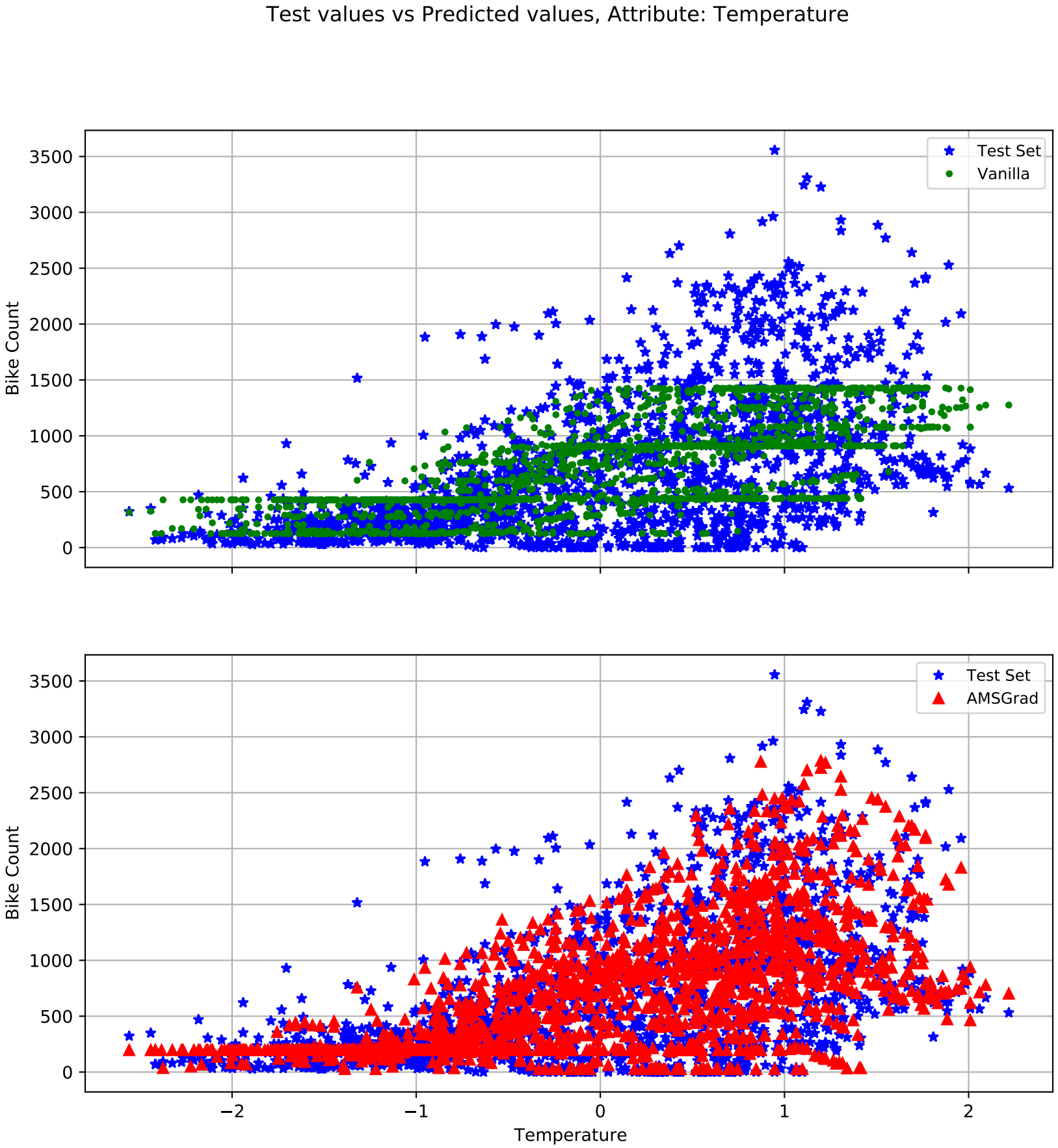
Activation	Learning Rate	Iterations	MSE	R2
relu	0.2892857142857143	373	70179.496466884	0.6625023209262666
relu	0.2892857142857143	300	70728.15724627646	0.6598637762092763
relu	0.6464285714285715	200	73559.83494591291	0.6462460573650963
relu	0.6464285714285715	237	73559.83494591291	0.6462460573650963
relu	0.2892857142857143	200	73900.49209330209	0.6446078153942667
relu	0.5571428571428572	237	76392.22194056491	0.632624927475902
relu	0.5571428571428572	200	76392.22194056491	0.632624927475902
relu	0.2	237	76680.02753954833	0.6312408519756287
relu	0.2	200	76680.02753954833	0.6312408519756287
relu	0.37857142857142856	154	77073.01820636228	0.6293509347686859
tanh	0.825	300	125357.72875085475	0.3971466790541567
tanh	0.825	327	125357.72875085475	0.3971466790541567
tanh	1.0035714285714286	259	129582.75786391548	0.3768282443052009
tanh	1.0035714285714286	200	131220.6462018239	0.36895153471827946
tanh	1.2714285714285714	200	134795.80036165778	0.35175839011023713
tanh	1.2714285714285714	205	134795.80036165778	0.35175839011023713
tanh	1.3607142857142858	154	136284.37057630488	0.34459976091119826
tanh	0.9142857142857144	200	136369.25777435402	0.34419153295621174
tanh	0.9142857142857144	238	136369.25777435402	0.34419153295621174
tanh	0.7357142857142858	400	136971.67064998607	0.34129449097663733
tanh	1.1821428571428572	155	137578.0915902921	0.3383781739581535
tanh	0.7357142857142858	300	139906.56374181577	0.32718040272168025
tanh	0.825	200	141263.50606965274	0.3206547804340083
relu	0.46785714285714286	54	144144.058949547	0.3068020177275643
tanh	1.45	154	146137.68466463458	0.2972145443819658
tanh	0.7357142857142858	200	146364.83241036232	0.2961221763706605
tanh	1.092857142857143	154	146694.31685213884	0.2945376646544364
tanh	0.6464285714285715	259	146901.26205983266	0.29354245193835427
tanh	0.6464285714285715	200	150876.96425794842	0.2744230462414843
tanh	0.46785714285714286	259	205750.40236696214	0.010533179016129268
tanh	0.37857142857142856	259	206366.75582999428	0.007569095861720165
relu	1.1821428571428572	53	206662.93257366781	0.006144762992181807
tanh	0.46785714285714286	200	207688.35658521965	0.0012134334530973234

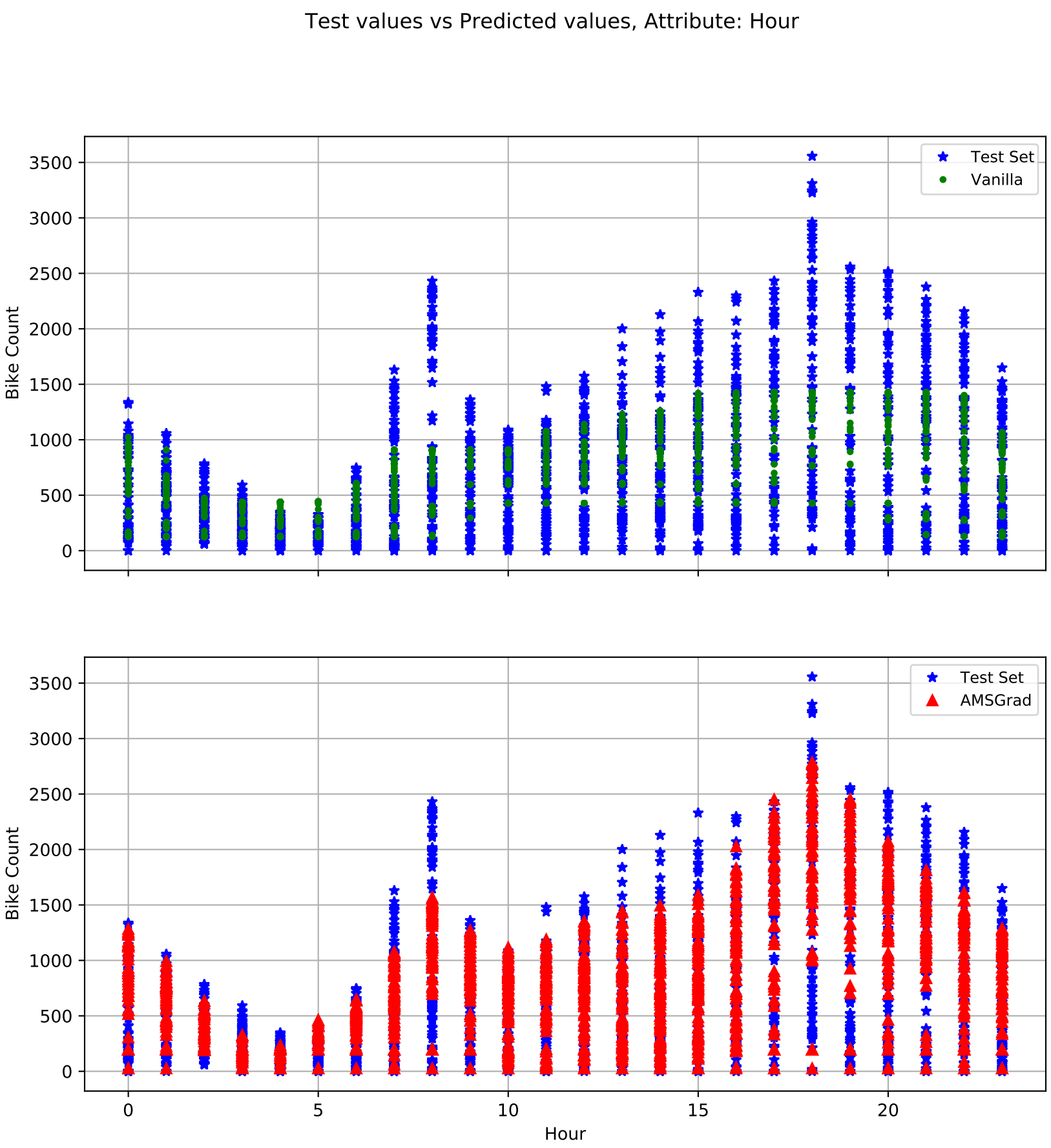
Performance Comparison Training Set Sorted by MSE

Algorithm	Activation	Learning Rate	Epochs	Layers	MSE	R2
AMSGrad	relu	0.2892857142857143	328	[6, 4]	50500.15476071578	0.7571415316060521
AMSGrad	relu	0.2892857142857143	373	[8]	60579.27408197471	0.7086704032952743
AMSGrad	relu	0.2892857142857143	236	[4, 4]	61101.28994474574	0.7061599956834439
AMSGrad	relu	0.2892857142857143	373	[6]	63220.03485856603	0.6959708161229856
AMSGrad	relu	0.2892857142857143	289	[2, 2]	66846.24992009952	0.6785321163787197
AMSGrad	tanh	0.825	300	[8]	68431.53828765663	0.6709083634074287
AMSGrad	relu	0.2892857142857143	373	[4, 2]	70179.496466884	0.6625023209262666
AMSGrad	relu	0.2892857142857143	142	[6, 2]	76430.52533296906	0.6324407240163752
Vanilla	sigmoid	8.000000000000001e-07	400	[8]	77917.24849769745	0.6252909773983191
Vanilla	tanh	1.15e-06	373	[8]	79836.17698104646	0.6160627278091477
Vanilla	sigmoid	8.000000000000001e-07	400	[6]	82521.67499607858	0.6031479964009772
AMSGrad	relu	0.2892857142857143	154	[4]	86071.80787899494	0.5860751807110984
Vanilla	sigmoid	8.000000000000001e-07	395	[6, 4]	91712.0660074954	0.5589508192725132
Vanilla	sigmoid	8.000000000000001e-07	400	[4]	92756.37059369602	0.5539286918445248
Vanilla	sigmoid	8.000000000000001e-07	400	[4, 4]	98109.60877522643	0.5281846276555213
Vanilla	tanh	1.15e-06	259	[6]	98736.85739598039	0.5251681489920337
Vanilla	tanh	1.15e-06	395	[4]	102484.33579216995	0.5071462861294276
AMSGrad	tanh	0.825	154	[6]	103807.10077262016	0.5007850248873691
AMSGrad	tanh	0.825	224	[6, 4]	109184.23679232682	0.4749260344688341
Vanilla	sigmoid	8.000000000000001e-07	400	[4, 2]	111641.51734863684	0.463108815390054
Vanilla	tanh	1.15e-06	327	[6, 4]	112812.586654712	0.45747706833097734
Vanilla	tanh	1.15e-06	400	[4, 2]	114722.22262100408	0.4482935070496076
AMSGrad	tanh	0.825	300	[4, 2]	125357.72875085475	0.3971466790541567
Vanilla	tanh	1.15e-06	259	[4, 4]	128369.22400277972	0.3826642061206861
AMSGrad	tanh	0.825	205	[4, 4]	129205.28310134365	0.3786435445380406
AMSGrad	tanh	0.825	154	[4]	130591.34978277363	0.3719778613745145
AMSGrad	tanh	0.825	259	[6, 2]	143319.01129191165	0.31076972458776597
Vanilla	sigmoid	8.000000000000001e-07	301	[2, 2]	143941.6728462625	0.3077753054194692
Vanilla	sigmoid	8.000000000000001e-07	395	[6, 2]	149673.04958056533	0.28021274878817215
AMSGrad	tanh	0.825	154	[2, 2]	153229.03193258104	0.263111802627939

Performance Comparison Testing set

Algorithm	Activation	Learning Rate	Layers	MSE	R2
Vanilla	sigmoid	8.000000000000001e-07	[8]	80047.75991884114	0.6153625783046273
AMSGrad	relu	0.2892857142857143	[6, 4]	54783.58496961835	0.7367594434209317


Test values vs Predicted values, Attribute: Temperature
Test Set
Vanilla
Bike Count
3500
3000
2500
2000
1500
1000
500
0
Test Set
AMSGrad
Bike Count
3500
3000
2500
2000
1500
1000
500
0
−2
−1
0
1
2
Temperature


Test values vs Predicted values, Attribute: Hour
Test Set
Vanilla
Bike Count
3500
3000
2500
2000
1500
1000
500
0
Test Set
AMSGrad
Bike Count
3500
3000
2500
2000
1500
1000
500
0
0
5
10
15
20
Hour