STAT 9100 HOMEWORK 4

Peng Shao

March 13, 2016

1

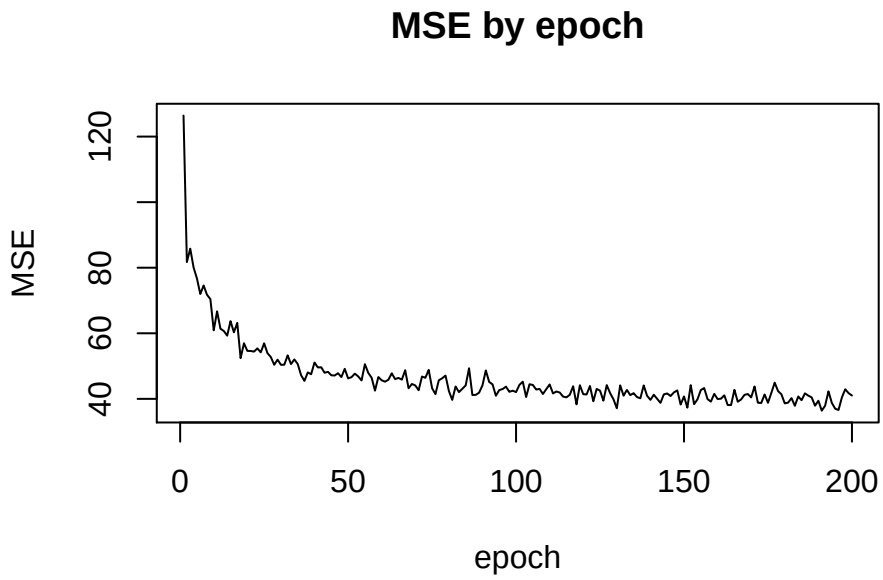
a.

Firstly, I tried two candidates for each parameter to explore how the values of parameters affect the model. The comparisons of performances are shown as below, where n is the number of hidden units, eta is learning rate and niter is the number of iterations.

batch_size	n	eta	niter	mse	time
1000	100	0.001	200	82.72997	25.0333816329638 mins
500	100	0.001	200	81.86168	15.0094160517057 mins
1000	50	0.001	200	95.27384	25.9833889007568 mins
500	50	0.001	200	95.29577	14.5915921171506 mins
1000	100	0.010	200	51.97142	24.1317209005356 mins
500	100	0.010	200	52.22719	11.1662522514661 mins
1000	50	0.010	200	60.37581	21.4933004339536 mins
500	50	0.010	200	60.00229	9.93551119963328 mins
1000	100	0.001	100	99.09260	12.4043050169945 mins
500	100	0.001	100	100.88258	5.58180388212204 mins
1000	50	0.001	100	129.29261	10.5478509346644 mins
500	50	0.001	100	124.47313	4.98161446650823 mins
1000	100	0.010	100	59.19046	12.3960566163063 mins
500	100	0.010	100	59.09987	5.60734366575877 mins
1000	50	0.010	100	71.91141	10.4691515167554 mins
500	50	0.010	100	70.86252	4.99019093513489 mins

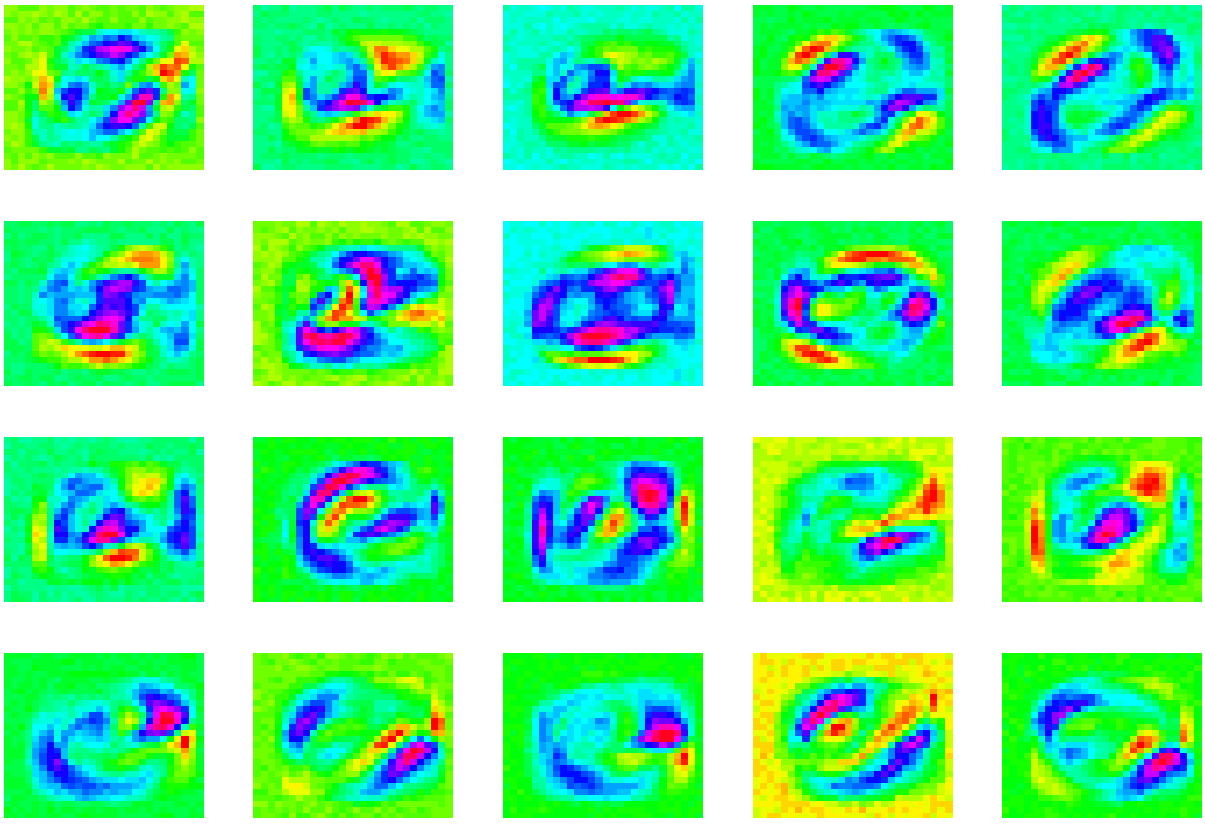
We can see that the learning rate has the most important effect and larger learning rate will lead the smaller mse and shorter runtime. So I try to increase the learning rate, and when the learning rate is larger than 0.5, there is no significant difference any more. Then I change the other three parameters and find that increasing these parameters does not greatly reduce the mse but will cost much more time. Hence the parameters of the final model I choose are: `batch_size=50`, `n=20`, `eta = 0.5` and `niter = 200`. The final MSE is about 41 on average and the program only takes 55 seconds.

b.



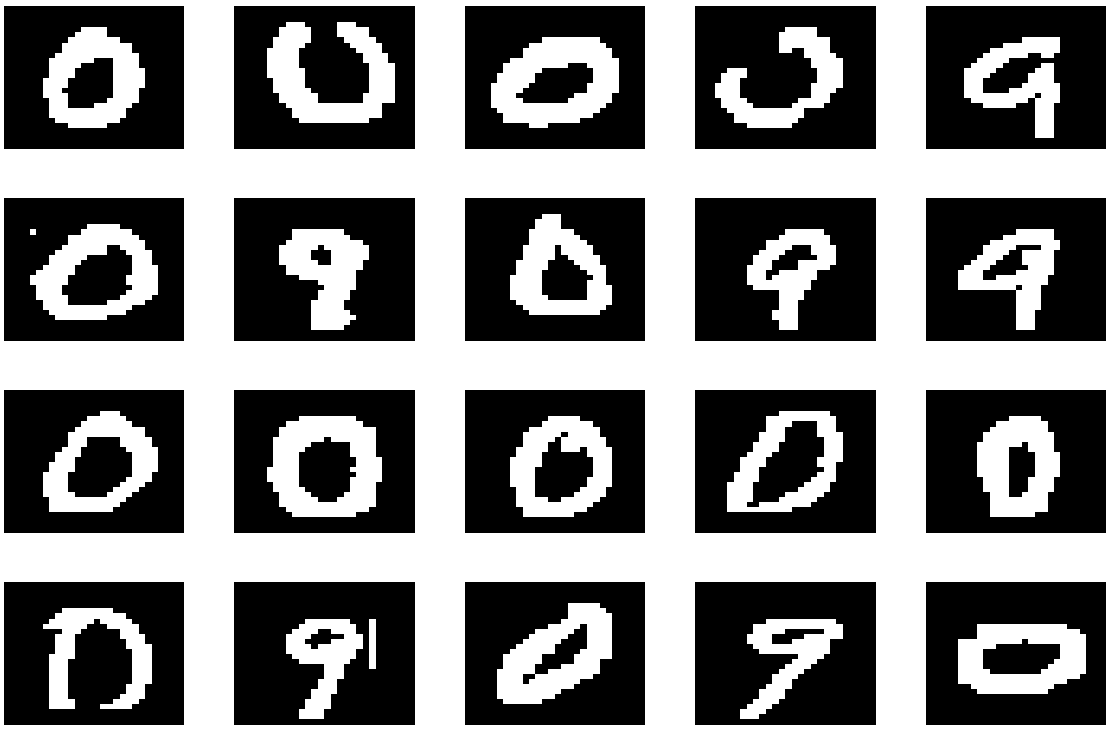
From the figure we can see that the algorithm converges very fast at the beginning and the MSE becomes small enough after about 50 iterations. Then more iterations does not improve the result so much.

c.



d.

I firstly draw the true images,



Then I draw the predicted images,