

alpha	0.1	0.2	0.3	0.4	0.5	0.6
accuracy	0.90167364	0.90794979	0.91422594	0.93305439	0.92050209	0.94142259

you can see with the increase of learning rate the accuracy begin to oscillate around the solut  
convex problem the local optima is also global optia

0.7	0.8	0.9
0.93933054	0.93514644	0.92887029

and the solution periodically. The gradient descent can converge to a local optimal. When the function is