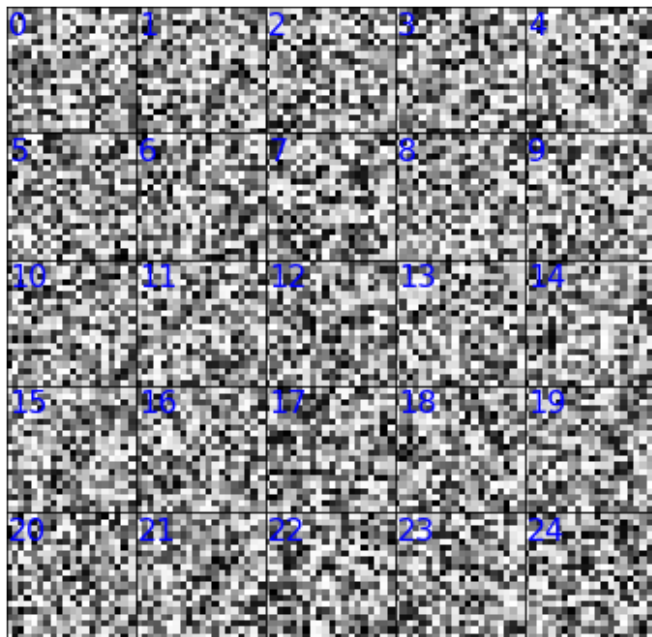
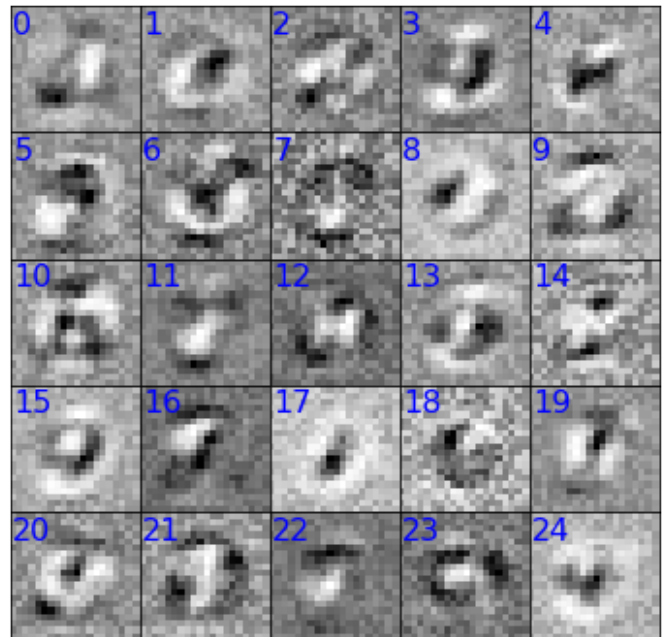


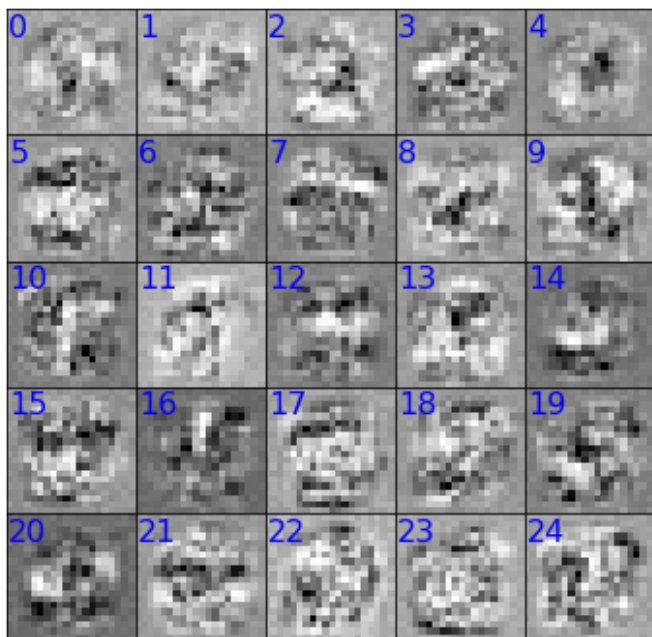
Visualization of the hidden layer



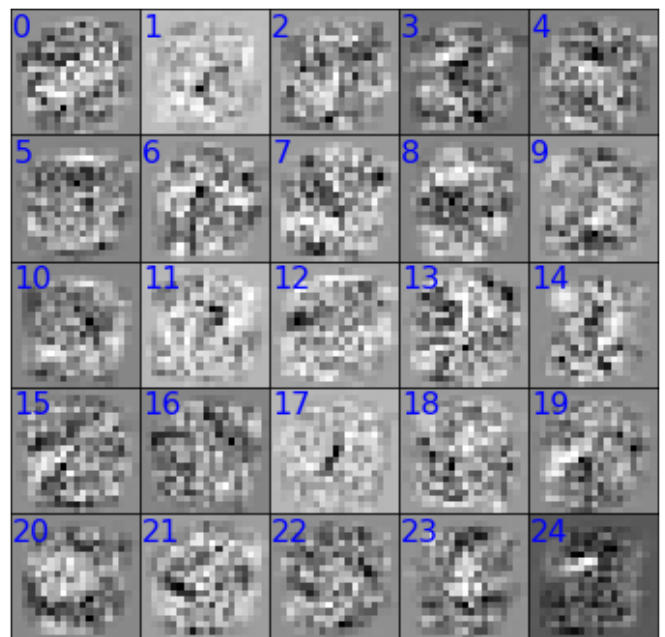
lambda=0.1, #iteration 1



lambda=0.1 #iteration 10

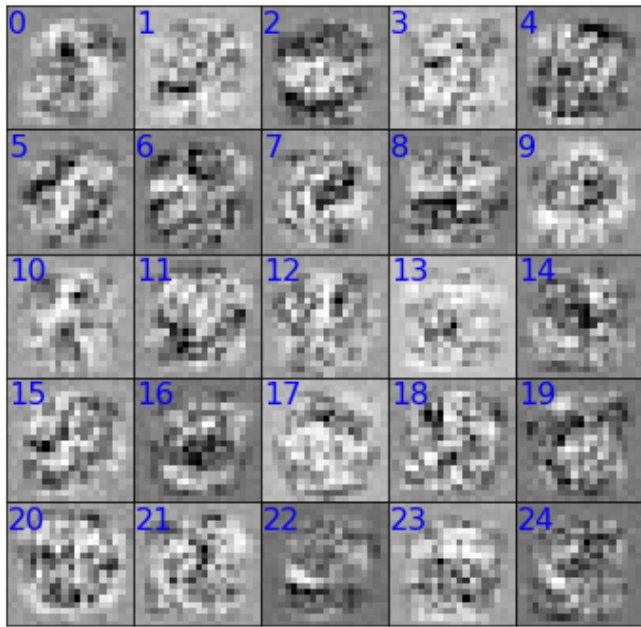


lambda=0.1 #iteration 100

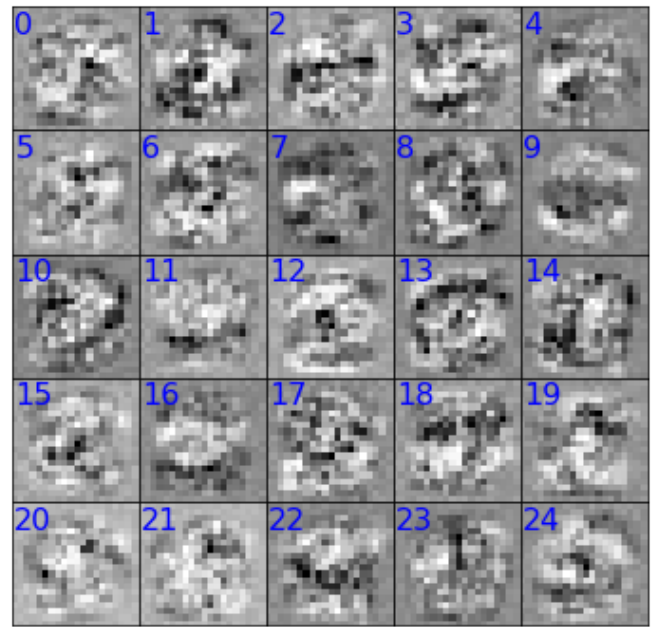


lambda=0.1 #iteration 300

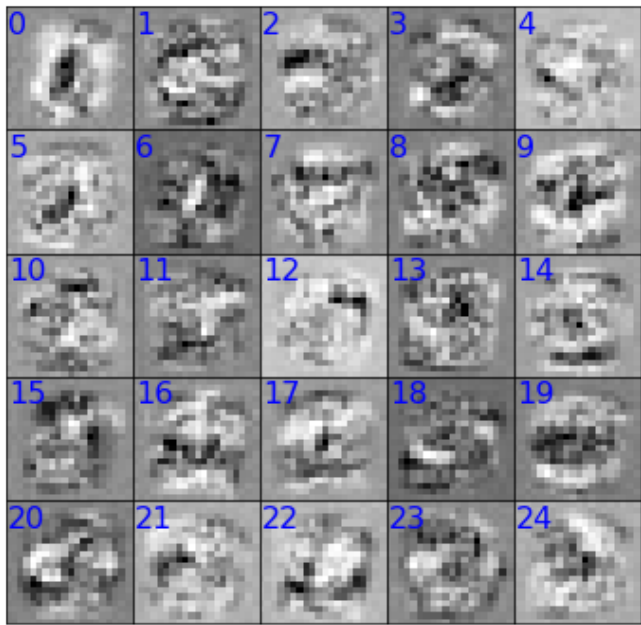
Fig. 1 The dependence of the parameter Θ_1 on the number of iterations



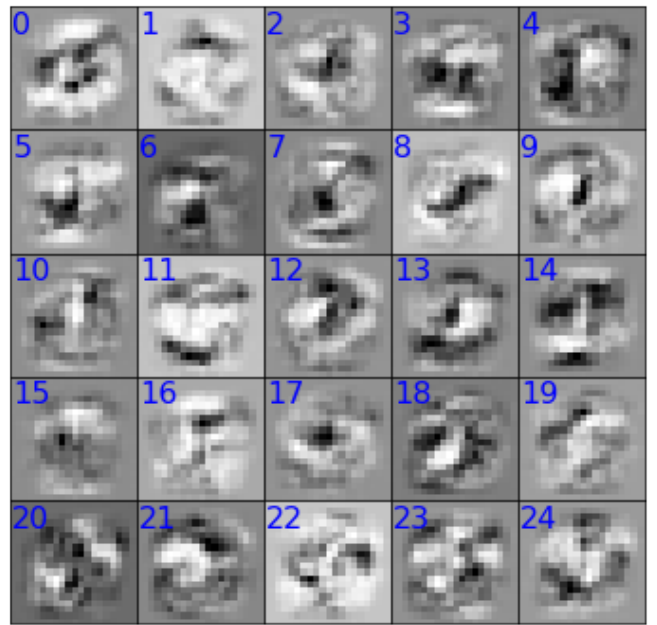
lambda=0 #iteration100



lambda=0.1 #iteration100



lambda=1 #iteration100



lambda=10 #iteration100

Fig. 2 The dependence of the parameter Θ_1 on the parameter λ

Training accuracy

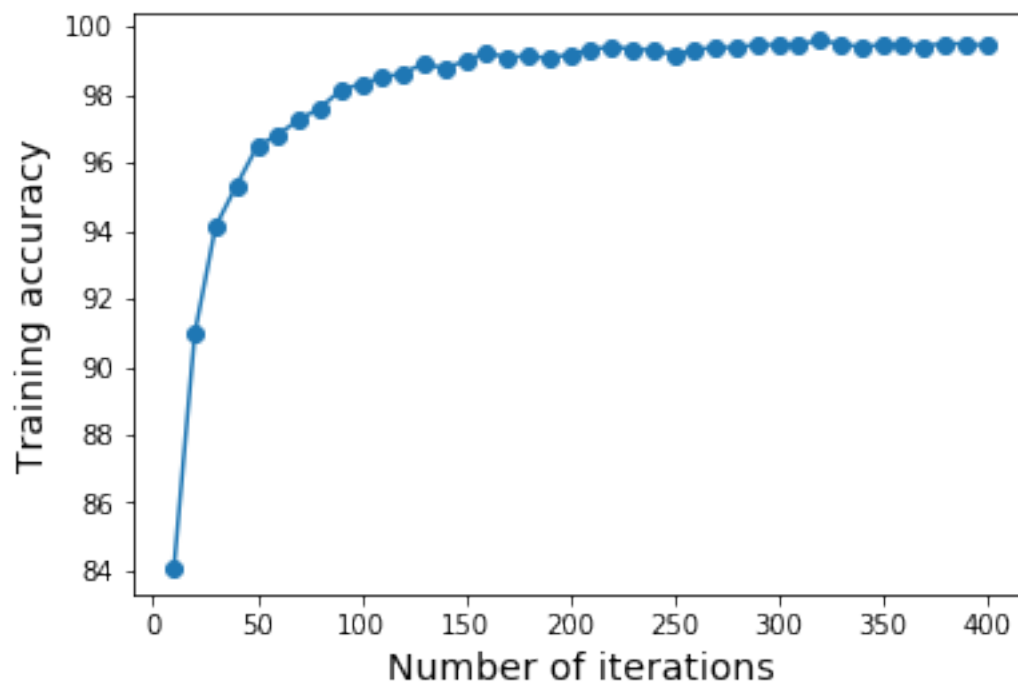


Fig. 3 The dependence of the training accuracy on the number of iterations ($\lambda = 1$)

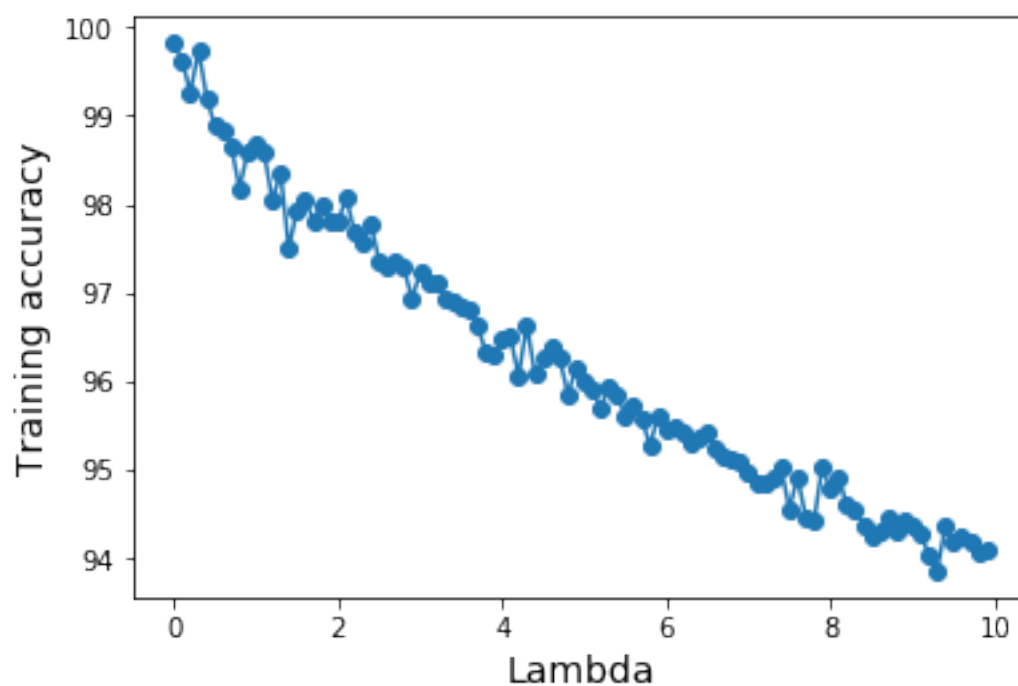


Fig. 4-1 The dependence of the training accuracy on the parameter λ (100 iterations)

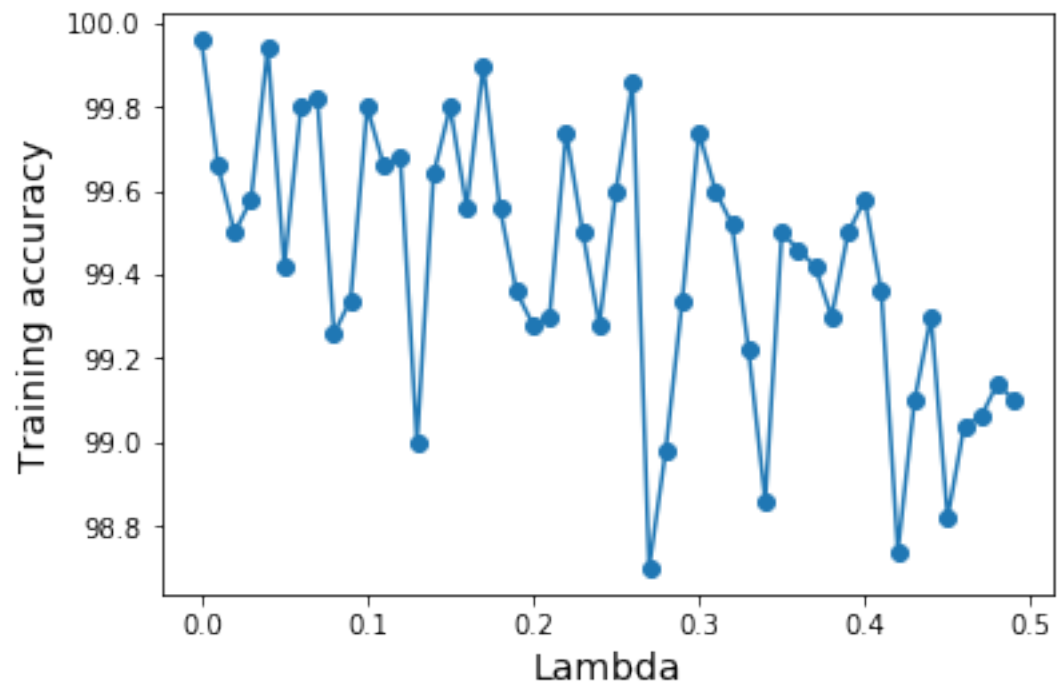


Fig. 4-2 The dependence of the training accuracy on the parameter λ (100 iterations)