

Results of optimization of H1 and H2

Common parameter :

early stopping is triggered when the average error on validation set over 4 epochs stops decreasing significantly, ie, $\text{new_value} > 0.95 * \text{old_value}$

Main result :

Error on validation set (averaged over 3 redundancy)

$\mu=0.01, v=0$							
H1 / H2	10	20	30	40	60	80	100
10	0.012						
20	0.010	0.009					
30	0.012	0.010	0.010				
40	0.013	0.011	0.008	0.009			
60	0.009	0.008	0.008	0.008	0.008		
80	0.008	0.009	0.008	0.008	0.009	0.009	
100	0.009	0.009	0.008	0.008	0.009	0.009	0.009

$\mu=0.01, v=0.05$							
H1 / H2	10	20	30	40	60	80	100
10	0.019						
20	0.012	0.010					
30	0.012	0.011	0.011				
40	0.010	0.010	0.010	0.009			
60	0.011	0.010	0.009	0.010	0.010		
80	0.009	0.009	0.011	0.008	0.008	0.008	
100	0.010	0.009	0.008	0.008	0.008	0.008	0.009

$\mu=0.001, v=0$							
H1 / H2	10	20	30	40	60	80	100
10	0.035						
20	0.042	0.030					
30	0.031	0.023	0.028				
40	0.024	0.023	0.023	0.027			
60	0.024	0.024	0.030	0.021	0.031		
80	0.026	0.022	0.024	0.027	0.025	0.022	
100	0.024	0.023	0.021	0.026	0.022	0.025	0.023

$\mu=0.001, v=0.05$							
H1 / H2	10	20	30	40	60	80	100
10	0.039						
20	0.031	0.033					
30	0.028	0.026	0.027				
40	0.028	0.027	0.026	0.027			
60	0.027	0.023	0.026	0.025	0.025		
80	0.025	0.025	0.024	0.025	0.025	0.024	
100	0.026	0.024	0.024	0.025	0.024	0.023	0.024

Results of optimization of H1 and H2

Results for control :

number of epochs before converging

$\mu=0.01, v=0$							
H1 / H2	10	20	30	40	60	80	100
10	30.7						
20	34.7	36.0					
30	30.7	38.7	36.0				
40	26.7	28.0	38.7	36.0			
60	36.0	41.3	33.3	38.7	36.0		
80	37.3	36.0	34.7	38.7	32.0	37.3	
100	38.7	34.7	36.0	33.3	32.0	37.3	30.7

$\mu=0.01, v=0.05$							
H1 / H2	10	20	30	#####	60	80	100
10	25.3						
20	34.7	38.7					
30	29.3	33.3	33.3				
40	36.0	38.7	32.0	37.3			
60	33.3	33.3	38.7	38.7	40.0		
80	37.3	30.7	29.3	41.3	37.3	41.3	
100	37.3	29.3	38.7	36.0	40.0	40.0	36.0

$\mu=0.001, v=0$							
H1 \ H2	10	20	30	40	60	80	100
10	49.0						
20	40.0	40.0					
30	48.0	49.0	49.0				
40	48.0	49.0	49.0	48.0			
60	49.0	49.0	32.0	48.0	32.0		
80	49.0	48.0	44.0	48.0	40.0	49.0	
100	44.0	48.0	49.0	40.0	49.0	48.0	48.0

$\mu=0.001, v=0.05$							
H1 / H2	10	20	30	40	60	80	100
10	41.3						
20	43.0	47.3					
30	46.7	49.0	47.0				
40	44.0	46.0	47.3	48.3			
60	45.3	48.0	47.0	48.3	47.0		
80	46.7	44.3	48.7	46.0	47.0	48.0	
100	48.7	45.7	47.3	45.7	47.0	45.3	47.0

49 means that the process has been stopped earlier (limit of 50 epochs)

standard deviation of error

$\mu=0.01, v=0$							
H1 / H2	10	20	30	40	60	80	100
10	0.003						
20	0.002	0.001					
30	0.001	0.001	0.001				
40	0.002	0.000	0.001	0.001			
60	0.001	0.001	0.001	0.001	0.001		
80	0.000	0.002	0.001	0.002	0.002	0.001	
100	0.001	0.001	0.000	0.001	0.001	0.001	0.001

$\mu=0.01, v=0.05$							
H1 / H2	10	20	30	40	60	80	100
10	0.010						
20	0.001	0.001					
30	0.001	0.001	0.002				
40	0.001	0.001	0.000	0.001			
60	0.002	0.001	0.000	0.001	0.002		
80	0.000	0.002	0.003	0.001	0.000	0.001	
100	0.002	0.001	0.001	0.001	0.000	0.001	0.001

$\mu=0.001, v=0$							
H1 / H2	10	20	30	40	60	80	100
10	NC						
20	NC	NC					
30	NC	NC	NC				
40	NC	NC	NC	NC			
60	NC	NC	NC	NC	NC		
80	NC	NC	NC	NC	NC	NC	
100	NC	NC	NC	NC	NC	NC	NC

$\mu=0.001, v=0.05$							
H1 / H2	10	20	30	40	60	80	100
10	0.002						
20	0.001	0.004					
30	0.003	0.002	0.001				
40	0.003	0.003	0.001	0.001			
60	0.000	0.000	0.003	0.000	0.000		
80	0.004	0.004	0.003	0.003	0.001	0.002	
100	0.002	0.003	0.002	0.002	0.002	0.001	0.001