

Adagrad

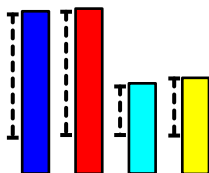
Constant

Iteration 0

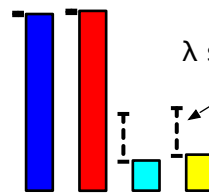


λ scaled up

Iteration 5000

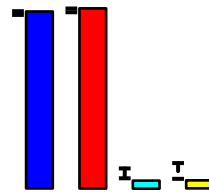
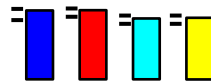


Iteration 10000



λ scaled down

Iteration 20000



λ of bottom words



λ of top words