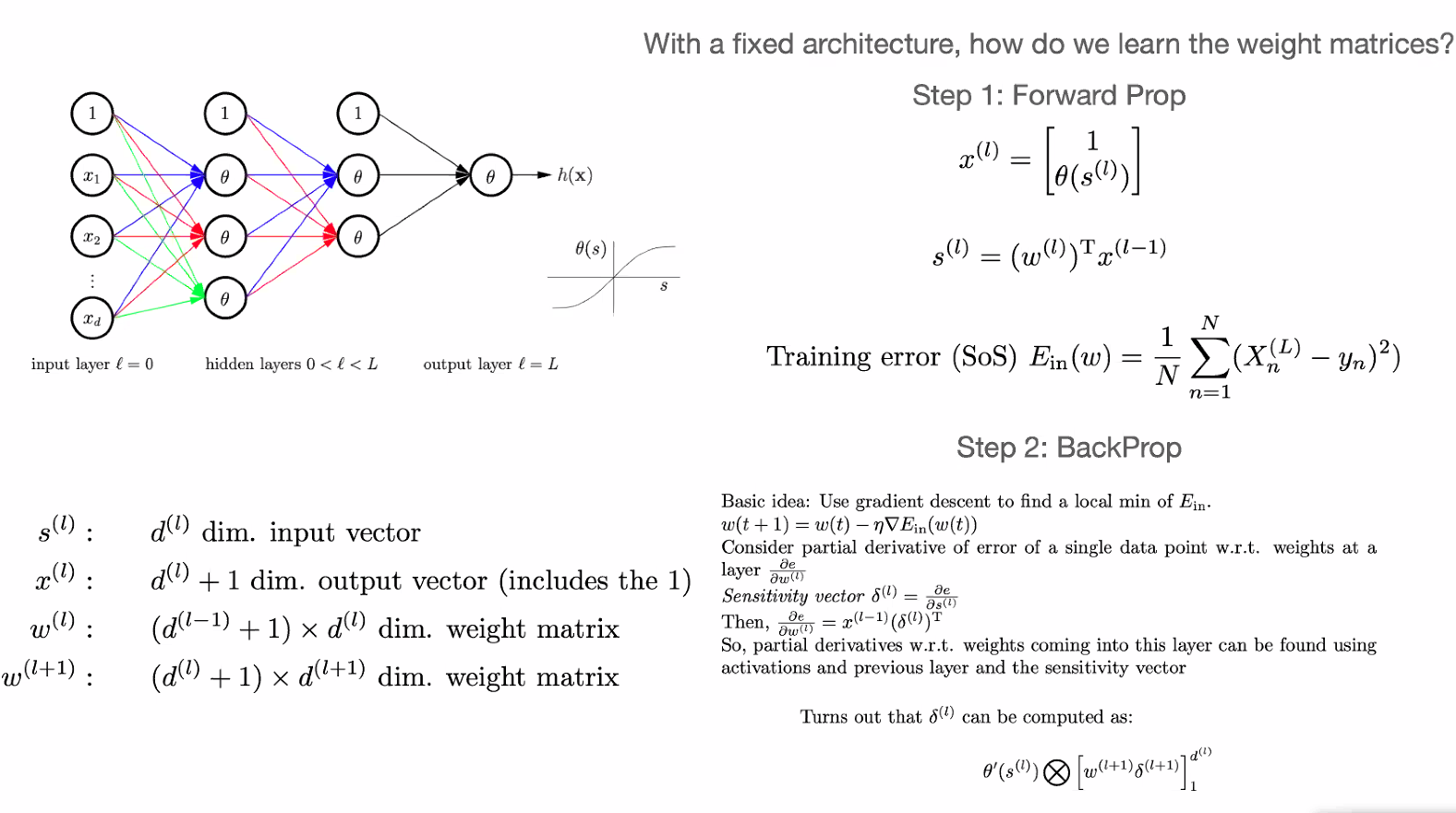
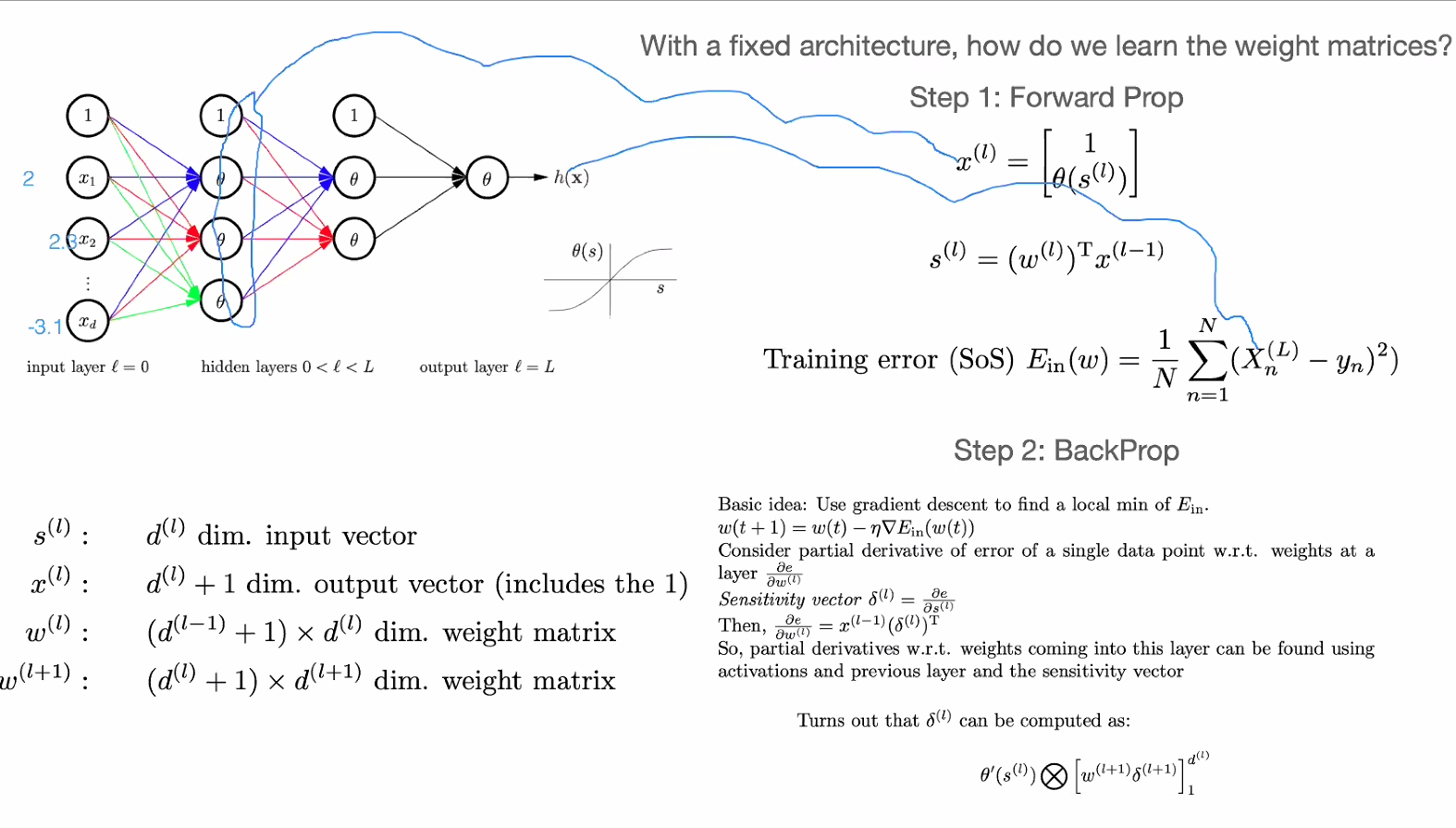
Multi-Layer neural network



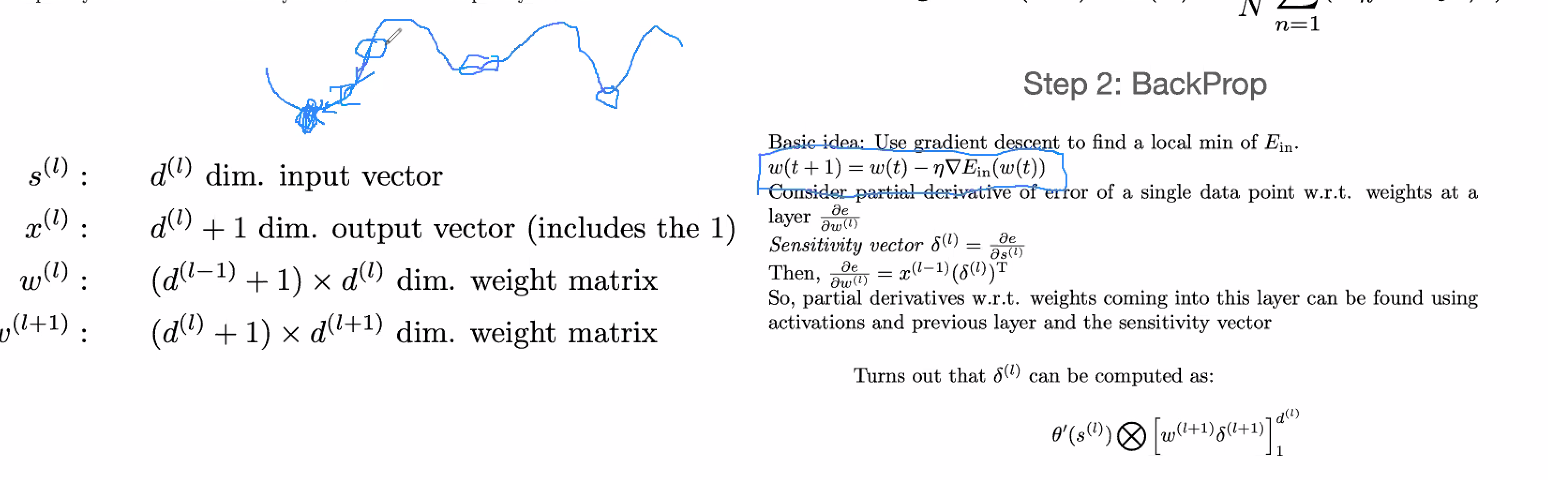
Input layer = feature vertex.

Theta function = hidden layers activate a function.

Gradient decent y-axis training error x-axis weight, goal minimize train error, use calculus. LocalMinmax.



Gradient decent:



Sensitivity vector:  
input and output of each layer, partial derivative.

