Christopher Morgan
CS478 : Brother Christophe
Sep 25, 2013

1. Consider the following simple dataset.

A	B	T
1	0	1
0	1	0

Show your results in the form of a table as we did in class.

a	b	W(A-h)	W(B-h)	W(h-T)	h	t	Target t	E(t)	dW(h-t)	E(h)	dW(A-h)	dW(B-h)	C
Init		0.100000	0.100000	0.100000									0.3000
1.000000	0.000000	0.100000	0.100000	0.100000	0.524979	0.513121	1.000000	0.12164	0.01916	0.00303	0.00091	0.00000	
Update Weights		0.100910	0.100000	0.119157									
0.000000	1.000000	0.100910	0.100000	0.119157	0.524979	0.515634	0.000000	-0.12878	-0.02028	-0.00383	0.00000	-0.00115	
Update Weights		0.100910	0.098852	0.098874									
1.000000	0.000000	0.100910	0.098852	0.098874	0.525206	0.512979	1.000000	0.12167	0.01917	0.00300	0.00090	0.00000	
Update Weights		0.101810	0.098852	0.118045									
0.000000	1.000000	0.101810	0.098852	0.118045	0.524693	0.515479	0.000000	-0.12875	-0.02027	-0.00379	0.00000	-0.00114	
Update Weights		0.101810	0.097715	0.097780									
1.000000	0.000000	0.101810	0.097715	0.097780	0.525431	0.512841	1.000000	0.12171	0.01918	0.00297	0.00089	0.00000	
Update Weights		0.102700	0.097715	0.116965									
0.000000	1.000000	0.102700	0.097715	0.116965	0.524409	0.515330	0.000000	-0.12871	-0.02025	-0.00375	0.00000	-0.00113	

3. Assume that the units of a neural network are modified so they compute the squashing function tanh (instead of the sigmoid function). What is the resulting backpropagation weight update rule for the output layer? (Note, tanh'(x) = 1 – tanh^2 (x)).

$$\Delta W_{jk} = C * O_j * (T_k - O_k)(1 - \tanh^2(x))$$