Introduction to Neural Networks

Johns Hopkins University
Engineering for Professionals Program

605-447/625-438

Dr. Mark Fleischer

Module 4.2: MOSD Example

This Sub-Module Covers …

- A closer look at the Method of Steepest Descent.
- Provides an example of how it performs using an Excel spreadsheet.
- Illustrates various performance issues.

A Quick Rehash

$$f(x) = f(x_0) + f'(x_0)(x - x_0)$$

First 2 terms of Taylor Series

$$f(x_{k+1}) = f(x_k) + f'(x_k)(x_{k+1} - x_k)$$

Substitute x with x_{k+1} and x_0 with x_k

$$0 = f(x_k) + f'(x_k)(x_{k+1} - x_k)$$

Setting $f(x_{k+1}) = 0$

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Rearranging and solving for x_{k+1}

$$x_{k+1} = x_k - \boxed{\frac{f(x_k)}{\left[f'(x_k)\right]^2}} f'(x_k)$$

Equivalently

$$x_{k+1} = x_k - \boxed{\eta} f'(x_k)$$

Method of Steepest Descent

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta \nabla E \quad \Rightarrow \quad w_{ij}(k+1) = w_{ij}(k) + \eta e_j \left[1 - y_j\right] y_j x_i$$

The Perceptron Delta Method

The Perceptron Delta Function

Putting it all together,

$$\frac{\partial E}{\partial \omega_{ij}} = \boxed{-e_j[1 - y_j]y_j} x_i$$

and letting $\delta_j = -e_j[1 - y_j]y_j$ then

$$\boxed{\Delta\omega_{ij} = \eta \frac{\partial E}{\partial \omega_{ij}} = \eta \delta_j x_i.}$$

Let’s Use the PDF in Excel

Desired Output	0.9														
eta	0.2														
x1	1														
x2	2														
w1	0.3	0.3	0.30777	0.315091	0.322002	0.328538	0.334729	0.340603	0.346184	0.351496	0.356559	0.36139	0.366007	0.370424	0.374654
w2	0.3	0.3	0.31554	0.330183	0.344005	0.357076	0.369457	0.381205	0.392369	0.402993	0.413118	0.422781	0.432014	0.440847	0.449308
Activity		0.9	0.93885	0.975456	1.010012	1.04269	1.073644	1.103013	1.130922	1.157482	1.182795	1.206952	1.230034	1.252118	1.273269
f(x)		0.71095	0.718867	0.726206	0.733023	0.739369	0.745289	0.750824	0.756009	0.760875	0.76545	0.769759	0.773825	0.777666	0.781302
error		-0.18905	-0.18113	-0.17379	-0.16698	-0.16063	-0.15471	-0.14918	-0.14399	-0.13913	-0.13455	-0.13024	-0.12618	-0.12233	-0.1187
delta_i		-0.03885	-0.03661	-0.03456	-0.03268	-0.03095	-0.02937	-0.02791	-0.02656	-0.02531	-0.02416	-0.02308	-0.02208	-0.02115	-0.02028

w1	0.37871	0.382604	0.386345	0.389943	0.393407	0.396744	0.399963	0.403068	0.406068	0.408967	0.41177	0.414484	0.417112	0.419658
w2	0.457421	0.465208	0.47269	0.479887	0.486814	0.493489	0.499925	0.506137	0.512135	0.517933	0.523541	0.528968	0.534223	0.539316
Activity	1.293551	1.31302	1.331726	1.349717	1.367036	1.383722	1.399813	1.415341	1.430338	1.444833	1.458852	1.472419	1.485558	1.49829
f(x)	0.784748	0.788018	0.791126	0.794083	0.796901	0.799588	0.802154	0.804607	0.806954	0.809202	0.811357	0.813425	0.815411	0.817319
error	-0.11525	-0.11198	-0.10887	-0.10592	-0.1031	-0.10041	-0.09785	-0.09539	-0.09305	-0.0908	-0.08864	-0.08658	-0.08459	-0.08268
delta_i	-0.01947	-0.01871	-0.01799	-0.01732	-0.01669	-0.01609	-0.01553	-0.015	-0.01449	-0.01402	-0.01357	-0.01314	-0.01273	-0.01234

0.71095	
0.71095	
0.8	0.8
0.8	0.8
	1.137519
	0.757224
	0.142776