The Johns Hopkins University

JHU ENGINEERING FOR PROFESSIONALS PROGRAM

NEURAL NETWORKS: 625-438.71/605-447.71

Problem Set #2

2.1 Show that for $f(x) = \dfrac{1}{1+e^{-x}}$, $\dfrac{df(x)}{dx} = f'(x) = f(x)\cdot(1 - f(x))$.

Ans:

$$f(x) = \frac{1}{1+e^{-x}} = (1+e^{-x})^{-1}$$

$$f'(x) = \frac{df}{dx} = -(1+e^{-x})^{-2}(e^{-x})(-1)$$

$$= f^2(e^{-x}) \qquad 1+e^{-x} = \frac{1}{f}$$

$$= f^2\,\frac{1-f}{f} \qquad \Rightarrow e^{-x} = \frac{1}{f} - 1 = \frac{1-f}{f}$$

$$= f(1-f)$$

2.2 For $f(x,y,z) = 3x^2 + 2xy + 3xz + 4yz + y^2 + xyz + 2z^3$, find the expressions for each of $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}, \dfrac{\partial f}{\partial z}$.

Ans:

$$\frac{\partial F}{\partial x} = 6x + 2y + 3z + yz$$

$$\frac{\partial F}{\partial y} = 2x + 4z + 2y + xz$$

$$\frac{\partial F}{\partial z} = 3x + 4y + xy + 6z^2$$

2.3 Let $I_j = \sum_{i=1}^{9} w_{ji}\theta_i$ for $j = 1,\ldots,4$ and where w_{ji} are the variables.

Determine $\dfrac{\partial I_2}{\partial w_{25}}, \dfrac{\partial I_3}{\partial w_{32}}$, and $\dfrac{\partial I_j}{\partial w_{rs}}$ for $r = j$, and $\dfrac{\partial I_j}{\partial w_{rs}}$ for $r \neq j$.

Ans:

$$\frac{\partial \tilde{I}_2}{\partial w_{25}} = \theta_5$$

$$\frac{\partial \tilde{I}_3}{\partial w_{32}} = \theta_2$$

$$\frac{\partial \tilde{I}_j}{\partial w_{rs}} = \theta_s \quad \text{for} \quad r = j$$

$$\frac{\partial \tilde{I}_j}{\partial w_{rs}} = 0 \quad \text{for} \quad r \neq j$$

2.4 Given the vectors $\mathbf{x} = (1,-1, 0)$, $\mathbf{y} = (3, 1, -2)$, find $|\mathbf{x}|$, $|\mathbf{y}|$, the inner product $\langle \mathbf{x},\mathbf{y} \rangle$, the angle between the vectors $\mathbf{x}$ and $\mathbf{y}$, the outer product $\mathbf{x}^T\mathbf{y}$, and a unit vector in the $\mathbf{x}$ direction.

Ans: |x| = sqrt(2); |y| = sqrt(14); <x,y> = 2;

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}\right) = \frac{2}{\sqrt{2}\sqrt{14}} = 67.7923^\circ$$

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} 3 & 1 & -2 \\ -3 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\hat{\mathbf{x}} = \left(\frac{\sqrt{2}}{2}, \quad \frac{-\sqrt{2}}{2}, \quad 0 \right)$$

2.5 Given the function $f(x, y, z) = 2xy - 3xz^2$, calculate the directional derivative in the direction of point (3,2,1) at the point (1,2,3).

Ans:

$$\nabla f(1,2,3) = (2y - 3z^2, 2x, -6xz) = \left(-23, \quad 2, \quad -18 \right).$$

$$\mathbf{d} = \left(3, \quad 2, \quad 1 \right) - \left(1, \quad 2, \quad 3 \right) = \left(2, \quad 0, \quad -2 \right)$$

$$\therefore \hat{\mathbf{d}} = \left(\frac{\sqrt{2}}{2}, \quad 0, \quad \frac{-\sqrt{2}}{2} \right)$$

$$\therefore \nabla f \cdot \hat{\mathbf{d}} = \left(\frac{-23 \cdot \sqrt{2}}{2} + 0 + 9 \cdot \sqrt{2} \right) = -3.535533...$$

2.6 A plane or hyperplane has the form $w_1x_1 + w_2x_2 + \ldots + w_nx_n = A$ which can be rewritten in vector form as $\mathbf{w} \cdot \mathbf{x} - A = 0$. Prove that the vector **w** is orthogonal (perpendicular) to the hyperplane.

Ans: Define a vector in the hyperplane **P** by defining two points $\mathbf{x_1}$ and $\mathbf{x_2}$ that satisfy the equation $\mathbf{w} \cdot \mathbf{x} - A = 0$ and then taking their difference (i.e., a vector that lies in the hyperplane). Thus, $\mathbf{x}^* = \mathbf{x_2} - \mathbf{x_1}$ is a vector who that lies in the hyperplane **P.** Thus, we can write

$$
\begin{aligned}
\mathbf{w} \cdot \mathbf{x}^* &= \mathbf{w} \cdot (\mathbf{x}_2 - \mathbf{x}_1) \\
&= \mathbf{w} \cdot \mathbf{x}_2 - \mathbf{w} \cdot \mathbf{x}_1 \\
&\text{ and now substituting from the eqn. of the line} \\
&= A - A = 0 \\
&\text{ hence the dot product is zero and } \mathbf{w} \text{ and } \mathbf{x}^* \text{are orthogonal.}
\end{aligned}
$$