Homework 1 for Introduction to Natural Language Processing

Part 1

1

设 $f(x) = ax + b$

则

$g_k(x) = y_k = f(\sum_j (a \sum_i w_{ji} x_i + a w_{j0} + b) + w_{k0})$
$= a \sum_j w_{kj}(a \sum_i w_{ji} x_i + a w_{j0} + b) + a w_{k0} + b$
$= a^2 \sum_j \sum_i w_{kj} w_{ji} x_i + a^2 \sum_j w_{kj} w_{j0} + ab \sum_j w_{kj} + a w_{k0} + b$

2

(1) Write the weight-updating fromula(Denote learning rate as λ)

$z = W^T X$

$\frac{\partial E}{\partial W} = \frac{\partial E}{\partial y} \times \frac{\partial y}{\partial z} \times \frac{\partial z}{\partial W}$
$= -(g - y) \times \frac{1}{e^z + e^{-z} + 2} \times X$
$= -(g - \frac{1}{1 + e^{-W^T X}}) \times \frac{1}{e^{W^T X} + e^{-W^T X} + 2} \times X$
$W_{new} = W + \lambda(g - \frac{1}{1 + e^{-W^T X}}) \times \frac{1}{e^{W^T X} + e^{-W^T X} + 2} \times X$

(2)

$W = (0.5, 1, 1)^T$
$X = (1, 2, 0.5)^T, g = 1, \lambda = 0.1$
$W_{new} = W + \lambda(g - \frac{1}{1 + e^{-W^T X}}) \times \frac{1}{e^{W^T X} + e^{-W^T X} + 2} \times X$
$= (0.5, 1, 1)^T + 0.1 \times (1 - \frac{1}{1 + e^{-3}}) \times \frac{1}{e^3 + e^{-3} + 2} \times (1, 2, 0.5)^T$
$\approx (0.5002142542, 1.0004285084, 1.0001071271)$