**HOMEWORK03**

**Problem 2:**

Given,

() = 1/(1+e-)

P(y = 1|x) = (wTx)

P(y = -1|x) = 1 - (wTx)

P(y|x) = P(y = 1|x)(y+1)/2 P(y = -1|x)(1-y)/2

l() = logP(yn|xn) = log P(yn|xn)

(a) Show that

|  |
| --- |
| (d/d) = () (1 - ()) |

() = (1+e-)-1 | Let, 1+e- = t

⇒ = t-1 | ⇒ e- = t - 1

⇒ d() = -t-2. dt | ⇒ - = ln (t - 1)

| ⇒ = ln (1/(t - 1))

Substituting the values of ‘t’ and ‘dt’, |

|

⇒ d() = -(1+e-)-2 \* (-e-) d | Derivating 1+e- = t w.r.t t

| ⇒ dt = (-1)(e-)d

Substituting (1+e-)-1 =() |

|

⇒ d() = ()2 \* (()-1 - 1) d |

⇒ (d/d) = ()2 \* (1 - ())/() |

⇒ **(d/d) = () (1 - ())**

(b) Derive the gradient of log-likelihood,

l() = logP(yn|xn) = log P(yn|xn)

⇒ log [P(y = 1|x)(y+1)/2 P(y = -1|x)(1-y)/2 ]

⇒ log [P(y = 1|x)(y+1)/2 ] + log [P(y = -1|x)(1-y)/2 ]

⇒ ()y=1 log P(y = 1|x) + ()y=-1 log P(y = -1|x)

⇒ ()y=1 log (Tx) + ()y=-1 log [1 - (Tx)]

⇒ ()y=1 log (1+e-)-1 + ()y=-1 log [e-/(1+e-)]

⇒ -log (1+e-) + log e- - log (1+e-)

⇒ log [(e-) \* (1+e-)-2]

This looks similar to

⇒ log [d(Tx)/d]

Now, Gradient (delta) of l() is

⇒ []

⇒

(c) Upgrade step for gradient ascent of l() is nothing but a local maxima point for l()

To find out, we’ve to derivate the result of (b) and equate it to 0