To achieve the best result; I decide to use hand to write down the written assignment part (Typing in math symbols and Equations in Latex really interrupts my logic)

Shirui le CS542 5.3 P(t1x, W) TN(tn/y(xnw), E) (Used Gaussion tan In/y(xnw) is u, z is 62) To mak Tinto E, take the In mIP(t(X,W,E) = -1 & (tn-yn) TE -1tn-yn) - N(ln 121+km (2711) constant we can ignore it won't affect maximization we used accussion, and in the law In = y(xn, w), E(w)= 立製(tn-yn) を1(tn-yn)- 立型(tn-yn) を1(tn-yn) を加了 =- 2Hz-12/th-yn)ttn-yn)TJ-N/n/z1 To maximize, take derivate of 5-1

The maximum likelihood is $\Sigma = \frac{1}{N} \sum_{n=1}^{N} (t_n - y_n) t_n - y_n) T$

Shirui Ye

CS 542

PS3

5.4 target values are $t \in \{0,1\}$.

Let $t \in \{0,1\}$ to be time class tabeling.

According to question, we have a network output y(x, w) that represents P(t=|x|) we can get below:

P($t=|x| \ge E(|-y(x,w)| + (|-E|y(x,w)|)$)

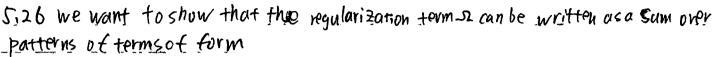
Then we can get conditional prob P(t|x) to be below:

P(t|x) = (t=|x|) - t=|x|) - t=|x|Finally, the error function (a Wesponding to the negative log || Welthwood | S|)

Ewil = t=|x| (In II- elly (Xn, w) - | In Ely(Xn, w) > 1. (In II) + thin I elly (Xn, w) + | In Elly (Xn, w) > 1.

Thus, the overall likelihood to the emer function is obtained when 6=0.

Shirui Ye CS542 PS3



 $\Omega_{n=\frac{1}{2}} \sum_{k} (Gy_{k})^{2} \qquad \Omega_{n=\frac{1}{2}} \sum_{k} (\sum_{n=1}^{\infty} \sum_{n=1}^{\infty} \sum_{n=1}$

use Jacobian Matrix Iki = 34k

-On= = = = = = [= Ini Jnki)2

Next by acting forward propagation equations, with G, we want to show an can be evaluated by using equations 5.204

Applying $G = \{i \}_{i=1}^{\infty}$ in $\{i \}_{i=1}^{\infty}$ $\{i \}_{i=1}^{\infty}$, and we have $\{i \}_{i=1}^{\infty}$, $\{i \}_{i=1}^{\infty}$, we can get $\{i \}_{i=1}^{\infty}$ $\{i \}_{i=1}^{\infty}$, $\{i \}_{i=1}^$

According to these 2, We can get Bri = Ewith ni = Ewit Trai axni = Ewithi

Finally, we want to show that the derivatives of In with respect to a weight was can be written

in the form $\frac{\partial Dn}{\partial ws} = \frac{\partial P}{\partial ws} (\Phi_{Kr}^2 + \delta_{Kr} ds)$

To do so, we use chain rule, $\Omega_n = \frac{1}{2} \sum_{k} (94k)^2$, $\theta_j = 92j$, $\theta_j = 90j$ and $\theta_k = \frac{34k}{34k}$, $\theta_k = 98kr + \frac{34j}{34k}$

We canget 32n = & dnx (pnxxx x + snk dns)

According to Byk = hilai) & Windyk, we know sukr = h'(anr) & war Snkl

Now, we can get backpropagation of Quer, by using above with $dj = g_{\overline{z}j}$, Bj = gaj $Skr = \frac{34k}{3ar}, \phi kr = 45kr$

-Pnyr=gsnkr= h'(ank) Zwirpnki+h"(ank)Brizwir Snke