I. Lagaris, A. Likas, D.I. Fotiadis
IEEE Trans. on Neural Networks
Vol. 9, No. 5, Sept. 1998, pp. 981-1000

EE434
03/05/08

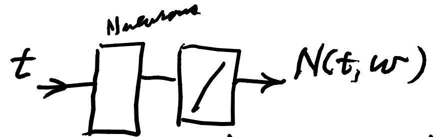
$\frac{dx}{dt} = \dot{x} = f(t, x)$; $x(0)$ they take x to be a scalar

$x_n(t) = x(0) + t N(t, w)$ w = weights

$@ t = 0, x_n(t) = x(0)$

$\parallel$ network output

$N(t,w) =$ 1 hidden with activation & a linear output layer



$\alpha(w_H)$ $\uparrow$ w_0 $\Rightarrow \sum_{j=1}^{n=1} w_{ij} \cdot t_j = z_i$, $i = 1, \ldots, H$

$\parallel$

$\sigma(z_i)$; $z = w_H \cdot t + u$; $N(t,w) = w_0 \cdot \sigma = \sum_{i=1}^{H} v_i \sigma(z_i)$

$$\frac{dx_n}{dt} = N(t,w) + t\frac{dN(t,w)}{dt} \quad ; \quad \frac{dN(t,w)}{dt_j} = \sum_{i=1}^{H} v_i w_{ij} \sigma_i^{(1)} = N_g(t,w)$$

$$\frac{dN(t,w)}{dt} = \frac{d w_0 \sigma}{dt} = w_0 \frac{d\sigma}{dt} = w_0 \frac{\partial \sigma(z)}{\partial z} \cdot \frac{\partial z}{\partial t} = w_0 \sigma^{(1)}\left[w_H + \frac{du}{dt}\right]$$

Their example:

$$\frac{dx}{dt} + \left(t + \frac{1+3t^2}{1+t+t^3}\right)x = t^3 + 2t + t^2\left(\frac{1+3t^2}{1+t+t^2}\right)$$

desire gradient with respect to weights of the error squared

$$E(w) = \sum_{k=1}^{m} \left(x(t_k) - x_n(t_k)\right)^2 = E(W) + \frac{\partial E}{\partial w} \cdot (W - w) + \frac{1}{2}\frac{\partial^2 E}{\partial w^2}(W-w)^2$$

gradient; expanding about; Hessian

Taking the gradient is not bad

$$w(l) = w(l-1) + \underline{\Delta w(l)} + \beta \frac{\partial^2 E}{\partial w^2} \text{ at } l\text{th iteration on weights}$$

$$-\alpha \frac{\partial E}{\partial w} \quad ; \quad 0 \le \alpha \le 1 \qquad 0 \le \beta \le 1$$

choose the negative of the gradient for weight update