Stable Learning

Wen Yu

https://www.ctrl.cinvestav.mx/~yuw/
yuw@ctrl.cinvestav.mx

Stable learning

Nonlinear system

$$y(k) = f[X(k)]$$

The neural networks is

$$\hat{y}(k) = W_k \phi [X(k)]$$

If the identified nonlinear system can be represented as (matching condition)

$$y(k) = W^* \phi [X(k)]$$

The training law is

$$W_{k+1} = W_k - \eta \phi e(k)$$

where

$$e(k) = \hat{y}(k) - y(k)$$

Error dynamic

$$\hat{y}(k) = W_k \phi$$
$$y(k) = W^* \phi$$

Modeling error is

$$e(k) = \hat{y}(k) - y(k)$$
$$= [W_k - W^*] \phi$$
$$= \tilde{W}_k \phi$$

where

$$\tilde{W}_k = W_k - W^*$$

Lyapunov

We select Lyapunov function as

$$L_k = \left\| \tilde{W}_k \right\|^2 = tr\left\{ \tilde{W}_k^T \tilde{W}_k \right\}$$

Becasue

$$\tilde{W}_{k+1} = \tilde{W}_k - \eta_k \phi e(k)$$

From the updating law

$$\begin{aligned}
\Delta L_k = L_{k+1} - L_k &= \left\| \tilde{W}_k - \eta_k \phi e(k) \right\|^2 - \left\| \tilde{W}_k \right\|^2 \\
&= \left\| \tilde{W}_k \right\|^2 - 2\eta_k e(k) \phi \tilde{W}_k + \left\| \eta_k \phi e(k) \right\|^2 - \left\| \tilde{W}_k \right\|^2 \\
&= \eta_k^2 e^2(k) \left\| \phi \right\|^2 - 2\eta_k e(k) \phi \tilde{W}_k \\
&= \eta_k^2 e^2(k) \left\| \phi \right\|^2 - 2\eta_k e^2(k) \\
&= \left[\eta_k \left\| \phi \right\|^2 - 2 \right] \eta_k e^2(k)
\end{aligned}$$

We select

$$\eta_k = \frac{2\eta_0}{1 + \|\phi\|^2}, \qquad 0 \leq \eta_0 \leq 1$$

then

$$\left[\eta_k \|\phi\|^2 - 2\right] = 2\left[\frac{\|\phi\|^2}{1 + \|\phi\|^2}\eta_0 - 1\right] \leq 2\left[\eta_0 - 1\right] \leq 0$$

so

$$\Delta L_k \leq 0$$

Unmodeled dynamic

Nonlinear system

$$y(k) = f[X(k)]$$

The neural networks is

$$\hat{y}(k) = W_k \phi [X(k)]$$

If the identified nonlinear system can be represented as

$$y(k) = W^* \phi + \xi(k), \qquad |\xi|^2 \leq \bar{\xi}$$

Unmodeled dynamic

The error dynamicr is

$$
\begin{aligned}
e(k) &= \hat{y}(k) - y(k) \\
&= W_k\phi - W^*\phi - \xi(k) \\
&= \tilde{W}_k\phi - \xi(k)
\end{aligned}
$$

Lyapunov

$$
\begin{aligned}
\Delta L_k &= \eta_k^2 e^2(k) \|\phi\|^2 - 2\eta_k e(k)\phi\tilde{W}_k \\
&= \eta_k^2 e^2(k) \|\phi\|^2 - 2\eta_k e(k)\left[e(k) + \xi(k)\right] \\
&= \left[\eta_k \|\phi\|^2 - 2\right]\eta_k e^2(k) - 2\eta_k e(k)\xi(k)
\end{aligned}
$$

Here

$$
-2\eta_k e(k)\xi(k) \le \eta_k e^2(k) + \eta_k \xi^2(k), \qquad \eta_k > 0
$$

Unmodeled dynamic

So

$$\Delta L_k = \left[\eta_k \left\|\phi\left[X\left(k\right)\right]\right\|^2 - 1\right] \eta_k e^2\left(k\right) + \eta_k \xi^2\left(k\right)$$

We select

$$\eta_k = \frac{\eta_0}{1 + \left\|\phi\left[X\left(k\right)\right]\right\|^2}, \qquad 0 \leq \eta_0 \leq 1$$

then

$$\begin{array}{c} \left[\eta_k \left\|\phi\right\|^2 - 1\right] \eta_k \\ = \left[\frac{\left\|\phi\right\|^2}{1+\left\|\phi\right\|^2} \eta_0 - 1\right] \eta_k \\ \leq \left(\eta_0 - 1\right) \eta_k = -\pi_k \leq 0 \end{array}$$

where

$$\pi_k = \left(1 - \eta_0\right) \eta_k$$

Unmodeled dynamic

So

$$\Delta L_k \leq -\pi_k e^2(k) + \bar{\bar{\xi}} \leq -[1-\eta_0]\,\bar{\eta} e^2(k) + \bar{\bar{\xi}}$$

where

$$\begin{aligned} \pi_k &= \left[1 - \frac{\|\phi\|^2}{1+\|\phi\|^2}\eta_0\right]\eta_k \\ &\geq [1-\eta_0]\,\eta_k \\ &\geq [1-\eta_0]\,\bar{\eta} \end{aligned}$$

here

$$\eta_k = \frac{\eta_0}{1+\|\phi[X(k)]\|^2} \geq \frac{\eta_0}{1+\bar{\phi}} = \bar{\eta}, \qquad \bar{\phi} = \max\left[\|\phi[X(k)]\|^2\right]$$

Dead-zone learning

$$W_{k+1} = W_k - s_k \eta_k \phi e(k)$$

where

$$\eta_k = \frac{\eta_0}{1 + \|\phi\|^2}, \qquad 0 \leq \eta_0 \leq 1$$

$$s_k = \begin{cases} 1 & \text{if } e^2(k) > \frac{\bar{\xi}}{[1-\eta_0]\bar{\eta}} \\ 0 & \text{if } e^2(k) \leq \frac{\bar{\xi}}{[1-\eta_0]\bar{\eta}} \end{cases}$$

Becasue

$$\Delta L_k \leq -[1 - \eta_0]\, \bar{\eta} e^2(k) + \bar{\xi}$$

- If $e^2(k) > \frac{\bar{\xi}}{[1-\eta_0]\bar{\eta}}$, then $\Delta V_k \leq 0$
- If $e^2(k) \leq \frac{\bar{\xi}}{[1-\eta_0]\bar{\eta}}$, $W_{k+1} = W_k$, then $\Delta V_k = 0$

Nonlinear active function

Nonlinear system

$$y(k) = f[X(k)]$$

The neural networks is

$$\hat{y}(k) = \phi[W_k X(k)]$$

We use Talyor seris

$$\begin{aligned}\phi[W_k X(k)] &= \phi\left(W^0 X\right) + \phi'\left(W_k - W^0\right) X(k) + \mathcal{O}\left(\tilde{W}\right) \\ &= \phi\left(W^0 X\right) + \phi'\left(W_k - W^0\right) X(k) + \delta(k)\end{aligned}$$

The identified nonlinear system can be represented as

$$y(k) = \phi\left[W^0 X(k)\right] - \xi(k), \qquad |\xi|^2 \leq \bar{\xi}$$

Nonlinear active function

Error dynamic is

$$
\begin{aligned}
& e(k) = \hat{y}(k) - y(k) = \phi(W_k X) - \phi\left(W^0 X\right) + \xi(k) \\
& = \phi' \left(W_k - W^0\right) X(k) + \delta(k) + \xi(k) \\
& = \phi' \tilde{W}_k X(k) + \delta(k) + \xi(k)
\end{aligned}
$$

Training law should be

$$
W_{k+1} = W_k - \eta_k \phi' X(k) e(k)
$$

Becasue

$$\tilde{W}_{k+1} = \tilde{W}_k - \eta_k \phi' X(k) e(k)$$
$$e(k) = \phi' \tilde{W}_k X(k) + \delta(k) + \xi(k)$$

Lyapunov

$$\begin{aligned}
&\Delta L_k = \eta_k^2 e^2(k) \|\phi' X(k)\|^2 - 2\eta_k e(k) \phi' \tilde{W}_k X(k) \\
&= \eta_k^2 e^2(k) \|\phi' X(k)\|^2 \\
&-2\eta_k e(k) [e(k) - \delta(k) - \xi(k)] \\
&= \left[\eta_k \|\phi' X(k)\|^2 - 2\right] \eta_k e^2(k) \\
&+2\eta_k e(k) [\delta(k) + \xi(k)]
\end{aligned}$$

Here

$$2\eta_k e(k) [\delta(k) + \xi(k)] \leq \eta_k e^2(k) + \eta_k \left(\bar{\xi} + \bar{\delta}\right), \qquad \eta_k > 0$$

Nonlinear active function

So

$$\Delta L_k \leq -\left[1-\eta_0\right]\bar{\eta} e^2\left(k\right) + \left(\bar{\xi}+\bar{\delta}\right)$$

Then the training law is

$$W_{k+1} = W_k - s_k \eta_k \phi' X\left(k\right) e\left(k\right)$$

where

$$\eta_k = \frac{\eta_0}{1+\left\|\phi\right\|^2}, \qquad 0 \leq \eta_0 \leq 1$$

$$s_k = \begin{cases} 1 & \text{if } e^2\left(k\right) > \frac{\bar{\xi}+\bar{\delta}}{\left[1-\eta_0\right]\bar{\eta}} \\ 0 & \text{if } e^2\left(k\right) \leq \frac{\bar{\xi}+\bar{\delta}}{\left[1-\eta_0\right]\bar{\eta}} \end{cases}$$

$\bar{\eta} = \frac{\eta_0}{1+\bar{\phi}}$, $\bar{\phi} = \max\left[\left\|\phi\left[X\left(k\right)\right]\right\|^2\right]$

MLP

Nonlinear system

$$y(k) = f[X(k)]$$

MLP neural networks is

$$\hat{y}(k) = W_k \phi [V_k X(k)]$$

The identified nonlinear system can be represented as

$$y(k) = W^0 \phi \left[V^0 X(k) \right] - \xi(k), \qquad |\xi|^2 \leq \bar{\xi}$$

We use Talyor seris

$$\phi[V_k X(k)] = \phi\left(V^0 X\right) + \phi' \left(V_k - V^0\right) X(k) + \delta(k)$$
$$\phi[V_k X(k)] - \phi\left(V^0 X\right) = \phi' \tilde{V}_k X(k) + \delta(k)$$

MLP

The error dyanmic is

$$
\begin{aligned}
& e(k) = W_k \phi\left[V_k X(k)\right] - W^0 \phi\left[V^0 X(k)\right] + \xi(k) \\
& = W_k \phi\left[V_k X(k)\right] - W^0 \phi\left[V^0 X(k)\right] \\
& + W^0 \phi\left[V_k X(k)\right] - W^0 \phi\left[V_k X(k)\right] + \xi(k) \\
& = \tilde{W}_k \phi\left[V_k X(k)\right] + W^0\left[\phi\left[V_k X(k)\right] - \phi\left[V^0 X(k)\right]\right] + \xi(k) \\
& = \tilde{W}_k \phi\left[V_k X(k)\right] + W^0 \phi' \tilde{V}_k X(k) + \delta(k) + \xi(k)
\end{aligned}
$$

Training law should be

$$
\begin{aligned}
W_{k+1} &= W_k - \eta_k \phi\left[V_k X(k)\right] e(k) \\
V_{k+1} &= V_k - \eta_k W^0 \phi' X(k)\, e(k)
\end{aligned}
$$

MLP

Lyapunov

$$L_k = \left\| \tilde{W}_k \right\|^2 + \left\| \tilde{V}_k \right\|^2$$

$$\begin{aligned}
\Delta L_k &= \left\| \tilde{W}_{k+1} \right\|^2 + \left\| \tilde{V}_{k+1} \right\|^2 - \left(\left\| \tilde{W}_k \right\|^2 + \left\| \tilde{V}_k \right\|^2 \right) \\
&= \left\| \tilde{W}_k - \eta_k \phi \left[V_k X \left(k \right) \right] e \left(k \right) \right\|^2 + \left\| \tilde{V}_k - \eta_k W^0 \phi' X \left(k \right) e \left(k \right) \right\|^2 \\
&- \left(\left\| \tilde{W}_k \right\|^2 + \left\| \tilde{V}_k \right\|^2 \right) \\
&= \left\| \eta_k \phi \left[V_k X \left(k \right) \right] e \left(k \right) \right\|^2 + \left\| \eta_k W^0 \phi' X \left(k \right) e \left(k \right) \right\|^2 \\
&- 2\eta_k \phi \left[V_k X \left(k \right) \right] e \left(k \right) \tilde{W}_k - 2\eta_k W^0 \phi' X \left(k \right) e \left(k \right) \tilde{V}_k \\
&= \left\| \eta_k \phi \left[V_k X \left(k \right) \right] e \left(k \right) \right\|^2 + \left\| \eta_k W^0 \phi' X \left(k \right) e \left(k \right) \right\|^2 \\
&- 2\eta_k \left\{ \tilde{W}_k \phi \left[V_k X \left(k \right) \right] + W^0 \phi' \tilde{V}_k X \left(k \right) \right\} e \left(k \right)
\end{aligned}$$

MLP

Because

$$\tilde{W}_k\phi\left[V_kX\left(k\right)\right]+W^0\phi'\tilde{V}_kX\left(k\right)=e\left(k\right)-\delta\left(k\right)-\xi\left(k\right)$$

$$\begin{array}{l}\Delta L_k=\left\|\eta_k\phi\left[V_kX\left(k\right)\right]e\left(k\right)\right\|^2+\left\|\eta_kW^0\phi'X\left(k\right)e\left(k\right)\right\|^2\\-2\eta_k\left\{e\left(k\right)-\delta\left(k\right)-\xi\left(k\right)\right\}e\left(k\right)\\=-\eta_k\left[2-\eta_k\left(\left\|\phi V_kX\left(k\right)\right\|^2+\left\|W^0\phi'X\left(k\right)\right\|^2\right)\right]e^2\left(k\right)\\+2\eta_ke\left(k\right)\left[\delta\left(k\right)+\xi\left(k\right)\right]\end{array}$$

Let

$$\eta_k=\frac{\eta_0}{1+\left\|\phi V_kX\left(k\right)\right\|^2+\left\|W^0\phi'X\left(k\right)\right\|^2},\qquad 1\geq\eta_0>0$$

Becasue

$$2\eta_k e(k)\left[\delta(k)+\xi(k)\right] \leq \eta_k e^2(k)+\eta_k\left(\bar{\bar{\xi}}+\bar{\delta}\right), \qquad 1 \geq \eta_k > 0$$

So

$$\begin{aligned}
&\Delta L_k \leq \eta_k\left[\eta_k\left(\left\|\phi V_k X(k)\right\|^2+\left\|W^0\phi' X(k)\right\|^2\right)-1\right]e^2(k)\\
&+\eta_k\left(\bar{\bar{\xi}}+\bar{\delta}\right)\\
&\leq \eta_k\left[\frac{\left(\left\|\phi V_k X(k)\right\|^2+\left\|W^0\phi' X(k)\right\|^2\right)}{1+\left\|\phi V_k X(k)\right\|^2+\left\|W^0\phi' X(k)\right\|^2}\eta_0-1\right]e^2(k)+\bar{\bar{\xi}}+\bar{\delta}\\
&\leq -\pi_k e^2(k)+\bar{\bar{\xi}}+\bar{\delta}\\
&\leq -\left[1-\eta_0\right]\bar{\eta}e^2(k)+\bar{\bar{\xi}}+\bar{\delta}
\end{aligned}$$

where

$$\bar{\eta}=\frac{\eta_0}{1+\bar{\phi}}, \bar{\phi}=\max\left[\left\|\phi V_k X(k)\right\|^2+\left\|W^0\phi' X(k)\right\|^2\right]$$

Then the training law is

$$W_{k+1} = W_k - s_k \eta_k \phi \left[V_k X(k) \right] e(k)$$
$$V_{k+1} = V_k - s_k \eta_k W^0 \phi' X(k) e(k)$$

where

$$\eta_k = \frac{\eta_0}{1 + \left\| \phi V_k X(k) \right\|^2 + \left\| W^0 \phi' X(k) \right\|^2}, \qquad 0 \leq \eta_0 \leq 1$$

$$s_k = \begin{cases} 1 & \text{if } e^2(k) > \frac{\bar{\xi} + \bar{\delta}}{[1 - \eta_0] \bar{\eta}} \\ 0 & \text{if } e^2(k) \leq \frac{\bar{\xi} + \bar{\delta}}{[1 - \eta_0] \bar{\eta}} \end{cases}$$

The average of the identification error

Becasue

$$\Delta L_k \leq -\left[1-\eta_0\right]\bar{\eta} e^2\left(k\right) + \bar{\xi} + \bar{\delta}$$

Summarizing from 1 up to T, and by using $L_T > 0$ and L_1 is a constant, we obtain

$$\begin{array}{l} L_T - L_1 \leq -\left[1-\eta_0\right]\bar{\eta}\sum_{K=1}^{T} e^2\left(k\right) + T\left(\bar{\xi}+\bar{\delta}\right) \\ \left[1-\eta_0\right]\bar{\eta}\sum_{K=1}^{T} e^2\left(k\right) \leq L_1 - L_T + T\left(\bar{\xi}+\bar{\delta}\right) \\ \leq L_1 + T\left(\bar{\xi}+\bar{\delta}\right) \end{array}$$

so

$$\frac{1}{T}\sum_{K=1}^{T} e^2\left(k\right) \leq \frac{\bar{\xi}+\bar{\delta}}{\left[1-\eta_0\right]\bar{\eta}} + \frac{1}{T}\frac{1}{\left[1-\eta_0\right]\bar{\eta}} L_1$$

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{k=1}^{T} e^2\left(k\right) \leq \frac{\bar{\xi}+\bar{\delta}}{\left[1-\eta_0\right]\bar{\eta}}$$

Remarks

W^0 does not effect the stability property of the neuro identification, but it influences the identification accuracy.

1. Start from any initial value for $W^0 = W_0$.
2. Do identification with this W^0 until T_0.
3. If the $\|e(T_0)\| < \|e(0)\|$, let W_T as a new W^0, *i.e.*, $W^0 = W_{T_0}$, go to 2 to repeat the identification process.
4. If the $\|e(T_0)\| \geq \|e(0)\|$, stop this off-line identification, now W_{T_0} is the final value for W^0.

With this prior knowledge W^0, we may start the on-line identification .