

УДК 519.71

Т.С. Сакало

Харьковский национальный университет радиоэлектроники

## СИНТЕЗ НЕЙРОСЕТЕВОЙ МУЛЬТИПЛИКАТИВНОЙ МОДЕЛИ ГАММЕРШТЕЙНА ОБЪЕКТА СО МНОГИМИ ВХОДАМИ

Решена актуальная задача построения нелинейной мультипликативной модели Гаммерштейна для динамического объекта со многими входами с использованием радиально-базисной искусственной нейронной сети. Предложена структура специализированной ИНС, решающая данную задачу. Показано, что для обучения данной сети могут применяться алгоритмы, основанные на рекуррентном методе наименьших квадратов.

*ИНС, модель Гаммерштейна, динамический объект, алгоритм, рекуррентный метод*

### Введение

Модель Гаммерштейна наряду с моделью Винера нашла в настоящее время достаточно широкое применение при исследовании нелинейных динамических объектов [1,2]. Обе модели представляют собой последовательно соединенные нелинейные статические и линейные динамические звенья и используют аппроксимацию статической нелинейности, однако в отличие от модели Винера выход модели Гаммерштейна представляет прошедший через линейную динамическую часть сигнал, являющийся нелинейным преобразованием входного сигнала. Это несколько упрощает анализ данной модели, поэтому ее применение оказывается более предпочтительным. Преимущество модели Гаммерштейна становится ощутимым, если исследуемый объект является многомерным, т.е. имеет несколько входных и выходных сигналов. Частным случаем такого объекта является MISO-объект с несколькими входами и одним выходом. Как показывают результаты исследований [1, 4], эффективность применения модели Гаммерштейна может быть существенно увеличена, если для аппроксимации нелинейной статической характеристики использовать не ряды, а искусственную нейронную сеть (ИНС).

Целью данной работы является построение мультипликативной модели Гаммерштейна MISO-объекта с использованием радиально-базисной ИНС.

### 1. Модель Гаммерштейна

Рассмотрим динамический MISO-объект, представленный на рис. 1 и описываемый уравнением:

$$y(k) = \prod_{j=1}^N f_j[y_j(k-1), \dots, y_j(k-m), u_j(k-1), \dots, u_j(k-m)], \quad (1)$$

где  $u(i), y_j(i)$  – выходные сигналы объекта и  $j$ -й модели соответственно в момент времени  $i$ ,  $u_j(i)$  – входной сигнал  $j$ -й модели,  $f_j[\cdot]$  – неизвестная нелинейная функция  $j$ -й модели.

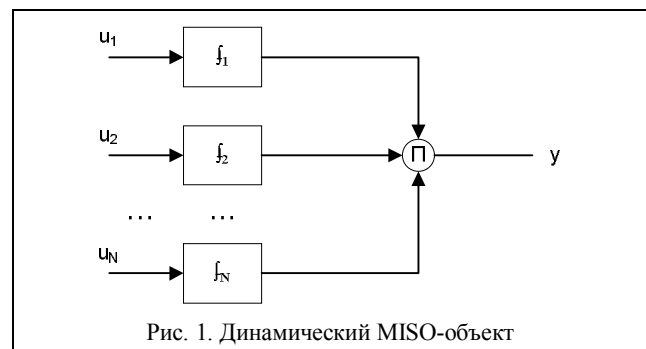


Рис. 1. Динамический MISO-объект

Предположим, что каждый канал может быть представлен моделью Гаммерштейна:

$$x_j(k) = g_j(u_j(k)); \quad (2)$$

$$\hat{y}_j(k) = \sum_{i=1}^m b_{ji}x(k-i) - \sum_{j=1}^m a_{ji}\hat{y}_j(k-i), \quad j = \overline{1, N}, \quad (3)$$

где  $\hat{y}_j(k)$  – выходные сигналы  $j$ -й модели и статической части  $j$ -й модели соответственно;  $g_j(\cdot)$  – нелинейная функция  $j$ -й модели;  $a_{ji}, b_{ji}$  – коэффициенты  $j$ -й модели.

Задача идентификации заключается в определении вида  $g(x)$  и коэффициентов модели  $a_i, b_i (i = \overline{1, m})$  на основе наблюдений  $u(k), y(k)$  и сводится к минимизации некоторого функционала от ошибки идентификации  $e(k) = y(k) - \hat{y}(k)$ .

В классической модели Гаммерштейна [2] статическая нелинейность  $g_j(\cdot)$  представляется рядом:

$$g_j(u_j) = \sum_{p=0}^L \alpha_{jp} u_j^p, \quad (4)$$

где  $\alpha_{jp}$  – некоторые весовые коэффициенты;  $p$  – степень используемого многочлена, определяемая в процессе решения задачи идентификации.

Подставляя (4) в (3), имеем

$$\hat{y}_j(k) = \alpha_{j0} \sum_{i=1}^m b_{ji} + \alpha_{j1} \sum_{i=1}^m b_{ji} u_j(k-i) + \dots$$

$$\dots + \alpha_{ji} \sum_{i=1}^m b_{ji} u_j^L(k-i) - \sum_{i=1}^m a_{ji} \hat{f}_j(k-i). \quad (5)$$

Рассмотрим объект с двумя входными сигналами  $u_1(k)$  и  $u_2(k)$ . Пусть по каждому из каналов он представлен соответствующей нелинейной моделью Гаммерштейна:

$$y_1(k) = \alpha_{10} \sum_{i_1=1}^m b_{1i_1} + \alpha_{11} \sum_{i_1=1}^m b_{1i_1} u_1(k-i_1) + \dots + \alpha_{1L} \sum_{i_1=1}^m b_{1i_1} u_1^L(k-i_1) - \sum_{i_1=1}^m a_{1i_1} y_1(k-i_1); \quad (6)$$

$$y_2(k) = \alpha_{20} \sum_{i_2=1}^m b_{2i_2} + \alpha_{21} \sum_{i_2=1}^m b_{2i_2} u_2(k-i_2) + \dots + \alpha_{2L} \sum_{i_2=1}^m b_{2i_2} u_2^L(k-i_2) - \sum_{i_2=1}^m a_{2i_2} y_2(k-i_2).$$

Подставив соотношения (6) в (1), получим

$$y(k) = \alpha_{10} \alpha_{20} \sum_{i_1=1}^m \sum_{i_2=1}^m b_{1i_1} b_{2i_2} + \alpha_{10} \alpha_{20} \sum_{i_1=1}^m \sum_{i_2=1}^m b_{1i_1} b_{2i_2} u_2(k-i_2) + \dots + \alpha_{10} \alpha_{2L} \sum_{i_1=1}^m \sum_{i_2=1}^m b_{1i_1} b_{2i_2} u_2^L(k-i_2) + \alpha_{11} \alpha_{2L} \sum_{i_1=1}^m \sum_{i_2=1}^m b_{1i_1} b_{2i_2} u_1(k-i_1) u_2^L(k-i_2) + \alpha_{1L} \alpha_{2L} \sum_{i_1=1}^m \sum_{i_2=1}^m b_{1i_1} b_{2i_2} u_1^L(k-i_1) u_2^L(k-i_2) - \alpha_{1L} \sum_{i_1=1}^m \sum_{i_2=1}^m b_{1i_1} a_{2i_2} u_1^L(k-i_1) y_2(k-i_2) + \dots + \sum_{i_1=1}^m \sum_{i_2=1}^m a_{1i_1} a_{2i_2} y_1(k-i_1) y_2(k-i_2). \quad (7)$$

Задача построения математической модели вида (7) состоит в определении степени полинома  $L$  и вычислении соответствующих коэффициентов разложения  $\alpha_{ij}, b_{ijk}, a_{ijk}$ . Так как всякое представление в виде ряда представляет собой аппроксимацию реальной нелинейности, повысить точность аппроксимации можно путем увеличения степени используемого полинома. При этом, однако, возникают существенные вычислительные проблемы, обуславливаемые как значительным увеличением количества идентифицируемых параметров, так и ухудшением свойств используемых при этом матриц наблюдения.

Кроме того, существенный недостаток данного подхода состоит в том, что увеличение числа членов используемого полинома требует нового пересчета оценок всех параметров. Задача построения модели может быть существенно упрощена, если для аппроксимации нелинейной статической части модели

использовать ИНС персептронного типа или радиально-базисную сеть [3].

## 2. Нейросетевая модель Гаммерштейна

Аппроксимация нелинейной статической функции  $g_j(u_j(k))$  (2) гауссовскими радиально-базисными функциями

$$\varphi_p(x) = \exp\left\{-\frac{\|x - \mu_p\|^2}{\sigma_p^2}\right\}, \quad (8)$$

где  $\mu_p$  и  $\sigma_p$  – центры и радиусы базисных функций соответственно;  $\|\cdot\|$  – евклидова норма, дает:

$$u_j(k) = \sum_{p=1}^M w_{jp} \varphi_p(u_j(k)). \quad (9)$$

Тогда по каждому из сигналов исследуемый объект может быть представлен следующим образом:

$$y_n(k) = \sum_{i_n=1}^m \beta_{ni_n} u_n(k-i_n) + \dots + \sum_{i_n=1}^m \sum_{j_n=1}^M \beta_{ni_n j_n} \varphi_{j_n}(u_n(k-i_n)) - \sum_{i_n=1}^m a_{ni_n} y_n(k-i_n), \quad (10)$$

где  $\beta_{ni_n} = b_{ni_n}$ ;  $\beta_{ni_n j_n} = b_{ni_n} w_{nj_n}$ .

Подставляя (10) в (1), получаем соотношения, описывающие нейросетевую мультипликативную модель Гаммерштейна

$$\begin{aligned} \hat{f}(k) = & \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_n=1}^m \beta_{1i_1} \beta_{2i_2} \dots \beta_{ni_n} u_1(k-i_1) u_2(k-i_2) \dots u_n(k-i_n) + \\ & + \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_n=1}^m \sum_{j_n=1}^N \beta_{1i_1} \beta_{2i_2} \dots \beta_{ni_n j_n} u_1(k-i_1) u_2(k-i_2) \dots \varphi_{j_n}(u_n(k-i_n)) - \\ & - \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_n=1}^m \beta_{1i_1} \beta_{2i_2} \dots a_{ni_n} u_1(k-i_1) u_2(k-i_2) \dots y_n(k-i_n) + \\ & + \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_n=1}^m \sum_{j_n=1}^N \beta_{1i_1} \beta_{2i_2} \dots \beta_{ni_n} u_1(k-i_1) u_2(k-i_2) \dots \varphi_{j_n}(u_n(k-j_n)) + \dots \\ & \dots + (-1)^n \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{j_n=1}^N a_{1i_1} a_{2i_2} \dots a_{ni_n} y_1(k-i_1) y_2(k-i_2) \dots y_{j_n}(k-i_n). \quad (11) \end{aligned}$$

Уравнение (11) может быть представлено как

$$\hat{f}(k) = \theta^T x(k), \quad (12)$$

где  $\theta$  – вектор-столбец неизвестных параметров, представляющий из себя произведения коэффициентов  $\beta_{ij} \dots \beta_{ni_n j_n}, \dots \beta_{ij_n} a_{nj_n}, a_{nj_n} a_{mj_n}$  и т.д.;  $x(k)$  – вектор обобщенных входов, образованный соответствующими произведениями входных  $u_i$  и выходных  $y_j$  сигналов.

Таким образом, задаваясь структурой модели (10), т.е. выбирая  $m, n$  и  $N$ , можно получить необходимые уравнения. Например, при  $m = n = 2$  и  $N = 3$  имеем

$$\begin{aligned} \xi_n = & \beta_{n1}u_n(k-1) + \beta_{n2}u_n(k-2) + \\ & + \beta_{n11}\varphi_1(u_n(k-1)) + \beta_{n12}\varphi_2(u_n(k-1)) + \\ & + \beta_{n13}\varphi_3(u_n(k-1)) + \beta_{n21}\varphi_1(u_n(k-2)) + \\ & + \beta_{n22}\varphi_2(u_n(k-2)) + \beta_{n23}\varphi_3(u_n(k-2)) - \\ & - a_{n1}\xi_n(k-1) - a_{n2}\xi_n(k-2), \quad n = 1, 2. \end{aligned}$$

В этом случае вектор параметров  $\theta$  и вектор обобщенных входов  $x(k)$  будут иметь соответственно вид:

$$\begin{aligned} \theta = & (\beta_{11}\beta_{21} \beta_{11}\beta_{22} \dots \beta_{12}\beta_{22} \beta_{11}\beta_{211} \dots a_{12}a_{22})^T; \\ x(k) = & (u_1(k-1)u_2(k-1) \ u_1(k-1)u_2(k-2) \dots \\ & \dots u_1(k-2)u_2(k-1) \ u_1(k-1)\varphi_2(u_2(k-1)) \dots \\ & \dots u_1(k-2)u_2(k-2))^T. \end{aligned}$$

Задача идентификации нелинейного объекта (1) сводится таким образом к обучению ИНС, заключающемся в настройке ее параметров на основании сравнения выходного сигнала объекта  $y(k)$  и реакции сети  $\hat{y}(k)$ . Выбирая в качестве критерия квадратичный, можно получить алгоритм обучения на основе рекуррентного метода наименьших квадратов [4]:

$$\theta(k) = \theta(k-1) + \frac{P(k-1)\nabla F(k)}{\lambda + \nabla^T F(u)P(u-1)\nabla F(u)} e(k);$$

$$P(u) = \frac{1}{\lambda} \left[ P(u-1) - \frac{P(u-1)\nabla F(u)\nabla^T F(u)P(u-1)}{\lambda + \nabla^T F(k)P(u-1)\nabla F(u)} \right],$$

$$\begin{aligned} \nabla F(u) = & (1, \varphi_1(x(u)), 2\varphi_1(x(u))\beta_{11}(x(u) - \\ & - \mu_1)^T \sigma_1^{-2}, 2\varphi_1(x(u))\beta_{11}\|x(u) - \mu_1\|^2 \sigma_1^{-3}, \dots \\ & , \varphi_N(x(u)), 2\varphi_N(x(u))\beta_{mN}(x(u-m) - \\ & - \mu_N)^T \sigma_N^{-2}, 2\varphi_N(x(k-m))\beta_{mN}(x(k-m) - \\ & - \mu_N)^T \sigma_N^{-3}, y(k-1), \dots, y(k-m))^T, \end{aligned}$$

$\lambda \in (0, 1]$ ,  $P(0) = \gamma I$ ,  $\gamma \gg 0$ ,  $I$  – единичная матрица.

Следует отметить, что при использовании базисных функций (7) с настраиваемыми центрами и радиусами эти параметры также следует ввести в вектор  $\theta_j$ , что приводит к существенному увеличению количества обучаемых параметров и, как следствие, к затягиванию процесса обучения.

Нейросетевая реализация модели Гаммерштейна (10) приведена на рис. 2.

### Заключение

С целью получения адекватного математического описания нелинейного динамического объекта со многими входами (MISO) в работе решена задача построения нейросетевой мультипликативной модели Гаммерштейна.

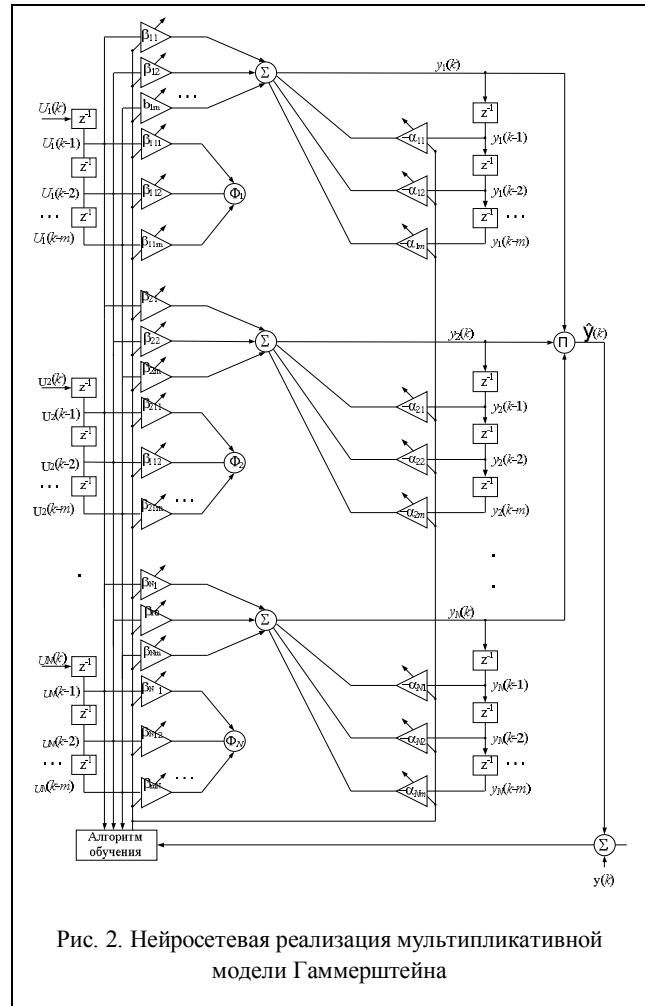


Рис. 2. Нейросетевая реализация мультипликативной модели Гаммерштейна

Предложенная специализированная структура радиально-базисной сети, реализующая модель Гаммерштейна, не отличается сложностью. Для обучения данной сети может быть применен рекуррентный метод наименьших квадратов.

### Список литературы

1. Nelles O. *Nonlinear System identification*. – Berlin: Springer, 2000. – 785 p.
2. Gallman P.G., Narendra K.S. *Representation of nonlinear systems via Stone-Weierstrass theorem* // *Automatica*. – 1976. – V. 12, N 8. – P. 619-622.
3. Руденко О.Г., Бодянский Е.В. *Искусственные нейронные сети*. – Х.: ООО «Компания СМІТ», 2005. – 408 с.
4. Сакало Т.С. *Синтез нелинейной модели Гаммерштейна на основе радиально-базисной сети* // *Міжнародна науково-практична конференція «Інтелектуальні системи прийняття рішень та прикладні аспекти інформаційних технологій» ISDMIT'2006*. Зб. наук. пр. у 4-х томах. – Євпаторія. – 2006. – Т. 2. – С.268-272.

Поступила в редколлегию 02.02.2007

Рецензент: д-р. физ.-мат. наук, проф. А.Г. Руткас, Харьковский национальный университет им. В.Н. Каразина, Харьков.