

用 B-样条神经网络设计非线性观测器

张洪钺 金宏

(北京航空航天大学自动控制系 北京 100083 E-mail: buaa301@mimi.cnc.ac.cn)

陈志炜 张祺忠

(香港大学机械工程系 香港 E-mail: mechan@hkucc.hku.hk)

摘要 对线性部分已知、非线性部分未知的一类非线性系统,提出了一种新的状态观测器的设计方法.首先针对线性部分设计线性观测器,随后在线性观测器中加入非线性补偿项.该补偿项先由“反卷积”的方式确定,再用 B-样条神经网络拟合.对三个非线性系统设计了观测器,通过与已有的解析方法进行比较,说明了该方法的优越性.

关键词 非线性系统,观测器,神经网络.

NONLINEAR OBSERVER DESIGN FOR SYSTEMS WITH UNKNOWN NONLINEARITY VIA B-SPLINE NETWORK APPROACH

ZHANG Hongyue JIN Hong

(Department of Automatic Control, Beijing University of Aeronautics and
Astronautics, Beijing 100083, E-mail: buaa301@mimi.cnc.ac.cn)

C. W. Chan K. C. Cheung

(Department of Mechanical Engineering, University of Hong Kong, Hong Kong, E-mail: mechan@hkucc.hku.hk)

Abstract A novel approach is proposed for the state estimation of a class of nonlinear systems which consist of known linear part and unknown nonlinear part. A linear observer is first designed, then a nonlinear compensation term in the nonlinear observer is estimated by using a “deconvolution method”. The B-spline neural network is used to model the obtained compensation term. Three simulation examples are given to show the effectiveness of the proposed approach

Key words Nonlinear system, observer, neural network.

1 引言

估计非线性系统的状态对系统控制、监控和故障诊断都是非常重要的.用解析方法设计非线性系统的观测器已有一些报道. Bestle, Zeitz 提出了典范观测器设计方法. Kerner 和 Respondek 提出了用李-代数方法将非线性系统转换为典范型的观测器设计方法. 然而,找到这样的一对一的转换不是易事,此外,控制对象的非线性模型必须已知. Bauman 和 Rugh 提出了用扩展线性化方法来设计观测器,但这种方法同样需要精确知道系统的

非线性特性,还要知道它的一阶导数. Walcott 和 Zak 提出了变结构观测器的设计. 这种方法只需知道非线性特性的界,然而在实践中这个界也许不能精确知道. 在文献[1]中,对上面几种方法进行了比较研究.

非线性观测器的设计问题尚有很多问题没有解决. 特别是当非线性特性未知,或非线性特性不满足上面解析设计方法的要求时,已有设计方法将不再适用.

神经网络可有效地逼近非线性函数,它已被成功的用于非线性系统的建模^[2]. 神经网络应用于状态估计问题的报道相对于建模问题要少得多^[3,4]. 尽管建模和状态估计都是函数逼近,但状态估计要求非线性系统满足可观性,因此更为困难.

本文针对具有任意未知非线性部分和已知线性部分的非线性系统,采用 B-样条神经网络设计了非线性观测器. 这个非线性观测器由一个线性观测器和一个非线性补偿项组成. 非线性补偿项先由本文提出的“反卷积”方法确定为时间函数,再用 B-样条神经网络对其进行建模. 仿真例子说明了本文所提出方法的有效性.

2 基于神经网络的非线性观测器的新的设计方法

考虑一类非线性系统

$$x(k+1) = Ax(k) + f(x(k)), \quad y(k) = Cx(k), \quad (1), (2)$$

其中 A 是系统的线性部分,设为已知; $f(x)$ 是未知的非线性函数,且不一定是光滑函数;测量方程是线性的, (A, C) 是可观的.

针对上述非线性系统提出如下非线性观测器的设计步骤:

1)根据式(1),(2)线性部分先设计一个线性观测器

$$\hat{x}_l(k+1) = A_0\hat{x}_l(k) + Ky(k), \quad A_0 = A - KC, \quad (3), (4)$$

式中 K 是观测器增益,可用已有的方法来设计; \hat{x}_l 是线性观测器的状态.

非线性观测器具有以下型式:

$$\hat{x}(k+1) = A_0\hat{x}(k) + Ky(k) + S(k), \quad \hat{y}(k) = C\hat{x}(k), \quad (5), (6)$$

式中 $S(k)$ 是针对式(1)中非线性项 $f(x)$ 而引入的补偿项,是 k 的函数,它与观测器特定的初始条件有关.

2)用“反卷积”的方法来确定非线性补偿项 $S(k)$. 因为非线性观测器的输出 \hat{y} 与 $S(k)$ 有一个卷积关系,将 \hat{y} 强迫等于系统的输出 $y(k)$,就可确定 $S(k)$.

3)用 B-样条神经网络对此非线性补偿项 $S(k)$ 进行建模,网络的输出为 $S(k)$,输入量的选择可以根据非线性特性 $f(x)$ 的已知输入量来定或通过实验来确定. 这样就求得了从 $\hat{x}(k)$ 到 $S(k)$ 的映射,即 $\hat{f}(\hat{x}(k)) = S(k)$.

3 用“反卷积”法确定非线性补偿项

将 $\hat{y}(k)$ 强制等于 $y(k)$ 就可求出非线性补偿项. 由式(5),(6)可得

$$\begin{aligned} \hat{y}(k) = C\hat{x}(k) &= C[A_0\hat{x}(k-1) + Ky(k-1) + S(k-1)] = \dots = \\ &CA_0^k\hat{x}(0) + \sum_{i=1}^k CA_0^{i-1}Ky(k-i) + \sum_{i=1}^k CA_0^{i-1}S(k-i), \end{aligned} \quad (7)$$

$$\tilde{y}(k) = y(k) - \hat{y}(k) = y^*(k) - \sum_{i=1}^k CA_0^{i-1} S(k-i), \quad (8)$$

$$y^*(k) = y(k) - CA_0^k \hat{x}(0) - \sum_{i=1}^k CA_0^{i-1} K y(k-i). \quad (9)$$

$\hat{x}(0)$ 可以任意给定,但需满足 $C\hat{x}(0) = \hat{y}(0) = y(0)$,因而 $y^*(k)$ 是已知的.强制使 $\tilde{y}(k) = 0$,并令 $C_i = CA_0^{i-1}$, $k=1, 2, \dots, M+1$,由(8)式可得

$$\begin{cases} C_1 S(0) = y^*(1), \\ C_1 S(1) + C_2 S(0) = y^*(2), \\ \vdots \\ C_1 S(M) + \dots + C_{M+1} S(0) = y^*(M+1). \end{cases} \quad (10)$$

当 $\text{rank}(C) = n$,且 $p \geq n$, p 为输出维数, n 为状态维数,可求出 $S(k)$ 如下

$$\begin{cases} S(0) = C^+ y^*(1), \\ S(k) = C^+ \left[y^*(k+1) - \sum_{i=0}^{k-1} C_{k+1-i} S(i) \right], \end{cases} \quad (11)$$

式中 $C^+ = (C'C)^{-1}C'$, $k=1, \dots, M$, C' 为 C 的转置.

当非线性特性只出现在一部分状态方程中时, $S(k)$ 可表示为 $S(k) = [0 \quad \bar{S}'(k)]'$,式中 $\bar{S}(k) \in R^{d \times n}$.将 C 相应地分割成 $C = [\tilde{C} \quad \bar{C}]$,式中 $\bar{C} \in R^{p \times d}$.

若 $\text{rank}(\bar{C}) = d$, $p \geq d$,可求出 $\bar{S}(k)$ 如下:

$$\begin{cases} \bar{S}(0) = \bar{C}^+ y^*(1), \\ \bar{S}(k) = \bar{C}^+ \left[y^*(k+1) - \sum_{i=0}^{k-1} C_{k+1-i} S(i) \right], \end{cases} \quad (12)$$

式中 $\bar{C}^+ = (\bar{C}'\bar{C})^{-1}\bar{C}'$, $k=1, \dots, M$.

4 用神经网络对非线性补偿项建模

本文采用 B-样条神经网络,其输出由下式给出^[5]

$$y(k) = \sum_{i=1}^p a_i(k) w_i, \quad (13)$$

式中 w_i 是相应于基函数 $a_i(k)$ 的权, $a_i(k)$ 的输入则为 $\hat{x}(k)$, p 是基函数的数目.对 B-样条网络来讲,它的指标函数(12)式是未知权系数的二次函数,因此只有一个极小值.此外 B-样条网络是局部网络,当输入在一定范围内变化时,只有对应范围内的基函数不为零,因此在训练时只有这些基函数的权值需要调整.这种局部调整很适合于在线训练网络.如果系统的非线性特性有多个输入,则需采用多变量基函数来作近似.多变量基函数是单变量基函数的张量(tensor)积.这时 B-样条网络的输出与这些多变量基函数成线性关系.

所求出的非线性补偿项作为网络的期望输出,而网络的输入即是非线性特性的输入.设网络输入变量为 $\hat{x}(k)$,而非线性补偿项 $S(k)$ 是标量函数(当 $S(k)$ 是向量函数时可以对各个分量来做),则网络的建模可以进行如下

$$\begin{cases} S(1) = \sum_{i=1}^p a_i(\hat{x}(1)) w_i + \epsilon(1), \\ \vdots \\ S(N) = \sum_{i=1}^p a_i(\hat{x}(N)) w_i + \epsilon(N), \end{cases} \quad (14)$$

$\varepsilon(k)$ 是网络的建模误差, $k=1, \dots, N$. 式(14)可以写成向量-矩阵形式

$$S = \bar{A}W + E, \quad (15)$$

$$\text{式中 } \bar{A} = \begin{bmatrix} a_1(\hat{x}(1)) & \cdots & a_p(\hat{x}(1)) \\ \cdots & \cdots & \cdots \\ a_1(\hat{x}(N)) & \cdots & a_p(\hat{x}(N)) \end{bmatrix}, E' = [\varepsilon(1) \quad \cdots \quad \varepsilon(N)], S' = [S(1) \quad \cdots \quad S(N)], W' = [w_1 \quad \cdots \quad w_p].$$

通常方程的数目大于未知数的数目, 即 $N > p$, 再假定 \bar{A} 列满秩, 因此可用最小二乘法来决定权系数如下:

$$W = (\bar{A}'\bar{A})^{-1}\bar{A}'S. \quad (16)$$

当 $S(k)$ 已求得, 且 $X(0)$ 给定时, $X(k)$ 可由式(5)递推求出, 故神经网络的输入输出样本可获得.

5 估计误差的分析

神经网络估计器的输出可写成

$$\hat{x}(k+1) = A_0\hat{x}(k) + Ky(k) + \hat{f}(\hat{x}(k)) \quad (17)$$

式中 $\hat{f}(\hat{x}(k))$ 是网络输出. 由式(1), (17)可得

$$e(k+1) = \hat{x}(k+1) - x(k+1) = A_0e(k) + [\hat{f}(\hat{x}) - f(x)] = A_0e(k) + \Delta\hat{f}(k) + \varepsilon(k), \quad (18)$$

式中 $\Delta\hat{f}(k) = \hat{f}[x(k)+e(k)] - \hat{f}[x(k)]$, $\varepsilon(k) = \hat{f}[x(k)] - f[x(k)]$.

因为 A_0 是 Hurwitz 矩阵, 故给定任意的正定实对称矩阵 Q , 存在着一个正定实对称矩阵 P , 使满足

$$A_0'PA_0 - P = -Q. \quad (19)$$

采用李亚普诺夫函数 $V(k) = e'(k)Pe(k)$, 并暂时先设网络的逼近无误差, 即 $\varepsilon(k) \equiv 0$. 由式(18)可得

$$V(k+1) = e'(k+1)Pe(k+1) = e'(k)A_0'PA_0e(k) + 2\Delta\hat{f}'(k)PA_0e(k) + \Delta\hat{f}'(k)P\Delta\hat{f}(k), \quad (20)$$

$$\Delta V(k) = V(k+1) - V(k) = -e'(k)Qe(k) + 2\Delta\hat{f}'(k)PA_0e(k) + \Delta\hat{f}'(k)P\Delta\hat{f}(k). \quad (21)$$

因 B-样条的输出 \hat{f} 满足李普希茨条件

$$\|\Delta\hat{f}(k)\| = \|\hat{f}(x(k)+e(k)) - \hat{f}(x(k))\| \leq L\|e(k)\|, \quad (22)$$

其中 L 是一个整数. 于是可得

$$\Delta V(k) \leq (-\lambda_{q\min} + 2L\lambda_{p\max}\sigma_{\max} + L^2\lambda_{p\max})\|e(k)\|^2, \quad (23)$$

式中 $\lambda_{q\min}$ 是 Q 的最小特征值, $\lambda_{p\max}$ 是 P 的最大特征值, σ_{\max} 是 A_0 的最大奇异值. 如果有

$$-\lambda_{q\min} + 2L\lambda_{p\max}\sigma_{\max} + L^2\lambda_{p\max} < 0, \quad (24)$$

则 $\Delta V(k) < 0$. 这意味着状态估计误差 $e(k)$ 渐进趋于零.

实际情况中, 网络的逼近总是有误差的, 即对所有 k , $\|\varepsilon(k)\|$ 不恒等于零. 这时可以给

出 e 的一个上界. 由式(18)可得

$$\begin{aligned} \|e(k+1)\| &\leq \|A_0 e(k)\| + \|\Delta \hat{f}(k)\| + \|\varepsilon(k)\| \\ &\leq (\sigma_{\max} + L)\|e(k)\| + \|\varepsilon(k)\|. \end{aligned} \quad (25)$$

当 $(\sigma_{\max} + L) < 1$, 式(25)是稳定的, 因而对足够大的 k , 有

$$\|e(k+1)\| \approx \|e(k)\|. \quad (26)$$

将式(26)代入式(25), 可得

$$\|e(k)\| \leq \frac{\|\varepsilon(k)\|}{1 - (\sigma_{\max} + L)}, \quad (\text{对大的 } k \text{ 值}). \quad (27)$$

由式(27)可见 $\|e(k)\|$ 正比于 $\|\varepsilon(k)\|$ 、 σ_{\max} 、 L . 这意味着网络逼近误差愈小, σ_{\max} 和 L 愈小, 则状态估计误差愈小.

6 仿真例子

本节先用解析的方法设计几个非线性系统的观测器, 再用本文的方法来设计, 并比较两种方法的结果, 然后对解析方法不适用的某些非线性系统用本文的方法来设计观测器.

例1^[1]. 系统和测量方程为

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\sin x_1 \end{bmatrix}, \quad y = x_1 + x_2. \quad (28), (29)$$

文献[1]中用四种解析方法来设计观测器, 并对结果进行了比较.

a) 李-代数(Lie-algebraic)法所设计的观测器具有下列型式

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} \hat{x}_2 \\ -\sin \hat{x}_1 \end{bmatrix} + \begin{bmatrix} g_1(\hat{x}) \\ g_2(\hat{x}) \end{bmatrix} (y - \hat{x}_1 - \hat{x}_2), \quad (30)$$

式中 $g_1(\hat{x})$, $g_2(\hat{x})$ 具有很复杂的表达式^[1].

b) 扩展线性化方法和 Tau 的方法, 这两种方法所设计的观测器具有相同的型式

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\sin \hat{x}_1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} y. \quad (31)$$

c) 变结构观测器方法所设计的观测器为

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \hat{S}(\hat{x}_1, \hat{x}_2, y) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} y. \quad (32)$$

式中 $\hat{S}(\hat{x}_1, \hat{x}_2, y) = -2\text{sgn}(\hat{x}_1 + \hat{x}_2 - y)$. 对相同的初始条件, 上面几个观测器的仿真输出基本相似. 文献[1]认为变结构观测器的性能最好. 图1表示变结构观测器的相应曲线, 由图1可见, \hat{x}_2 有明显的抖动, 这一点在文献[1]中未表示出来.

d) 用本文的方法先设计线性观测器, 它具有与 b), c) 方法相同的型式. 将线性观测器离散化, 采样频率为 0.05, 再加上一个非线性补偿项即可构成下面的离散非线性观测器

$$\begin{bmatrix} \hat{x}_1(k+1) \\ \hat{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.9512 & 0 \\ -0.0476 & 0.9512 \end{bmatrix} \begin{bmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \end{bmatrix} + \begin{bmatrix} 0.0488 \\ 0.0476 \end{bmatrix} y(k) + \begin{bmatrix} 0 \\ S(k) \end{bmatrix}, \quad (33)$$

式中 $S(k)$ 可用本文的“反卷积”方法给出, 再用 B-样条网络建模. 基于对系统非线性特性的知识, 网络的输入取为 \hat{x}_1 , 即 $S(k)$ 用网络输出 $\hat{S}(\hat{x}_1(k))$ 来近似. 当不具备这个知识时,

可以试探 \hat{x}_1, \hat{x}_2 或它们的适当组合作为输入, 使网络的输出能最好地逼近 $S(k)$.

在仿真中, 用9个单变量线性 B-样条函数, 节点均匀分布, 网络的输入范围为 $[-2, 2]$. 仿真的响应曲线示于图2中. 图中只比较了 $x_2(k), \hat{x}_2(k), x_1(k), S(k), y(k)$ 的估计效果也很好, 限于篇幅没有列出. 若进一步精化网络的结构和参数, 有可能使估计精度更高.

例2. 线性部分与例1一样, 非线性部分则为死区函数

$$f(x) = \begin{cases} K(x - b_2), & x \geq b_2, \\ 0, & b_1 < x < b_2, \\ K(x - b_1), & x \leq b_1 \end{cases} \quad (34)$$

在仿真中 $K=1, b_1=-0.8, b_2=0.8$.

李-代数方法和扩展线性化方法都不能用于设计这个非线性系统的观测器. 尽管变结构观测器原则上适用于这种非线性系统, 但它需要知道非线性函数的上界. 在例1中, 非线性函数为 $\sin(x_1)$, 因为 $\|\sin x_1\| < 1$ 故上界为1. 对上面的死区函数来说, 非线性函数 $f(x)$ 的上界取决于 x 的变化范围. 由于这个范围在设计前是不知道的, 因此在变结构观测器的设计中我们试用不同的上界值. 图3表示上界取得太小, $\hat{S}(\hat{x}_1, \hat{x}_2, y) = -0.1 \operatorname{sgn}(\hat{x}_1 + \hat{x}_2 - y)$, 这时 \hat{x}_2 跟踪 x_2 非常迟缓, 估计误差 $(\hat{x}_2(k) - x_2(k))$ 很大. 图4表示上界取得太大, $\hat{S}(\hat{x}_1, \hat{x}_2, y) = -8 \operatorname{sgn}(\hat{x}_1 + \hat{x}_2 - y)$, 这时尽管 \hat{x}_2 能很快跟踪 x_2 , 但抖动幅度很大, 因此估计误差仍较大. 图5表示本文方法估计 \hat{x}_2 的结果, 估计精度很高.

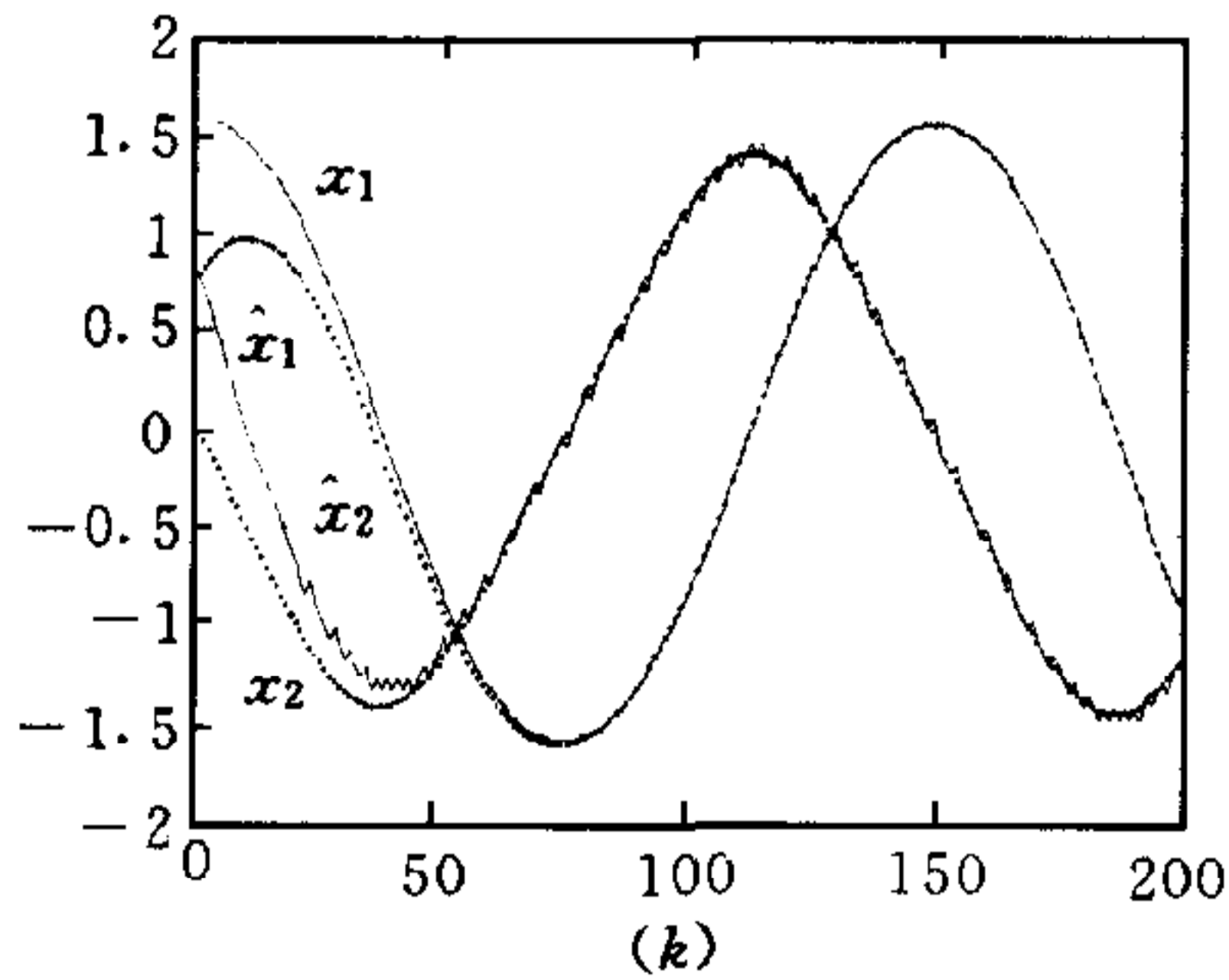


图1 $f(x) = \sin x$ 时变结构观测器的估计

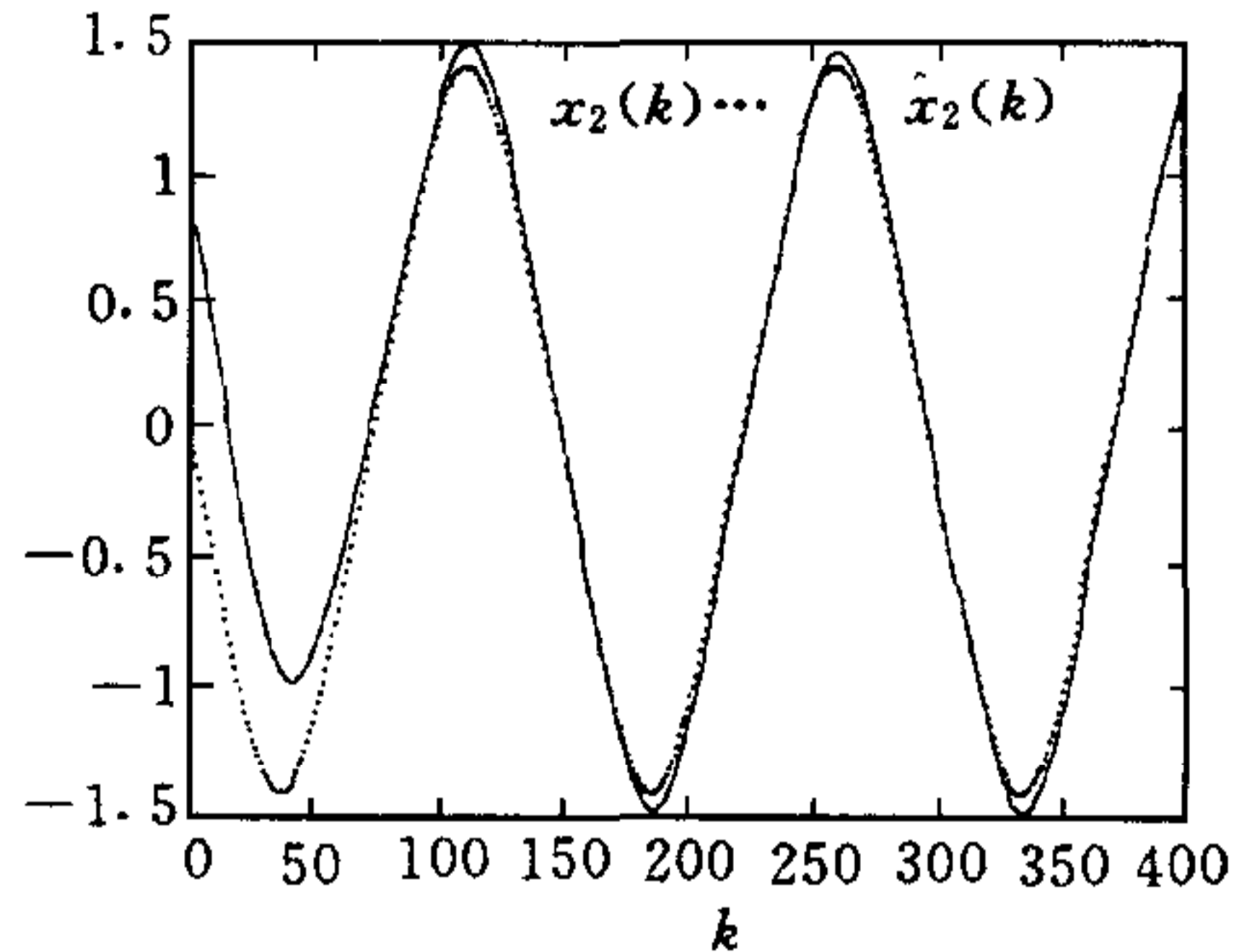


图2 $f(x) = \sin x$ 时, 本文方法结果

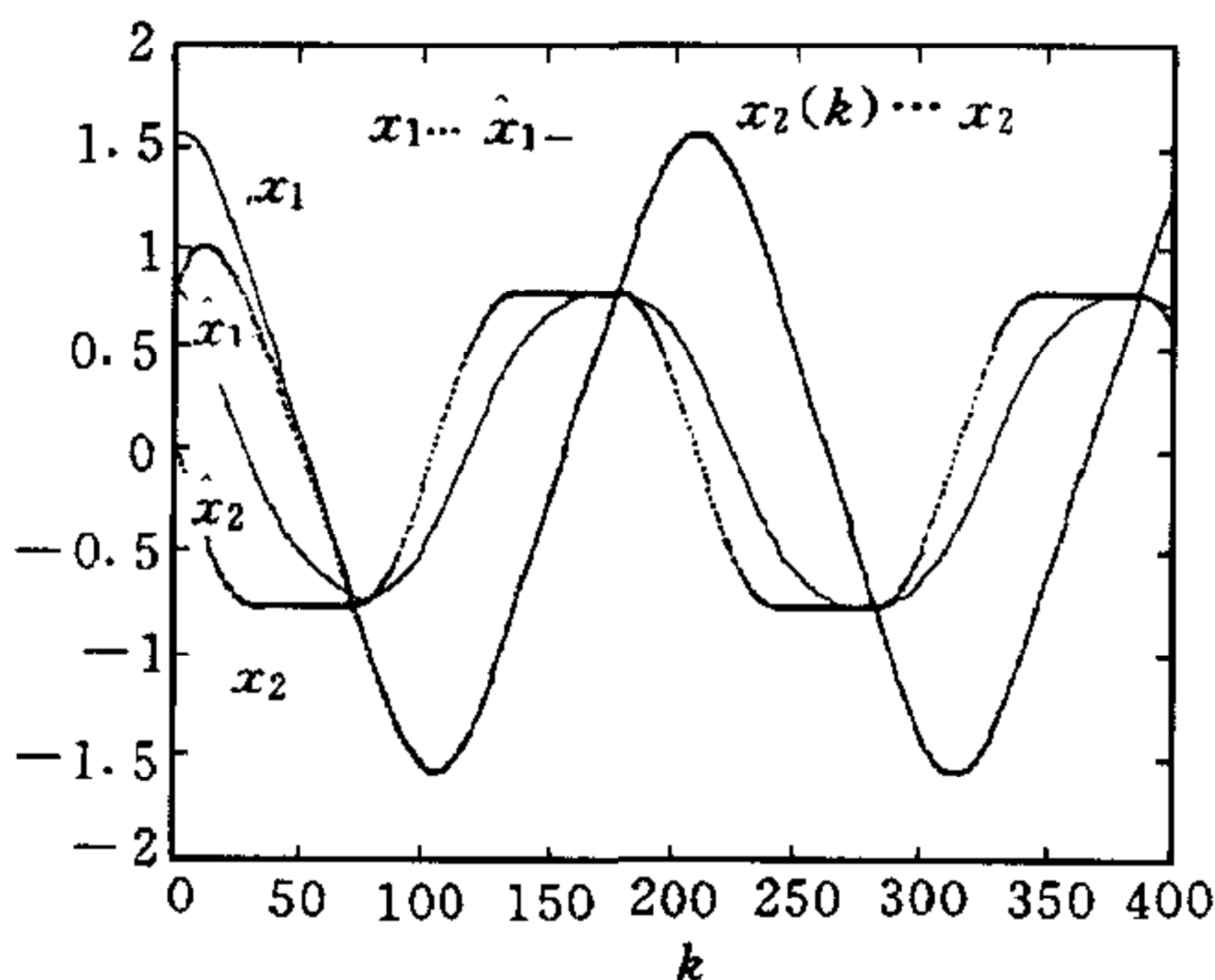


图3 $f(x)$ 为死区函数, \hat{S} 太小时变结构观测器的结果

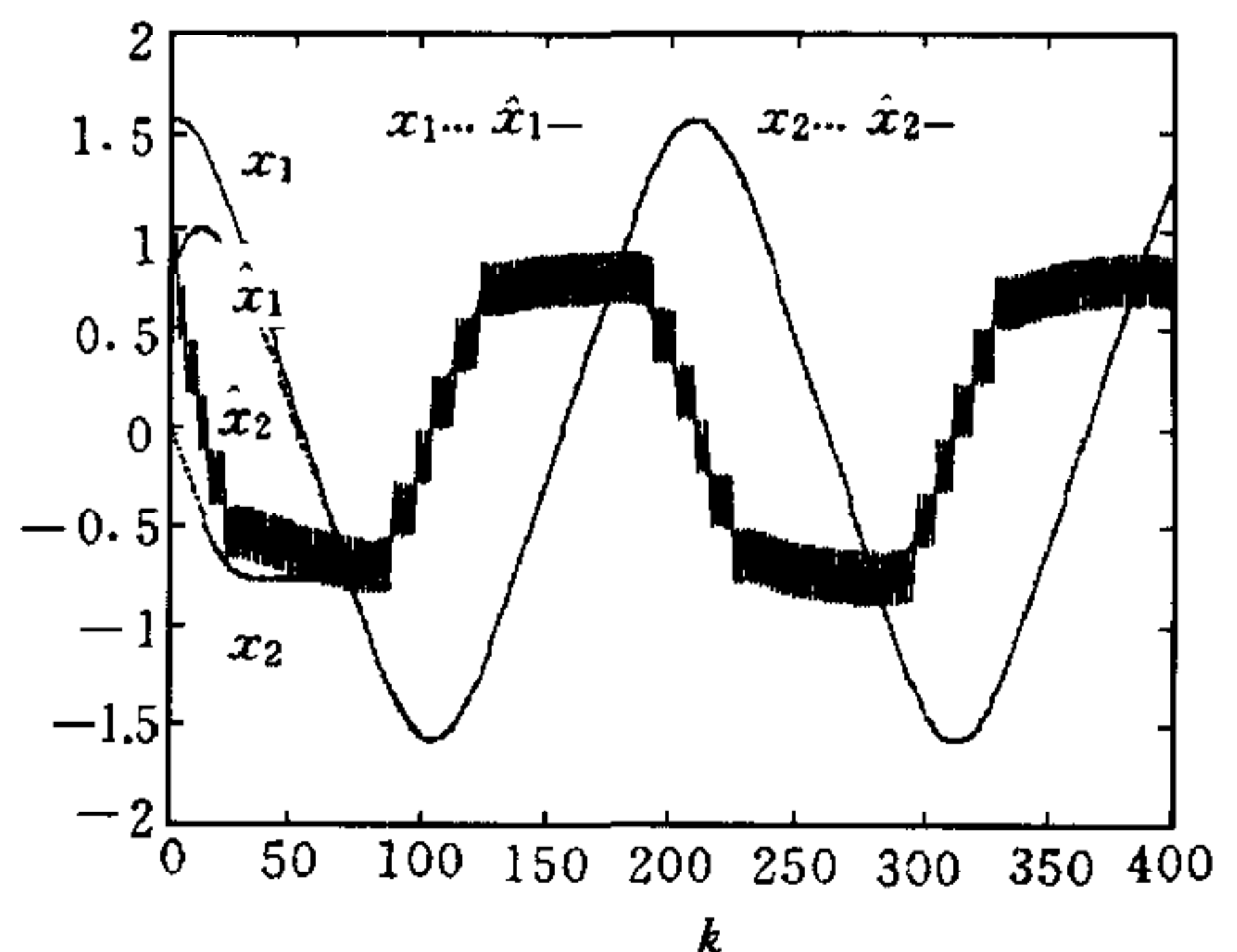


图4 $f(x)$ 为死区函数, \hat{S} 太大时变结构观测器的结果

例3. 线性部分同例1, 非线性部分是库伦摩擦函数

$$f(x) = \begin{cases} b + Kx, & x > 0, \\ -b + Kx, & x < 0. \end{cases} \quad (37)$$

这时李-代数法和扩展线性化法也不能用. 图6表示本文方法结果, 估计精度很高.

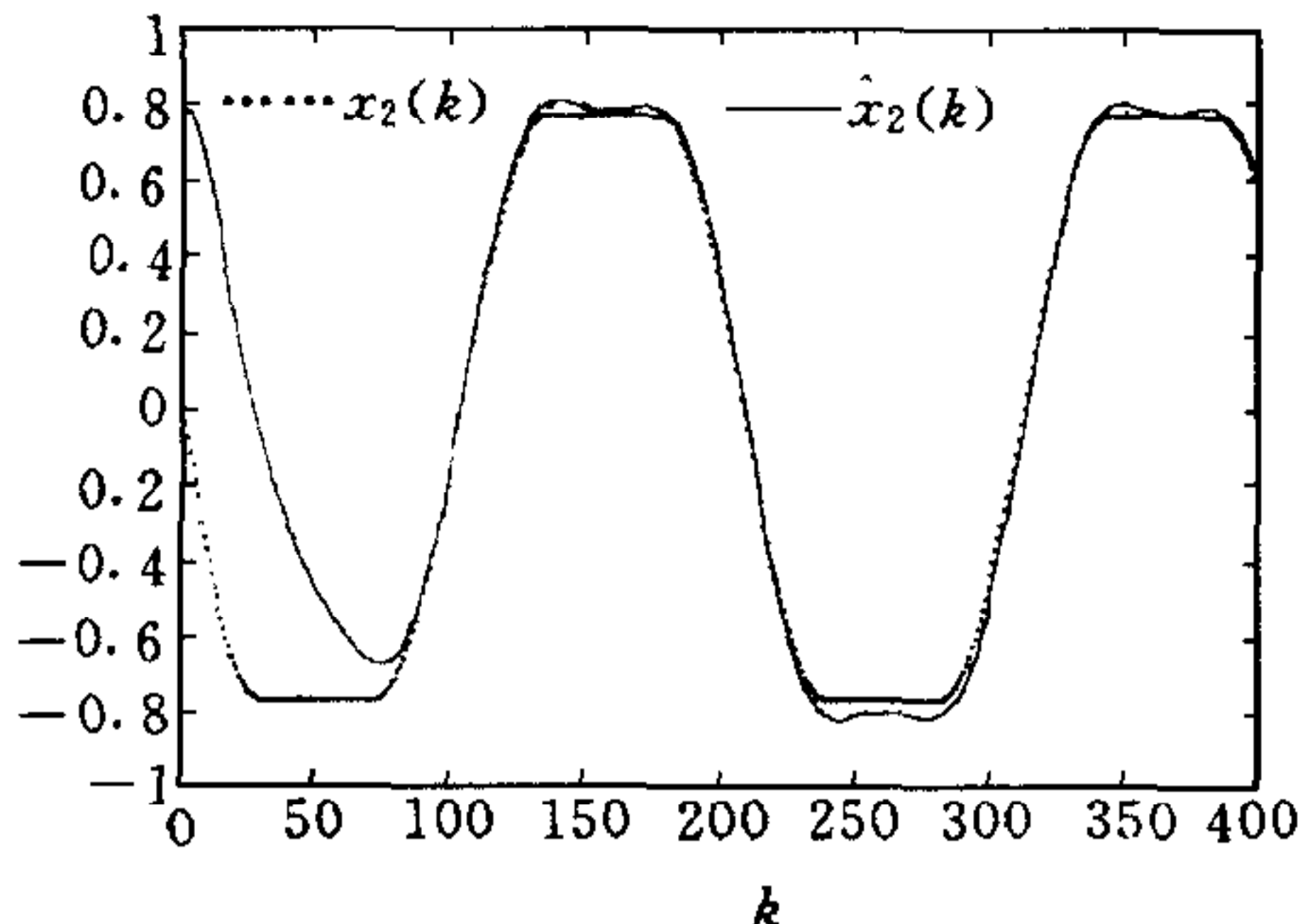


图5 $f(x)$ 为死区函数时本文方法结果

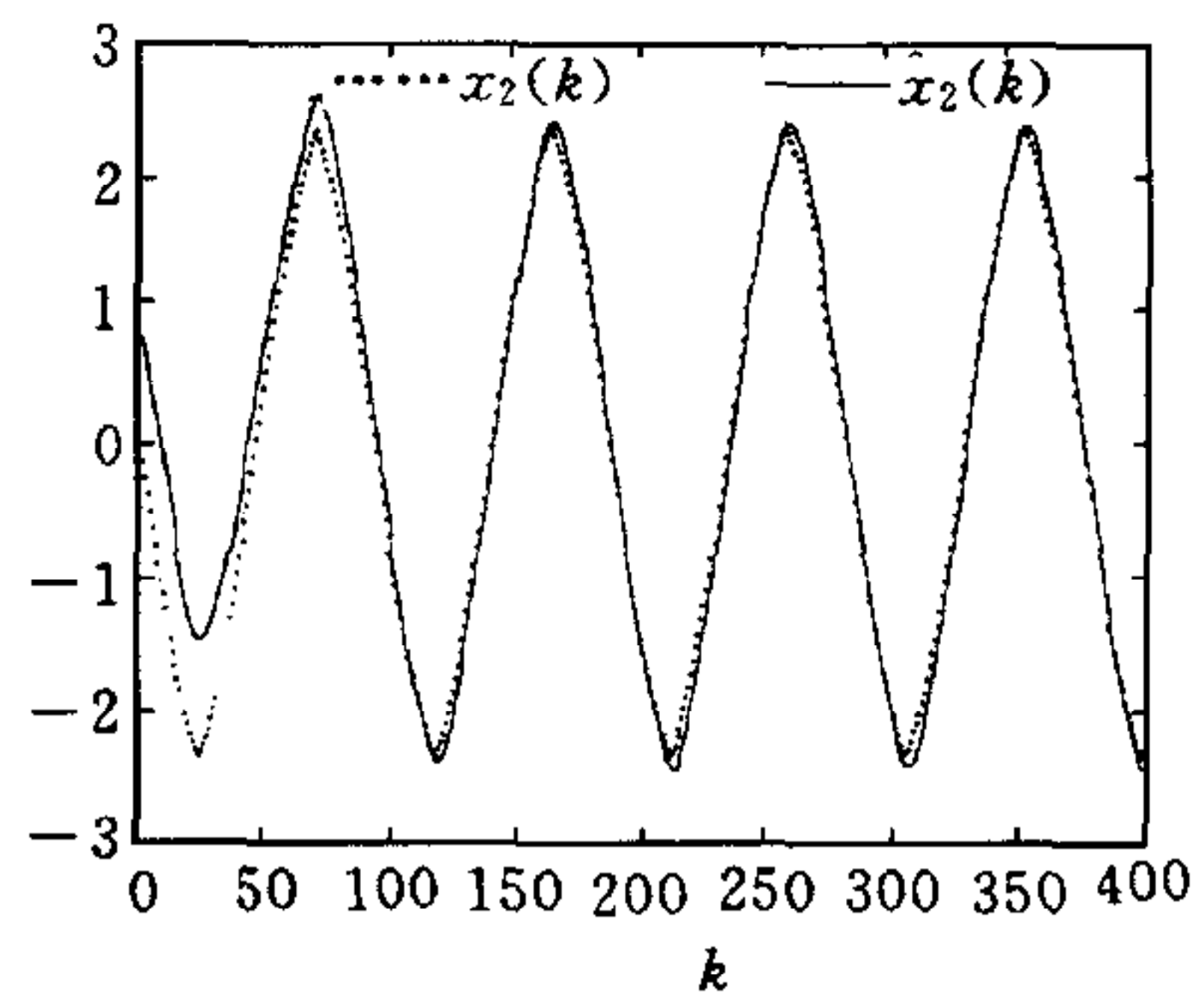


图6 $f(x)$ 为库伦摩擦函数时本文方法结果

7 结论

非线性观测器的设计并非易事, 已有的解析方法可以给出系统化的设计步骤, 但它们的应用是有局限性的. 本文所得出的方法可以应用于方程(1), (2)所描述的一类非线性系统, 其中线性部分已知并满足可观性条件, 非线性部分则可以未知且对其特性不加限制. 它的设计方法包括三个步骤: 1) 针对系统的线性部分设计一个线性观测器. 非线性观测器是由此线性观测器再加上一个非线性补偿项所构成, 如(5)式所示; 2) 用“反卷积”方法确定非线性补偿项 $S(k)$, 如式(11), (12)所示; 3) 用 B-样条神经网络对 $S(k)$ 建模, 网络系数由式(16)确定. 这样 $S(k)$ 被换成 $f(\hat{x}(k))$, 它是 $\hat{x}(k)$ 的函数. 仿真例子证明了本文方法的有效性, 在解析方法不能应用时, 本文方法仍有效.

参 考 文 献

- 1 Walcott B L, Corless M J, Zak S H. Comparative study of non-linear state-observation techniques. *Int. J. Control*, 1987, **45**:2109~2132
- 2 Narendra K S, Pathasarathy K. Identification and control of dynamical systems using neural network. *IEEE Trans. Neural Network*, 1990, **1**(1):4~27
- 3 Levin A U, Narendra K S. Control of nonlinear dynamical systems using neural network-part 2: Observability, Identification and Control. *IEEE Trans. on Neural Networks*, 1996, **7**(1)
- 4 LI Y T, Qin Z X, Zhang H Y(张洪钺). Nonlinear observer based on the neural network. In: Proceedings of the First Chinese World Congress on Intelligent Control and Intelligent Automation. Beijing:1993
- 5 Brown M, Harris C. *Neurofuzzy Adaptive Modelling and Control*. New York:Prentice Hall, 1994

张洪钺 教授、博士生导师. 主要研究领域有容错控制与导航、故障诊断、神经网络应用.

金 宏 博士 主要研究领域有导航系统的可靠性与精度分析、故障诊断、神经网络应用.

陈志炜 香港大学高级讲师.

张祺忠 香港大学副教授.