ISSN 0868–5886 *НАУЧНОЕ ПРИБОРОСТРОЕНИЕ, 2005, том 15, № 4, с. 98–101*

ОРИГИНАЛЬНЫЕ СТАТЬИ

УДК 519.816+519.23

© С. В. Соколов, И. В. Щербань, В. А. Бертенев

ИСПОЛЬЗОВАНИЕ НЕЛИНЕЙНЫХ ВЕРОЯТНОСТНЫХ КРИТЕРИЕВ ДЛЯ РЕШЕНИЯ ЗАДАЧ АДАПТИВНОЙ ФИЛЬТРАЦИИ

В статье рассмотрен общий метод решения задачи адаптивной параметрической фильтрации, обеспечивающий по сравнению с традиционными методами существенно меньший объем вычислительных затрат и потенциально большую точность процедуры оценивания-идентификации. Повышение точности достигается за счет использования вместо среднеквадратичного критерия более общих вероятностных критериев. Приведен пример идентификации стохастического нелинейного объекта.

ВВЕДЕНИЕ

Разработанные на сегодняшний день методы решения задачи идентификации-оценивания параметрически неопределенного динамического объекта, функционирующего в условиях возмущающих воздействий, предполагают, как известно, расширение его вектора состояния за счет вектора неизвестных параметров с последующим оцениванием всего расширенного вектора [1–4]. При подобном подходе N неизвестных параметров увеличивают размерность интегрируемой системы уравнений оценок с учетом симметрии матрицы ковариаций на величину $\dfrac{N^2}{2}+\left(\dfrac{3}{2}+M\right)N$, где M — размерность вектора состояния системы, что существенно влияет на число параметров, допускающих практическую возможность идентификации.

Более того, при этом, как правило, принимается весьма упрощающее допущение о постоянстве идентифицируемых параметров на интервале наблюдения, что для подавляющего большинства реальных ситуаций оказывается условием невыполнимым, а в общем случае существенно снижает потенциальную точность оценивания.

Таким образом, поиск новых путей решения задачи параметрической идентификации, свободных от вышеупомянутых ограничений, представляет собой очевидный научный и практический интерес. Ниже предлагается подход, позволяющий не только избежать перечисленных недостатков традиционного метода, но и повысить потенциальную точность идентификации за счет использования вместо среднеквадратичного критерия более общих вероятностных критериев, зависящих от плотности распределения и обеспечивающих достижение потенциально большей точности. Для дальнейшего решения поставленной задачи сформулируем ее следующим образом.

ПОСТАНОВКА ЗАДАЧИ

Пусть нелинейный стохастический динамический объект, наблюдаемый зашумленным нелинейным измерителем

$$\mathbf{Z}=\mathbf{H}(\mathbf{X},t)+\mathbf{W}, \qquad (1)$$

где $\mathbf{X}$ — наблюдаемый вектор состояния, $\mathbf{H}(\mathbf{X},t)$ — известная нелинейная вектор-функция наблюдения, $\mathbf{W}$ — центрированный белый гауссовский шум с матрицей интенсивности D_W, описывается векторным дифференциальным уравнением

$$\dot{\mathbf{X}}=\mathbf{U}_0(\mathbf{X},t)+\mathbf{U}(\mathbf{X},t)+\mathbf{V},$$

где $\mathbf{U}_0(\mathbf{X},t)$ — известная нелинейная вектор-функция, $\mathbf{U}(\mathbf{X},t)$ — вектор-функция с параметрической неопределенностью, $\mathbf{V}$ — центрированный белый гауссовский шум с матрицей интенсивности D_V.

В общем случае вектор $\mathbf{U}(\mathbf{X},t)$ можно представить в виде

$$\mathbf{U}(\mathbf{X},t)=S(\mathbf{X})\mathbf{a}(t),$$

где $S(\mathbf{X})$ — известная нелинейная функция-матрица, $\mathbf{a}(t)$ — искомый вектор неизвестных параметров, и записать уравнение объекта как

$$\dot{\mathbf{X}}=\mathbf{U}_0(\mathbf{X},t)+S(\mathbf{X})\mathbf{a}+\mathbf{V}. \qquad (2)$$

Для решения задачи адаптивной фильтрации (идентификации-оценивания) в соответствии с вышеизложенным необходимо, чтобы искомый вектор $\mathbf{a}(t)$ доставлял оптимум некоторому заданному обобщенному вероятностному функционалу J, зависящему от апостериорной плотности вероятности (АПВ) $\rho(\mathbf{X},t)$ процесса $\mathbf{X}$, причем в общем случае нелинейно.

Анализ физического существа решаемой задачи показывает, что в качестве наиболее адекватной формы минимизируемого функционала J целесообразно использовать аддитивную совокупность двух функционалов, оптимизация первого из которых $J_1 = \int_X \Phi_1\left[\rho(\mathbf{X},t)\right]\mathrm{d}\mathbf{X}$ должна обеспечить минимум неопределенности (максимальную информативность) идентифицируемого вектора $\mathbf{a}$, а второго $J_2 = \int_T \Phi_2\left[\mathbf{a}(\mathbf{X},t)\right]\mathrm{d}\mathbf{X}$ — минимум его "энергетики" (в соответствии с принципом Ферма) на заданном (конечном или текущем) интервале времени T , т. е.

$$J = \int_X \Phi_1\left[\rho(\mathbf{X},t)\right]\mathrm{d}\mathbf{X} + \int_T \Phi_2\left[\mathbf{a}(\mathbf{X},t)\right]\mathrm{d}\mathbf{X}.$$

При этом в соответствии с постановкой задачи функцию Φ_1 можно выбирать как ядро функционала энтропии Шеннона ($\Phi_1 = -\rho\ln\rho$) или Фишера ($\Phi_1 = -\rho\left[\frac{\partial\ln\rho}{\partial\mathbf{X}}\right]\left[\frac{\partial\ln\rho}{\partial\mathbf{X}}\right]^{\mathrm{T}}$) (т. н. "информационных функционалов"), а Φ_2 — в виде классической квадратичной "энергетической" формы, заданной на текущем интервале времени

$$J_2 = \int_{t_0}^{t} \mathbf{a}^{\mathrm{T}}\mathbf{a}\,\mathrm{d}t = \int_{t_0}^{t} \Phi_2(\mathbf{a})\,\mathrm{d}t.$$

Очевидно, что приведенные формы функционалов J_1, J_2 не являются единственными. Так, если *apriori* известна форма $g(\mathbf{X},t)$ функции АПВ, то Φ_1 можно выбирать, например, в виде положительно определенных функций $(\rho - g)$, $\rho\ln\frac{g}{\rho}$ (ядро функционала Кульбака), $|\rho - g|$ и др. Если известны пределы существования процесса (2) $X_* = [X_{\min}, X_{\max}]$, то функционал J_1 можно выбрать из условия минимума вероятности существования процесса (2) вне интервала X_*:

$$J_1 = 1 - \int_{X_*} \rho\,\mathrm{d}\mathbf{X} = \int_{-\infty}^{X_{\min}} \rho\,\mathrm{d}\mathbf{X} + \int_{X_{\max}}^{\infty} \rho\,\mathrm{d}\mathbf{X},$$

и т. д., исходя из особенностей решаемой задачи.

Таким образом, окончательно исследуемую задачу можно сформулировать как задачу поиска вектора $\mathbf{a}$, доставляющего минимум функционалу

$$J = \int_{X_*} \Phi_1(\mathbf{X},\rho)\mathrm{d}\mathbf{X} + \int_{t_0}^{t} \Phi_2(\mathbf{a})\mathrm{d}t, \qquad (3)$$

определенному на множестве функций АПВ ρ, удовлетворяющих решению известного уравнения Стратоновича для объекта (2) и наблюдателя (1):

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= -\mathrm{div}\left[\left(\mathbf{U}_0(\mathbf{X},t) + S(\mathbf{X})\mathbf{a}\right)\rho\right] + \\ &+ \frac{1}{2}\mathrm{div}\left[\overrightarrow{\mathrm{div}}(D_V\cdot\rho)\right] + \left[F(\mathbf{X},t) - F(t)\right]\rho = \\ &= -\mathrm{div}(S\mathbf{a}\rho) + Q(\rho,\mathbf{X},t),\end{aligned}$$

где div — символ операции дивергенции; $\overrightarrow{\mathrm{div}}$ — символ операции дивергенции, применяемой к каждой строчке матрицы;

$$F(\mathbf{X},t) = -\frac{1}{2}\left[\mathbf{Z} - \mathbf{H}(\mathbf{X},t)\right]^{\mathrm{T}} D_W^{-1}\left[\mathbf{Z} - \mathbf{H}(\mathbf{X},t)\right];$$

$$F(t) = \int_{-\infty}^{\infty} F(\mathbf{X},t)\rho(\mathbf{X},t)\mathrm{d}\mathbf{X}.$$

Так как

$$\begin{aligned}&-\mathrm{div}(S\mathbf{a}\rho) = \\ &= -\left[\mathrm{div}\left(S_{(1)}\rho\right) \;\ldots\; \mathrm{div}\left(S_{(2)}\rho\right) \;\ldots\; \mathrm{div}\left(S_{(k)}\rho\right)\right]\mathbf{a} = \\ &= S_0(\mathbf{X},\rho)\mathbf{a},\end{aligned}$$

где S_0 — вектор-строка, S_i — i-й столбец матрицы S, то приведенное уравнение АПВ будем использовать далее в виде:

$$\frac{\partial\rho}{\partial t} = S_0(\mathbf{X},\rho) + Q(\rho,\mathbf{X},t). \qquad (4)$$

РЕШЕНИЕ ЗАДАЧИ

Для решения поставленной задачи используем тот известный факт, что при неотрицательно определенной критериальной функции (как в рассмотренных выше вариантах) для обеспечения ее минимального значения в каждый момент времени достаточно, чтобы производная ее по времени, взятая с обратным знаком, имела максимум [3]. Отсюда для исследуемого случая получаем условие для определения идентифицируемого вектора параметров $\mathbf{a}$:

$$\max_{(\mathbf{a})}\left\{-\dot{J}\right\}=$$

$$=\max_{(\mathbf{a})}\left\{-\left(\int_{X_*}\frac{\partial\Phi_1(\mathbf{X},\rho)}{\partial\rho}\frac{\partial\rho(\mathbf{X},t)}{\partial t}\mathrm{d}\mathbf{X}+\Phi_2(\mathbf{a})\right)\right\}. \quad (5)$$

Подставляя в (5) выражение для правой части уравнения (4) имеем следующее уравнение относительно $\mathbf{a}$:

$$\frac{\partial}{\partial\mathbf{a}}\left\{\int_{X_*}\frac{\partial\Phi_1}{\partial\rho}\left(Q_0+S_0\mathbf{a}\right)\mathrm{d}\mathbf{X}+\Phi_2(\mathbf{a})\right\}=0 .$$

Из последнего вытекает окончательное уравнение для определения искомого вектора $\mathbf{a}$:

$$-\int_{X_*}\frac{\partial\Phi_1}{\partial\rho}S_0\mathrm{d}\mathbf{X}=\frac{\partial\Phi_2(\mathbf{a})}{\partial\mathbf{a}}, \quad (6)$$

решение которого осуществляется, исходя из конкретного вида функции Φ_2. Так например, для предложенной выше квадратичной формы функции $\Phi_2(\mathbf{a})$ уравнение (6) принимает вид

$$-\int_{X_*}\frac{\partial\Phi_1}{\partial\rho}S_0\mathrm{d}\mathbf{X}=2\mathbf{a}^{\mathrm{T}},$$

откуда

$$\mathbf{a}=-\frac{1}{2}\int_{X_*}\frac{\partial\Phi_1}{\partial\rho}\left(S_0\right)^{\mathrm{T}}\mathrm{d}\mathbf{X}. \quad (7)$$

Очевидно, что в результате проделанных построений функцию АПВ, обеспечивающую искомое оптимальное решение поставленной задачи адаптивной фильтрации, необходимо формировать из решения уравнения, полученного в результате подстановки выражения (7) в уравнение (4):

$$\frac{\partial\rho}{\partial t}=Q(\rho,\mathbf{X},t)-\frac{1}{2}S_0\int_{X_*}\frac{\partial\Phi_1(\mathbf{X},\rho)}{\partial\rho}S_0^{\mathrm{T}}(\mathbf{X})\mathrm{d}\mathbf{X}. \quad (8)$$

Важно при этом отметить, что уравнение (8) полностью по структуре совпадает с уравнением Стратоновича (4), являясь интегро-дифференциальным уравнением в частных производных. Более того, совпадение их размерностей приводит, по существу, к практически одинаковым вычислительным затратам при формировании тех или иных алгоритмов фильтрации на базе известных методов[1–4] (что при отсутствии расширения размерности оцениваемой системы за счет вектора неизвестных параметров $\mathbf{a}$ сокращает, как было отмечено выше, размерность интегрируемой системы оценок на $\frac{N^2}{2}+\left(\frac{3}{2}+M\right)N$ по сравнению с традиционной). Отсутствие же допущений о неизменности неизвестных параметров на интервале оценивания позволяет осуществить процедуру адаптивной фильтрации с принципиально большей точностью.

Для иллюстрации возможности эффективного практического использования предложенного подхода рассмотрим следующий пример.

ПРИМЕР

Стохастический нелинейный объект, описываемый уравнением

$$\dot{X}=a_1X+a_2X^2+a_3X^3+V, \quad X(0)=0.1,$$

где $a_1=3\cos 0.5t,\ a_2=2\sin t,\ a_3=4e^{-t}\sin 0.25t$ — параметры, неизвестные *apriori* наблюдателю; V — белый центрированный гауссовский шум с интенсивностью D_V, наблюдается с помощью измерителя

$$Z=1.5X^2+W,$$

где W — белый центрированный гауссовский шум с интенсивностью D_W.

Оценивание процесса X с одновременной идентификацией вектора $\mathbf{a}=\left[a_1,a_2,a_3\right]^{\mathrm{T}}$ требуется осуществить, исходя из условия минимума функционала, традиционно используемого в задачах идентификации динамических объектов с полностью параметрически неопределенной структурой:

$$J=\int_{-\infty}^{\infty}\left(Z-1.5X^2\right)^2D_W^{-1}\rho(X,t)\mathrm{d}X+\int_0^t\mathbf{a}^{\mathrm{T}}\mathbf{a}\mathrm{d}t.$$

Следуя приведенным рассуждениям, выражение для идентифицируемого вектора $\mathbf{a}$ было получено в виде

$$\mathbf{a}=\frac{1}{2}\int_{-\infty}^{\infty}\frac{\left(Z-1.5X^2\right)^2}{D_W}\begin{bmatrix}\rho+X\dfrac{\partial\rho}{\partial X}\\ 2X\rho+X^2\dfrac{\partial\rho}{\partial X}\\ 3X^2\rho+X^3\dfrac{\partial\rho}{\partial X}\end{bmatrix}\mathrm{d}X,$$

или, учитывая известное соотношение

$$\int_{-\infty}^{\infty}\Psi(X)\frac{\partial\rho}{\partial X}\mathrm{d}X=-\int_{-\infty}^{\infty}\frac{\partial\Psi}{\partial X}\rho\mathrm{d}X,$$

где Ψ — произвольная нелинейная функция,

$$\mathbf{a} = \frac{3}{D_W} \int_{-\infty}^{\infty} \left(Z - 1.5X^2\right) \begin{bmatrix} X^2 \\ X^3 \\ X^4 \end{bmatrix} \rho\left(X,t\right) \mathrm{d}X .$$

АПВ, в свою очередь, формировалась на основе решения уравнения

$$\frac{\partial \rho}{\partial t} = -\operatorname{div}\left(\frac{3}{D_W}\begin{bmatrix} X & X^2 & X^3 \end{bmatrix} \rho \int_{-\infty}^{\infty} \left(z - 1.5X^2\right) \begin{bmatrix} X^2 \\ X^3 \\ X^4 \end{bmatrix} \rho\left(X,t\right) \mathrm{d}X \right) +$$
$$+ \frac{1}{2} D_V \frac{\partial^2 \rho}{\partial X^2} - \frac{1}{2D_W}\left[\left(Z - 1.5X^2\right)^2 - \int_{-\infty}^{\infty} \left(Z - 1.5X^2\right)^2 \rho\left(X,t\right) \mathrm{d}X \right] \rho,$$

численное решение которого осуществлялось методом прямоугольных сеток с шагом $\Delta X = 0.5$ в области $X_* = \left[-30,30\right]$ на временнóм интервале $\left[0,40\right]$с с шагом формирования измерений $\Delta t = 0.1$с. Оценка процесса X при этом была сформирована по критерию максимума АПВ, определяемого на каждом временнóм шаге методом случайного поиска. Общая ошибка оценивания, усредняемая на интервале [30, 40]с, не превысила 18% от текущего значения процесса, при этом максимальные отклонения компонентов идентифицируемого вектора **a** от истинных значений составили: для a_1 — 7%, a_2 — 11%, a_3 — 5%. Затраченный при этом объем вычислений позволяет сделать вывод о возможности осуществления процедуры идентификации в масштабе времени поступления реальной измерительной информации, т.е. в реальном времени.

Таким образом, полученные результаты свидетельствуют об эффективности практического использования предложенного параметрического подхода как с точки зрения точности идентификации и оценивания, так и объема вычислительных затрат.

СПИСОК ЛИТЕРАТУРЫ

1. *Красовский А.А.* Справочник по теории автоматического управления. М.: Наука, 1987. 712 с.
2. *Сейдж Э.П., Мелса Дж.Л.* Идентификация систем управления. М.: Наука. 1974. 375 с.
3. *Казаков И.Е.* Статистическая теория систем управления в пространстве состояний. М.: Наука, 1975. 432 с.
4. *Первачев С.В., Перов А.И.* Адаптивная фильтрация сообщений. М.: Радио и связь, 1991. 220 с.

Ростовский военный институт ракетных войск, г. Ростов-на-Дону

Материал поступил в редакцию 14.04.2005.

THE USE OF NONLINEAR PROBABILISTIC CRITERIA TO SOLVE ADAPTIVE FILTERING PROBLEMS

S. V. Sokolov, I. V. Shcherban', V. A. Bertenev

Rostov Military Institute of Rocket Forces

The paper presents a general method for solution of the adaptive parametric filtering problem requiring much less computational effort at a potentially greater accuracy of estimation-identification as compared with the traditional methods. The accuracy increase is obtained at the expense of the use of more general probabilistic criteria instead of the mean -square criterion. An example of identifying a stochastic nonlinear object is given.