УДК 681.5.015

СЛОЖНЫЕ СТАТИСТИЧЕСКИЕ КРИТЕРИИ И МОДЕЛИ, ОПТИМАЛЬНЫЕ НА КЛАССЕ КРИТЕРИЕВ

И.С. Дургарян, Ф.Ф. Пащенко

Рассмотрены задачи моделирования и идентификации стохастических систем по сложным статистическим критериям. Исследованы условия эквивалентности критериев. Получены условия эквивалентности и оптимальности моделей, построенных по различным критериям идентификации.

ВВЕДЕНИЕ

Существенным моментом в решении задач идентификации и управления является выбор критерия. Заметим, что как в практических, так и в теоретических работах по идентификации выбор критерия определяется обычно не столько физическим смыслом задачи, сколько традициями, степенью полноты и характером исходных данных, степенью разработанности соответствующего математического аппарата и относительной сложностью необходимой вычислительной работы.

Замена одного критерия, физически обоснованного, другим, более удобным для решения, часто приводит к результатам, которые считаются приемлемыми. Видимо, этими причинами и объясняется тот факт, что обычно в качестве критериев идентификации выбираются критерии типа минимума среднеквадратической ошибки. Эти критерии, приводя к достаточно простым и отработанным вычислительным схемам, зачастую дают малочувствительные к изменению критерия качества функционирования объекта оценки параметров математических моделей. Причем эти оценки обладают рядом полезных статистических свойств. Однако критерии типа минимума среднеквадратической ошибки часто не связаны с конечной целью идентификации, т. е. с тем, для чего и как предполагается использовать математическую модель.

В связи с этим в идентификации и управлении возникает необходимость применения более сложных критериев. Например, при решении задач управления качеством управляющее воздействие обычно находится из условия равенства модельного (прогнозируемого) значения выходной переменной некоторому постоянному значению (заданию). Ясно, что, благодаря вероятностным свойствам объекта управления, задание в точности никогда не будет выполнено, и можно говорить лишь о вероятности выполнения задания. Требования к качеству выходного продукта задаются обычно не в виде конкретного значения того или иного показателя качества, а в форме допуска или допусков, соответствующих различным сортам. В этом случае в качестве критерия идентификации естественен критерий максимума вероятности нахождения ошибки $E(t) = Y(t) - Y_{\text{М}}(t)$, где $Y(t)$ и $Y_{\text{М}}(t)$ — выходные процессы объекта и модели, t — время, в некоторых допусках, а критерием для формирования управляющего воздействия (или воздействий) может служить максимум вероятности нахождения выходной переменной в допусках, определяемых требованиями к качеству выходного продукта [1, 2].

Известен хрестоматийный пример: средняя температура больных в больнице, не отражающая истиной картины ни общего состояния больных, ни состояния конкретного больного. В диагностике нас обычно интересует вероятность данного состояния, в задаче поражения цели — вероятность ее поражения и т. п. Вероятностные критерии, на наш взгляд, более реалистичные, более полно отражают реальные условия функционирования объекта, например, при моделировании состояния человека или в диагностических задачах.

Они допускают обобщения, позволяющие охватить широкий круг практически встречающихся задач моделирования [1, 3, 4]. Если выходной продукт подразделяется на несколько сортов, то применение вероятностного критерия для формирования управления позволяет максимизировать вероятность получения продукта требуемого сорта, минимизировать вероятность получения бракованной продукции и т. п. Если каждому сорту поставлена в соответствие некоторая экономическая характеристика (например, отпускная цена), то вероятностному критерию может быть придан экономический смысл.

Многие задачи такого типа встречаются в биологии, экономике и медицине. Аналогичные вопросы возникают при идентификации следящих и измерительных систем, создании систем принятия решений и построении прогнозирующих моделей.

Интересно отметить, что вероятностное обоснование метода наименьших квадратов, данное К. Гауссом, состоит в том, что получаемая оценка значений зависимой переменной является наилучшей в смысле критерия максимума вероятности нахождения ошибки в заданных пределах.

Безусловно, идентификация объектов по сложным статистическим критериям сложнее, однако они глубже отражают конечную цель идентификации. Например, при построении математической модели функционирования человеческого организма в критических ситуациях (предынфарктное состояние и т. п.) более естественно принять в качестве критерия идентификации критерий максимума вероятного нахождения ошибки в заданных пределах.

1. КРИТЕРИИ ИДЕНТИФИКАЦИИ

Рассмотрим следующий критерий идентификации:

$$J = M[\varphi(E)] \to \text{extr}, \tag{1}$$

где M – знак математического ожидания, $\varphi(E)$ – заданная функция. Из него при $\varphi(E) = E^2$ следует критерий минимума дисперсии ошибки идентификации (при $M(E) = 0$) или критерий минимума квадратической ошибки (если $M(E) \neq 0$). В том случае, когда

$$\varphi(E) = \begin{cases} 0, & E \,\overline{\in}\, [d_1, d_2] \\ 1, & E \in [d_1, d_2], \end{cases}$$

этот критерий представляет собой вероятность того, что ошибка идентификации находится в заданных допусках: $M[\varphi(E)] = P(d_1 \leqslant E \leqslant d_2)$, где d_1 и d_2 – границы допуска.

Критерий (1) обладает достаточно большой общностью, но его применение для решения задач идентификации сопряжено со значительными трудностями, обойти которые позволяют методы теории оптимальных статистических систем, развитые в работах В.С. Пугачева, Н.И. Андреева, И.Е. Казакова, В.В. Солодовникова и др.

Задача идентификации по вероятностному критерию является частным случаем более общей задачи идентификации по сложному статистическому критерию вида

$$J = f(J_1, \ldots, J_n) \to \text{extr}, \tag{2}$$

где f – заданная функция некоторых функционалов J_1, ..., J_n от параметров модели и объекта.

Например, при $J_1 = m_E^2$, $J_2 = D_E$ (m_E – математическое ожидание, а D_E – дисперсия ошибки) и гауссовских сигналах на входе и выходе системы вероятностный критерий при малом $d = d_1 = d_2$ примет вид: $J = f(J_1, J_2) = 2d\exp\{-J/2J_2\}/\sqrt{2\pi D_E}$, а критерий минимума среднеквадратической ошибки запишется в виде: $J = J_1 + J_2$.

В настоящей статье для решения задач идентификации по критерию (2) существенно используются результаты Н.И. Андреева по линейным статистическим оптимальным системам [1].

В пп. 5 и 6 рассматриваются случаи, когда построенные по разным критериям модели являются равносильными, т. е. оптимальными на классе критериев.

2. ИДЕНТИФИКАЦИЯ ЛИНЕЙНЫХ ОБЪЕКТОВ

Рассмотрим стандартную схему идентификации. В качестве критерия близости оценки оператора объекта к его истинному значению примем критерий экстремума квадратического функционала вида:

$$J = m_E + \sum_{i=1}^{n} \xi_i K_E(\tau_i) \tag{3}$$

где $K_E(\tau_i)$ – значение корреляционной функции ошибки в момент t, ξ_i – некоторые постоянные коэффициенты, удовлетворяющие условию $\sum_{i=1}^{n} \xi_i^2 \neq 0$. Оптимальную оценку оператора объекта будем искать в классе линейных интегральных стационарных операторов:

$$Y(t) = AX(t) = \int_{-\infty}^{t} g(\tau)X(t-\tau)d\tau, \tag{4}$$

где $g(\tau)$ – весовая функция оператора A. Предположим, что статистические характеристики входного $X(t)$ и выходного $Y(t)$ процессов объекта известны или вычислены по данным нормальной эксплуатации, а сами случайные процессы $X(t)$ и $Y(t)$ – стационарны и стационарно связаны. Выражая функционал (3) через исходные данные и учитывая физическую реализуемость системы – $g(\tau) = 0$ при $\tau < 0$, $\tau > T$, получаем:

$$J = m_E \int_0^T g(\tau)d\tau - m_Y + \sum_{i=1}^{n} \xi_i \Bigg[\int_0^T \int_0^T K_X(\tau_i - \lambda + \tau) g(\tau) d\tau d\lambda - 2\int_0^T K_{XY}(\tau_i + \tau) g(\tau) d\tau + K_Y(\tau_i) \Bigg], \tag{5}$$

где m_X и m_Y – математические ожидания входного и выходного сигналов, $K_X(\cdot)$ и $K_{XY}(\cdot)$ – авто- и взаимно корреляционная функции случайных процессов $X(t)$ и $Y(t)$. Выражение (5) представляет собой функционал от весовой функции модели типа:

$$J = \int_{t_0}^{t_1}\int_{t_0}^{t_1} \Psi[t, s, g(t), g(s)]dtds, \tag{6}$$

экстремали которого удовлетворяет уравнению [1]

$$\int_{t_0}^{t_1} \left(\frac{\partial \Psi}{\partial g_t} + \frac{\partial \Psi}{\partial g_s}\right) ds = 0, \quad 0 \leqslant t \leqslant T. \tag{7}$$

Решение этого уравнения не всегда принадлежит классу непрерывных функций, в общем случае оно лежит в классе обобщенных функций. Если функционал ψ представляет собой симметричную билинейную форму относительно функций $g(t)$ и $g(s)$, то интегральное уравнение (7), определяющее экстремаль $g_0 = g_0(t)$, принимает вид

$$\int_{t_0}^{t_1} \frac{\partial \Psi}{\partial g_t} ds = 0, \quad 0 \leqslant t \leqslant T. \tag{8}$$

При некоторых условиях выражение (8) представляет собой интегральное уравнение Фредгольма первого рода и решение его в общем случае лежит в классе обобщенных функций, содержащих разрывы и δ-функции. Если выражение (8) представляет собой уравнение Фредгольма второго ряда, то оно имеет решение в классе непрерывных функций. Функционалы (3), фигурирующие в критерии идентификации, являются частными случаями функционалов (6).

В нашем случае подынтегральная функция имеет вид:

$$\Psi(\tau, \lambda, g(\tau), g(\lambda)) = \sum_{i=1}^{n} \xi_i K_X(\tau - \lambda + \tau) g(\tau) g(\lambda) - $$
$$- \frac{1}{T} \sum_{i=1}^{n} \xi_i [K_{XY}(\tau_i + \tau) g(\tau) + K_{XY}(\tau_i + \lambda) g(\lambda)] +$$
$$+ \frac{m_X}{2T}[g(\tau) + g(\lambda)]$$

и представляет собой симметричную билинейную форму от функций $g(\tau)$ и $g(\lambda)$. Решение задачи идентификации системы (линейной или нелинейной) в классе моделей (4) по критерию (3) дается следующей теоремой.

Теорема 1. Пусть идентифицируемый объект — линейный, стационарный и физически реализуемый ($g(\tau) \neq 0$, $\tau \in [0, T]$, $g(\tau) = 0$, $\tau < 0$, $\tau > T$), случайные процессы на входе $X(t)$ и выходе $Y(t)$ объекта стационарны и стационарно связаны. Пусть выполняются условия: $\sum_{i=1}^{n} \xi_i^2 \neq 0$; $\sum_{i=1}^{n} |K_E(\tau_i)| \neq 0$; весовая функция $g(\tau)$ интегрируема на отрезке $[0, T]$. Тогда: существует модель объекта в классе линейных стационарных физически реализуемых операторов типа (4), оптимальная по критерию экстремума функционала (3); оценка $g_0(\lambda)$ весовой функции модели объекта удовлетворяет линейному интегральному уравнению

$$\int_0^T \sum_{i=1}^{n} \xi_i K_X(\tau_i - \lambda + \tau) g(\tau) d\tau\, d\lambda -$$
$$- \sum_{i=1}^{n} \xi_i K_{XY}(\tau_i + \tau) + \frac{m_X}{2} = 0. \tag{9}$$

Доказательство теоремы следует из решения уравнения (8) при условии, что функционал Ψ определяется выражением (5).

Уравнение идентификации (9) представляет собой линейное интегральное уравнение Фредгольма первого рода. Его решение можно найти известными методами [3].

3. ИДЕНТИФИКАЦИЯ НЕЛИНЕЙНЫХ ОБЪЕКТОВ

Методы идентификации по сложным статистическим критериям, рассмотренные в п. 2, применимы для построения оптимальных оценок операторов линейных и нелинейных объектов в классе линейных операторов, в классе линейных в среднем операторов и операторов Гаммерштейна. Однако для нелинейных стохастических объектов корреляционные функции не являются исчерпывающими характеристиками связи между входными и выходными сигналами. В общем случае изложенный выше корреляционный метод идентификации будет давать заниженную оценку статистической связи входных и выходных сигналов, а, следовательно, приводить к ошибкам решения задачи идентификации.

Рассмотрим состоятельный метод идентификации по сложным статистическим критериям, основанный на идеях метода функциональных преобразований [5]. Для простоты рассмотрим случай стационарного объекта, хотя все изложенное очевидным образом распространяется и на нестационарный случай. В качестве критерия близости модели и оператора объекта примем экстремум билинейного функционала вида:

$$J = m_E + \sum_{i=1}^{n} \xi_i R_E^{\phi}(\tau_i), \tag{10}$$

где $R_E^{\phi}(\tau_i)$ — значение функциональной корреляционной функции ошибки идентификации в момент τ_i [5], ξ_i — некоторые постоянные коэффициенты, удовлетворяющие условию $\sum_{i=1}^{n} \xi_i^2 \neq 0$. Весовые коэффициенты ξ_i учитывают вклад значений функции $R_E^{\phi}(\tau_i)$ в различные моменты времени τ_i или интервалы времени, если $\tau_i = i\Delta$, где Δ — интервал времени, а $i = 1, 2, ..., n$. Задавая различные значения коэффициентов ξ_i, получаем семейство функционалов вида (10), которые можно использовать для приближенной оценки качества идентификации. Критерии, использующие функционалы типа (10), эффективны не только при идентификации по вероятностным критериям, но и в том случае, когда надо построить модель, наилучшим образом приближающую идентифицируемый объект на заданном отрезке времени, или режиме функционирования, или в данный момент времени, при решении задачи терминального управления. Аналогичные задачи возникают не только при идентификации с целью диагностики критических режимов, но и при идентификации сложных объектов, имеющих несколько, в том числе и стационарных, режимов функционирования, например, ядерный реактор, доменная печь или социально-экономическая система.

Оптимальную модель нелинейного объекта будем искать в классе стационарных полулинейных систем $\pi(N_1 - L - N_2)$ [5]:

$$B[Y(t)] = \int_{-\infty}^{t} g(\tau) C[X(t - \tau)] d\tau. \tag{11}$$

В соответствии с методом функциональных преобразований, опираясь на результаты работы [5], сначала найдем преобразования B и C случайных процессов на выходе и входе объекта, определяющие нелинейные звенья модели. Далее, воспользуясь изложенной методикой, найдем оптимальную весовую функцию линейной части модели.

Пусть случайные процессы $X(t)$ и $Y(t)$ — ϕ-стационарны и ϕ-стационарно связаны [5]. Тогда, с учетом физической реализуемости линейного звена, функциональная (обобщенная) автокорреляционная функция ошибки идентификации имеет вид

$$R_E^{\phi}(\tau_i) = \int_0^T\int_0^T R_X^{\phi}(\tau_i - \lambda + \tau)g(\tau)g(\lambda)d\tau\,d\lambda - $$
$$- 2\int_0^T R_{XY}^{\phi}(\tau_i + \tau)g(\tau)d\tau + R_Y^{\phi}(\tau_i).$$

Функционал (10) в этом случае можно представить в виде

$$J = m_E + \sum_{i=1}^{n}\xi_i\left[\int_0^T\int_0^T R_X^{\phi}(\tau_i - \lambda + \tau)g(\tau)g(\lambda)d\tau\,d\lambda - \right.$$
$$\left. - 2\int_0^T R_{XY}^{\phi}(\tau_i + \tau)g(\tau)d\tau + R_Y^{\phi}(\tau_i)\right], \qquad (12)$$

где m_E — математическое ожидание ошибки идентификации, $R_X^{\phi}(\cdot)$, $R_Y^{\phi}(\cdot)$ и $R_{XY}^{\phi}(\cdot)$ — соответственно обобщенные авто- и взаимная корреляционные функции случайных процессов $X(t)$ и $Y(t)$ [5].

Выражение (12) представляет собой функционал от весовой функции модели типа (6). Учитывая, что математическое ожидание обобщенной ошибки идентификации при сделанных выше предположениях можно записать в виде

$$m_E = m_{CX}\int_0^T g(\tau)d\tau - m_{BY},$$

получим следующее выражение для подынтегральной функции функционала (6):

$$\Psi(\tau, \lambda, g(\tau), g(\lambda)) = \sum_{i=1}^{n}\xi_i R_X^{\phi}(\tau_i - \lambda + \tau)g(\tau)g(\lambda) - $$
$$- \frac{1}{T}\sum_{i=1}^{n}\xi_i[R_{XY}^{\phi}(\tau_i + \tau)g(\tau) + R_{XY}^{\phi}(\tau_i + \lambda)g(\lambda)] + $$
$$+ \frac{m_{CX}}{2T}[g(\tau) + g(\lambda)],$$

где $m_{CX} = M\{C(X(t))\}$; $m_{BY} = M\{B(Y(t))\}$.

Функция Ψ представляет собой симметричную билинейную форму от функций $g(\tau)$ и $g(\lambda)$. Решение задачи идентификации нелинейного объекта в этом случае определяется следующей теоремой.

Теорема 2. Пусть идентифицируемый объект представляет собой нелинейный динамический объект. Случайные процессы на входе $X(t)$ и выходе $Y(t)$ объекта ϕ-стационарны и ϕ-стационарно связаны. Модель объекта ищется на множестве полулинейных систем вида (11). Пусть выполняются условия: $\sum_{i=1}^{n}\xi_i^2 \neq 0$; существуют преобразования B и C такие, что автокорреляционная функция $R_E^{\phi}(\tau)$ не равна тождественно нулю на отрезке $[0, T]$; весовая функция линейной части модели (11) удовлетворяет условиям физической реализуемости и интегрируема на отрезке $[0, T]$.

Тогда: существует модель объекта в классе полулинейных систем вида (11), оптимальная по критерию экстремума функционала (10); оценка оптимальной весовой функции $g_0(\tau)$ линейной части модели (11) удовлетворяет линейному интегральному уравнению идентификации

$$\int_0^T\sum_{i=1}^{n}\xi_i R_X^{\phi}(\tau_i - \lambda + \tau)g(\lambda)\,d\lambda - $$
$$- \sum_{i=1}^{n}\xi_i R_{XY}^{\phi}(\tau_i + \tau) + \frac{m_{CX}}{2} = 0. \qquad (13)$$

Уравнение (13) представляет собой интегральное уравнение Фредгольма первого рода, решение которого в общем случае лежит в классе обобщенных функций, содержащих не только непрерывные, но и разрывные функции и δ-функции.

Замечание. Выполнение равенства (13) является необходимым и достаточным условием экстремума функционала (10). Однако это условие может соответствовать как минимуму, так и максимуму функционала (10). На существование минимума или максимума критерия влияют конкретные выбранные значения параметров ξ_i.

Если существует максимальная взаимная корреляционная функция случайных процессов $X(t)$ и $Y(t)$, то справедлива следующая

Теорема 3. Пусть выполняются условия теоремы 2. Пусть существует и не равна тождественно нулю на отрезке $[0, T]$ максимальная взаимная корреляционная функция $R_{XY}^{\max}(\tau)$. Тогда: существует модель объекта в классе полулинейных систем вида (11), оптимальная по критерию экстремума функционала

$$J = m_E + \sum_{i=1}^{n}\xi_i R_{XY}^{\max\phi}(\tau_i),$$

где обозначение $R_{XY}^{\max\phi}(\tau_i)$ указывает на то, что ошибка идентификации $E(t)$ определяется при условии, что преобразования $B(Y(t))$ и $C(X(t))$ удовлетворяют теореме 7.2.3 (см. работу [5]); оценка оптимальной весовой функции $g_0(\tau)$ линейной части модели (11) удовлетворяет линейному интегральному уравнению идентификации

$$\int_0^T\sum_{i=1}^{n}\xi_i R_X^{\phi}(\tau_i - \lambda + \tau)g(\lambda)\,d\lambda - $$
$$- \sum_{i=1}^{n}\xi_i R_{XY}^{\max}(\tau_i + \tau) + \frac{m_{CX}}{2} = 0. \qquad (14)$$

Доказательство теорем 2 и 3 осуществляется в два этапа. На первом этапе, учитывая результаты работ [6—8] находят преобразования B и C случайных процессов на выходе $Y(t)$ и входе $X(s)$ объекта, определяющие обобщенные авто- и взаимные корреляционные функции и максимальную взаимную корреляционную функцию случайных процессов $X(s)$ и $Y(t)$. На втором этапе по теореме 1 находят оптимальную весовую функцию линейной части модели.

Как важные частные случаи из теорем 2—3 следуют дисперсионные методы идентификации и синтеза оптимальных стохастических систем по сложным статистическим критериям, рассмотренным в работах [2, 9]. Заметим, что при идентификации нелинейных объектов, описываемых монотонными зависимостями, уравнения дисперсионной идентификации совпадают с уравнениями (13) и (14). Легко показать, что теорема 1 также является частным случаем теорем 2 и 3.

4. ИДЕНТИФИКАЦИЯ СИСТЕМ ПО ВЕКТОРНОМУ КРИТЕРИЮ

Постановка задачи идентификации, особенно в случае сложных иерархических систем, часто приводит к векторному критерию построения оптимальной модели:

$$J = (J_1, ..., J_n), \tag{15}$$

т. е. к задаче многокритериальной оптимизации. При этом для определения области решений, например, оптимальных по Парето, необходимо перейти от задачи векторной оптимизации к задаче нелинейной оптимизации со специально сконструированной скалярной функцией цели:

$$J = f(J_1, ..., J_n). \tag{16}$$

Существуют различные способы образования такой функции, наиболее распространенный среди них метод взвешенных сумм:

$$J = \sum_{i=1}^{n} k_i J_i, \text{ где } k_i \geqslant 0 \text{ и } \sum_{i=1}^{n} k_i = 1.$$

Вопросы построения обобщенных критериев оптимальности рассмотрены в работах В.С. Пугачева, Н.И. Андреева, И.Е. Казакова и др.

Функционалы (15) и (16) охватывают довольно широкий класс критериев. Как частные случаи, они включают в себя критерии вероятности выхода ошибки из заданных допусков, минимума средней квадратической ошибки и т. п. Кроме того, некоторые из критериев J_i могут иметь смысл производительности, себестоимости, экономической эффективности. Идентификация объектов по критерию экстремума функционала (16) представляет собой трудную задачу, которую можно значительно упростить, воспользовавшись следующей модификацией результатов Н.И. Андреева по теории статистически оптимальных систем.

Теорема 4. Пусть в функционале (16) $J_1, J_2, ..., J_n$ — действительные выпуклые функционалы, определенные в банаховом пространстве. Пусть функционалы $J_1, ..., J_n$ — сильно дифференцируемы в области G, т. е. существует дифференциал Фреше — $dJ(g)$, $g \in G$. Пусть функция $f(\cdot)$ — выпуклая, монотонная и дифференцируемая в области G. Тогда для того, чтобы функционал J достигал экстремума на функции g_0, необходимо, чтобы $dJ(g_0, h) = J'(g_0)h = 0$ при всех $h \in G$; функция g_0, на которой функционал J принимает экстремальное значение, принадлежит множеству решений задачи поиска экстремума приведенного функционала

$$J(g) = \sum_{i=1}^{n-1} k_i J_i(g) + J_n(g), \quad k_i \geqslant 0, \quad i = 1, ..., n-1. \tag{17}$$

Другими словами, для того чтобы функция $g(t)$, принадлежащая классу допустимых функций, соответствовала экстремуму функционала (16), необходимо, чтобы она соответствовала экстремуму приведенного функционала (17) при некоторых значениях согласующих параметров $k_1, ..., k_{n-1}$. Примерами выпуклых функционалов являются квадратические функционалы (3) и (10). Для выпуклых функционалов необходимое условие экстремума является и достаточным.

Функция $g_0(t, k_1, ..., k_{n-1})$, соответствующая экстремуму функционала (16), зависит от $n - 1$ согласующих параметров k_i. Подставляя найденное значение $g_0(t, k_1, ..., k_{n-1})$ в выражение (16), получим некоторую функцию параметров $k_1, ..., k_{n-1}$:

$$J = f[J_1(g(t, k_1, ..., k_{n-1})), ..., J_n(t, k_1, ..., k_{n-1}))]] = F(k_1, ..., k_{n-1}).$$

Определяя экстремум функционала J как функцию от $k_1, ..., k_{n-1}$, найдем соответствующие ему значения $k_1 = k_{1,0}, ..., k_{n-1,0}$. Тогда искомая оценка весовой функции объекта $g_0(t) = g(t, k_{1,0}, ..., k_{n-1,0})$. Способы определения значений согласующих параметров $k_1 = k_{1,0}, ..., k_{n-1} = k_{n-1,0}$, соответствующих экстремуму функционала (16), приведены в работах Н.И. Андреева по теории статистически оптимальных систем [1].

Если функция $F(k_1, ..., k_{n-1})$ имеет несколько максимумов или минимумов, то в качестве $\sup J$ или $\inf J$ следует брать, соответственно, наибольший из максимумов или наименьший из минимумов и соответствующие этим значениям оптимальные значения согласующих параметров. При этом необходимо, чтобы полученная в результате весовая функция модели принадлежала классу допустимых функций.

Отыскание экстремума функционала (17) по трудности равносильно отысканию экстремума функционала $\sum_{i=1}^{n} J_i$. Если, например, функционалы $J_1, ..., J_n$ — квадратические, то и функционал (17) тоже квадратический, и, следовательно, независимо от вида функции f, поиск экстремума функционала (17) сводится к задаче определения экстремума квадратического функционала, для решения которой имеется хорошо развитый аппарат. Таким образом, задача отыскания экстремума функционала (16) сводится к решению более простых задач: определению функции $g(t, k_1, ..., k_{n-1})$ и отысканию согласующих параметров $k_{1,0}, ..., k_{n-1,0}$.

В качестве функционалов $J_1, ..., J_n$ в выражении (16) могут использоваться функционалы типа (3) и (10), что приводит к критериям, имеющим смысл вероятности того, что ошибка идентификации находится в заданных допусках:

- в некоторый момент времени;
- в течение некоторого отрезка времени,
- к более сложным обобщениям этих критериев.

5. О МОДЕЛЯХ, ОПТИМАЛЬНЫХ ПО РАЗЛИЧНЫМ КРИТЕРИЯМ

Задача идентификации систем и построения моделей, оптимальных по вероятности выхода ошибки идентификации за пределы заданного интервала или по сложным статистическим критериям, рассмотренная в пп. 1—4, связана с серьезными математическими затруднениями, громоздкими и большими вычислительными затратами, в отличие от идентификации по среднеквадратическому критерию. Понятно, что критерий идентификации должен отражать цель создания модели, например, критерий последующей оптимизации системы или критерий управления идентифицируемым объектом. Однако, применение сложных статистических критериев не гарантирует, что в результате будет получен более существенный эффект, нежели при классическом среднеквадратическом критерии.

В работах В.С. Пугачева, Н.И. Андреева, О.М. Козлова, П. Эйкхоффа [1, 3, 10] и других исследователей показано, что для гауссовских процессов синтез оптимальных систем и идентификация объектов по различным критериям часто приводит соответственно к одному и тому же результату.

В связи с этим представляет интерес отыскание свойств систем, по которым можно было бы определить, будут совпадать или отличаться результаты идентификации по различным критериям.

Рассмотрим различные критерии близости выходных сигналов системы $Y(t)$ и модели $Y_{\text{М}}(t)$ вида:

$$J(\varphi) = J(\varphi(Y(t), AX(s)) = \\ = M[\varphi(Y_{\text{М}}(t) - Y(t))] \to \min_{Y_{\text{М}}}, \qquad (18)$$

или

$$J(\varphi) = M[\varphi(|Y_{\text{М}}(t) - Y(t)|)] \to \min_{Y_{\text{М}}}, \qquad (19)$$

где φ — произвольная неотрицательная неубывающая функция, $\varphi(\cdot) \neq \text{const}$, $Y_{\text{М}}$ — выходной сигнал модели $Y_{\text{М}}(t) = A[X(s), s \in T_t]$, где A — оператор модели, T_t — интервал времени.

Критерии вида (18) и (19) охватывают критерии минимума среднеквадратической ошибки $\varphi(\cdot) = E^2$, среднего значения модуля ошибки, среднее значение взвешенной суммы ошибки $\sum a_k|E|^k$, вероятности выхода ошибки идентификации из заданной области радиуса δ при

$$\varphi(E) = \begin{cases} 0, 0 \le E < \delta \\ 1, E \ge \delta \end{cases} \qquad (20)$$

и другие.

Поскольку при решении задач идентификации применяются методы, учитывающие знания о системе и действующих на нее сигналов, и дисперсионные методы идентификации [5], рассмотрим также функционалы вида

$$J(\varphi(\cdot)/X(s)) = M[\varphi(|Y_{\text{М}}(t) - Y(t)|)/X(s), s \in T_t] \quad (21)$$

и

$$J(\varphi(\cdot)/K(F, X(s), Y(t)) = \\ = M[\varphi(\cdot)/K(F(\cdot), X(s), Y(t)], \qquad (22)$$

где, как и ранее, M[·/·] означает условное математическое ожидание функционала $\varphi(\cdot)$ относительно реализации случайного сигнала $x(s)$ или знаний $K(\cdot)$ о системе $F(\cdot)$, входном $X(s)$ и выходном сигнале $Y(t)$ системы.

Пусть Φ — класс моделей (например, линейных, нелинейных, полулинейных), в котором ищется идентифицируемая модель системы. Требуется найти модель системы из класса Φ, оптимальную в смысле критериев (18) и (19) или (21) и (22).

6. СУЩЕСТВОВАНИЕ ОПТИМАЛЬНЫХ МОДЕЛЕЙ

Определение оптимальной модели по данному критерию или по набору критериев на заданных классах моделей, входных и выходных сигналов является одним из основных этапов решения задачи идентификации в широком смысле. Однако прежде чем приступить к построению оптимальной модели, необходимо ответить на вопрос: существует ли оптимальная модель идентифицируемой системы? В данной работе будем искать оптимальные модели в смысле следующих определений.

Определение 1. Модель A^0 называется оптимальной «в среднем» в классе Φ, если $A^0 \in \Phi$ и для любых $A \in \Phi$, $X(s)$ и любого критерия (18) или (19) выполняется неравенство $J(\varphi, A^0) \leqslant J(\varphi, A)$.

Определение 2. Модель A^0 называется оптимальной «в целом» в классе Φ, если $A^0 \in \Phi$ и для любых $A \in \Phi$, $X(s)$ и любого критерия (21) или (22) выполняются неравенства $J(\varphi, A^0/X(s)) \leqslant J(\varphi, A/X(s))$ и $J(\varphi, A^0/K(\cdot)) \leqslant J(\varphi, A/K(\cdot))$.

Из этих определений следует, что модель, оптимальная «в целом», дает наилучшую аппроксимацию выходного сигнала системы не только «в среднем» по всем возможным реализациям $X(s)$, но и для каждой конкретной реализации $x(s)$ или набора знаний $K(\cdot)$ в отдельности. Конечно, возникает вопрос: существуют ли такие модели? Ответ дают следующие теоремы.

Теорема 5. Пусть условное распределение процесса $Y(t)$ относительно процесса $X(s)$

$$F_{y/x}(u) = P(Y(t) \leqslant u/X(s), s \in T_t) \qquad (23)$$

унимодально и симметрично. Тогда существует оптимальная модель «в целом», ее выходная реакция $Y_{\text{М}}(t) = M(Y(t)/X(s), s \in T_t)$.

Справедливо и обратное утверждение.

Теорема 6. Пусть существует оптимальная «в целом» модель. Тогда соответствующие условные распределения (23) унимодальны и симметричны.

Доказательства теорем 5 и 6 следуют из работ [3, 4].

При доказательстве теорем существенен учет свойств унимодальности и симметричности (функция распределения $F(u)$ называется унимодальной, если существует такое число u_0, что $F(u)$ выпукла при $u < u_0$ и вогнута при $u > u_0$; функция распределения называется симметричной, если $F(u_0 - u) = 1 - F(u_0 + u)$).

Как известно, унимодальные функции распределения непрерывны и обладают производными слева и справа в каждой точке, за исключением, быть может, одной точки вершины, где возможен разрыв.

Примером систем, соответствующих теоремам 5 и 6 являются системы, случайные процессы на входе $X(s)$ и выходе $Y(t)$ которых представляют собой гауссовские процессы $N_1(m_x, \Sigma_1)$ и $N_2(m_y, \Sigma_2)$.

Следствие 1. Если условная функция распределения $F_{y/x}(u)$ имеет более одной точки разрыва, то $F_{y/x}(u)$ не унимодальна и оптимальной «в целом» модели не существует.

Следствие 2. Если условная функция распределения $F_{y/x}(u)$ несимметрична, то оптимальной «в целом» модели не существует.

В случае выполнения условий следствий 1 и 2 оптимальная модель по среднеквадратической ошибке не совпадает с моделью, полученной по критерию минимума выхода ошибки идентификации из заданной области (интервала). В частности, рассмотренный в работе [2] пример нелинейного квадратического преобразователя $Y(t) = X^2(t)$ является примером нарушения условия симметричности условной функции распределения $F_{y/x}(u)$. Следовательно, для идентификации по сложным статистическим критериям в этом случае следует применять методы идентификации, рассмотренные в работах [2, 9].

Таким образом, существование оптимальной модели, вообще говоря, не дает оснований заменять при идентификации одни критерии другими. Условия, при которых такая замена возможна, определяются следующими теоремами.

Теорема 7. Пусть выполнены условия теоремы 5 и условие а) плотность $p_{y/x}(u)$ условного распределения случайного процесса $Y(t)$ относительно наблюдаемых значений $X(s)$ при $u > u_0 = M[Y(t)/X(s),\ s \in T_t]$ строго убывающая функция или условие б) функция $\varphi_0(E)$ строго возрастающая. Тогда существует и единственна оптимальная «в целом» модель идентифицируемой системы, определяемая уравнением

$$J(\varphi_0, A^0/X(s)) = \min_A J(\varphi_0, A/X(s)).$$

Доказательство. Пусть выполняются условия теоремы 5 и условие а) и существуют оптимальные модели

$$A^0(X(\tau)) = M\{Y(t)/X(\tau),\ \tau \in T_t\} = a_0,$$

$$B^0(x(\tau)) = b_0 \neq a_0.$$

Пусть $p_{y/x}(u) = F'_{y/x}(u + a_0)$ – условная плотность распределения.

В силу условий (23) и а) теоремы

$$p(u) = p(-u), \quad p(u_2) > p(u_1) \text{ при } u_2 > u_1 > 0 \qquad (24)$$

Тогда имеем:

$$J(\varphi_0, A^0/X(\tau)) = \int_{-\infty}^{\infty} \varphi_0(y - a_0)p(y - a_0)dy = \int_{-\infty}^{\infty} \varphi_0(u)p(u)du,$$

$$J(\varphi_0, B^0/X(\tau)) = \int_{-\infty}^{\infty} \varphi_0(y - b_0)p(y - a_0)dy =$$

$$= \int_{-\infty}^{\infty} \varphi_0(u + 2c)p(u)du,$$

где $c = (a_0 - b_0)/2$. Отсюда

$$J(\varphi_0, A^0/X(\tau)) - J(\varphi_0, B^0/X(\tau)) =$$

$$= \int_{-\infty}^{\infty} [p(u - c) - p(u + c)][\varphi_0(u + c) - \varphi_0(u - c)]du = 0.$$

Поскольку последний интеграл равен нулю, то подынтегральное выражение (в силу своей не отрицательности) равно нулю

$$[p(u - c) - p(u + c)]\varphi_0(u + c) - \varphi_0(u - c)] = 0$$

почти при всех u. В силу условий (24) первый сомножитель строго больше нуля, следовательно

$$\varphi_0(u + c) = \varphi_0(u - c) \text{ – почти всюду.}$$

Последнее невозможно, так как функция $\varphi_0(u)$ отлична от нуля и монотонна (см. формулы (18) и (19)). Следовательно, оптимальная модель A^0 единственна. Аналогично доказывается единственность оптимальной модели при условиях (23) и б).

Теорема 8. Пусть выполняются условия теоремы 7. Пусть существует модель B^0, оптимальная «в среднем» по критерию (19):

$$J(\varphi_0, B^0) = J(\varphi_0, A^0) = \min_A J(\varphi, A). \qquad (25)$$

Тогда выходные сигналы моделей A^0 и B^0 тождественны почти всюду для всех реализаций входного сигнала $X(s)$; модель B^0 оптимальна по всем критериям $J(\varphi, A)$ (19).

Доказательство. Первое утверждение теоремы следует из теорем 5 и 7. Действительно, в силу теоремы 5 оптимальная модель существует, а в силу теоремы 7 она единственна.

Докажем второе утверждение. Из теоремы 7 следует

$$J(\varphi_0, B^0/X(\tau)) \geqslant J(\varphi_0, A^0/X(\tau)). \qquad (26)$$

В силу условия (25)

$$J(\varphi_0, B^0) = J(\varphi_0, A^0). \qquad (27)$$

Выражения (26) и (27) не противоречат друг другу только в случае, когда с вероятностью 1 выражение (26) превращается в равенство. Отсюда в силу единственности оптимальной модели $B^0(\cdot) = A^0(\cdot)$ с вероятностью 1.

Поскольку значения условного критерия $J(\varphi, A/x)$ на множестве реализаций $X(\tau)$ вероятностной меры нуль не влияют на значение критерия $J(\varphi, A)$, то

$$J(\varphi_0, B^0) = J(\varphi_0, A^0) = \min_A J(\varphi, A)$$

что и требовалось доказать.

Из теорем 5—8 следует, что если условные распределения унимодальны и симметричны, то модель, оптимальная по минимуму среднеквадратического отклонения ошибки идентификации, оптимальна по всем критериям с функционалом типа (18), в том числе и по критерию минимума вероятности выхода ошибки идентификации из заданного диапазона: $P\{|Y_{\text{M}}(t) - Y(t)| > \delta\}$ при любом δ.

Полученная при этом модель A^0 отличается от оптимальной «в целом» модели лишь на множестве входных сигналов меры нуль.

Если модель A^0 обеспечивает минимум среднеквадратического отклонения для каждой реализации входного сигнала $X(s)$, т. е. выполняется условие (20), то модель A^0 оптимальна «в целом».

Отметим, что условие а) теоремы 7 выполняется для гауссовского процесса, а условие б) — для степенной функции, т. е. для среднеквадратического, среднекубического отклонений и т. п. Поэтому все вышесказанное справедливо для моделей оптимальных по среднему значению любой степени n отклонения: $|Y_{\text{M}}(t) - Y(t)|^n$.

Необходимыми условиями существования модели, оптимальной «в среднем» по всем критериям типа (19), являются унимодальность и симметричность распределения следующей случайной величины: $\xi(t) = Y(t) - M\{Y(t)/X(s), s \in T_t\}$. Наличие этого свойства способствует высокой эффективности дисперсионных моделей и методов идентификации.

ЗАКЛЮЧЕНИЕ

Один из основных этапов решения задачи идентификации в широком смысле заключается в определении оптимальной модели системы на заданных классах моделей, входных и выходных сигналов. Рассмотрены вопросы выбора и обоснования критериев при моделировании систем. Предложены методы идентификации по сложным вероятностным критериям. Показано, что задача идентификации по векторным критериям в ряде случаев сводится к задаче идентификации по приведенным билинейным критериям, впервые рассмотренным Н.И. Андреевым [1].

Даны определения моделей, оптимальных «в среднем» и «в целом» на выбранном классе моделей. Получены условия существования моделей, оптимальных по разным критериям идентификации. Показано, что если существует оптимальная «в целом» модель, то условные распределения выходного сигнала системы относительно входного сигнала должны быть унимодальны и симметричны. Рассмотрен частный случай, когда входной и выходной сигналы системы являются гауссовскими. Если указанные выше условия не выполняются, идентификацию систем следует проводить методами, изложенными в пп. 2—4 и в работах [2, 9].

Изложенный подход к исследованию и применению сложных статистических критериев для решения задач идентификации может быть полезен и для решения задач синтеза оптимальных стохастических систем, прогнозирования, экстраполяции и фильтрации.

ЛИТЕРАТУРА

1. *Андреев Н.И.* Корреляционная теория статистически оптимальных систем. — М.: Наука, 1966. — 386 с.
2. *Дургарян И.С., Пащенко Ф.Ф.* Идентификация нелинейных объектов по сложным критериям // Автоматика и телемеханика. — 1980. — № 7. — С. 61—71.
3. *Пугачев В.С.* Теория случайных функций. — М.: Физматгиз, 1962.
4. *Бернацкий Ф.И., Пащенко Ф.Ф., Коновалова Т.Р.* Идентификация и управление технологическими объектами по сложным критериям. — Владивосток: ИАПУ ДВНЦ АН СССР, 1983. — С. 34.
5. *Прангишвили И.В., Пащенко Ф.Ф., Бусыгин Б.П.* Системные законы и закономерности в электродинамике, природе и обществе. — М.: Наука, 2001. 526 с.
6. *Сарманов О.В.* Собственные корреляционные функции и их применение в теории стационарных марковских процессов // Докл. АН СССР. 1960. — Т. 132, № 4. — С. 769—772.
7. *Пащенко Ф.Ф., Чернышев К.Р.* Применение метода функциональных преобразований в идентификации нелинейных систем // Автоматика и телемеханика. — 1992. — № 4. — С. 77—85.
8. *Breiman L., Friedman J.H.* Estimating Optimal Transformations for Multiple Regression and Correlation //J. of the American Statistical Association. — 1985. Vol. 80, № 391. — P. 580—598.
9. *Дургарян И.С., Пащенко Ф.Ф.* Дисперсионный критерий статистической оптимизации систем // Автоматика и телемеханика. — 1974. — № 12. С. 46—52.
10. *Эйкхофф П.* Основы идентификации систем управления. — М.: Мир, 1975.

☎ *(095) 334-90-20*

E-mail: feodor@ipu.rssu.ru

❑