



ПРОБЛЕМА МАТЕМАТИЧЕСКОЙ ИНТЕРПРЕТАЦИИ ДАННЫХ. I. СИСТЕМЫ С СОСРЕДОТОЧЕННЫМИ ПАРАМЕТРАМИ

Аннотация. Рассмотрена проблема интерпретации данных, полученных в экспериментальных исследованиях, как неклассическая математическая задача, которая в большинстве случаев является в широком смысле некорректно поставленной. Использована дополнительная информация об объекте в виде уравнений локальных связей, определяющих его замкнутую или незамкнутую математическую модель. Описаны процедуры регуляризации, позволяющие находить практически пригодные решения, согласованные с имеющимися данными.

Ключевые слова: экспериментальные данные, математическая модель, интерпретация, некорректность, регуляризация, редукция модели, обобщенное решение, вариационный метод.

ВВЕДЕНИЕ

В математике различают корректно и некорректно поставленные задачи. К первым относят такие, для которых установлена их точная разрешимость, т.е. доказано существование и единственность получаемого решения. Кроме того, если исходные данные содержат погрешность, то важнейшим фактором является устойчивость решения, а именно его чувствительность к допустимым вариациям данных. Это характеризует практическую пригодность решений, с которой ассоциируется так называемая некорректность по Адамару. К некорректно поставленным задачам часто относят также такие, для которых невозможно установить их корректность, т.е. доказать существование, единственность и устойчивость искомого решения. Объясняется это обстоятельство тем, что методы регуляризации, используемые для нахождения решения некорректных задач, как правило, обеспечивают сходимость к классическому решению, если исходная задача была корректной.

В работе формулируется и анализируется проблема математической интерпретации данных, получаемых при проведении экспериментов на исследуемом объекте, которая не рассматривается в рамках классической математики и во многих случаях сводится к необходимости решения некорректно поставленных задач.

ПОСТАНОВКА ЗАДАЧИ

Задачи, которые необходимо решать при математической интерпретации данных, отличаются от рассматриваемых в классической математике тем, что их сложно классифицировать и установить условия, при которых они разрешимы. Поэтому такие задачи можно трактовать как некорректно поставленные в указанном выше смысле. При этом исходными данными для их решения являются результаты измерения или наблюдения за состоянием либо поведением не-

которого реального физического объекта, полученные в процессе экспериментальных исследований. Данные измерений могут быть косвенными, т.е. задаваться в виде соотношений или операторов, связывающих их с переменными состояния объекта. Такую связь запишем в виде уравнения наблюдения

$$y = h(w, u), \quad (1)$$

где y — вектор измеряемых переменных, имеющий конечную размерность M , w — состояние объекта (конечномерное или бесконечномерное), u — вектор входного воздействия, который в большинстве случаев также конечномерный размерности R .

Задача интерпретации данных возникает, когда из (1) по y и u невозможно найти w , т.е. система (1) незамкнута. Например, в простейшем линейном случае формула (1) с вектором внутреннего состояния x принимает вид

$$y = Cx + Du, \quad (2)$$

где в конечномерном случае C — матрица размера $M \times N$ (N — порядок системы), матрица D имеет размер $M \times R$, а для бесконечномерных систем C — линейный оператор, осуществляющий отображение линейного бесконечномерного векторного пространства в конечномерное. При этом система (2) недоопределена относительно x , т.е. измерения неполные. Это означает, что из (1) или (2) невозможно определить вектор состояния. Тогда необходима дополнительная информация об исследуемой системе. Обычно в таком случае используется математическая модель, описывающая процессы или состояние системы. Чаще всего она имеет вид локальных связей между ее основными параметрами. Для динамических систем с сосредоточенными параметрами такие связи записываются в виде обыкновенных дифференциальных (разностных) уравнений или систем уравнений. В системах с распределенными параметрами локальная модель задается в виде уравнений или систем уравнений в частных производных.

Функционально-аналитическая форма в виде операторных уравнений позволяет объединить оба случая и записать локальную модель как

$$L(D)w = f, \quad (3)$$

где D — оператор дифференцирования, w — скалярное или векторное поле, характеризующее внутреннее состояние системы, $w = x$ в сосредоточенных системах, f — функция, зависящая от u и определяющая классическое или обобщенное входное воздействие на систему.

В классической математике формула (3) дополняется начальным условием, которое в общем виде запишем как

$$L_0(D)w = w_0, \quad t = t_0. \quad (4)$$

В пространственно-распределенных системах $w = w(z, t)$, где переменная z принадлежит открытой области G евклидова пространства, включая границу ∂G , а временная переменная принимает значения на числовой прямой $t \geq t_0$ или конечном временном интервале S . Кроме того, для систем с распределенными параметрами записываются граничные или краевые условия, которые в операторном представлении имеют вид

$$l_\Gamma(D)w = f_\Gamma(z, t), \quad z \in \partial G, \quad t \geq t_0 \quad \text{или} \quad t \in S, \quad (5)$$

где l_Γ — линейный оператор, $f_\Gamma(z, t)$ — функция или вектор-функция, классическая либо обобщенная, которую можно считать заданной. Заметим, что для обыкновенных дифференциальных уравнений помимо задачи Коши рассматриваются также краевые задачи.

В математике кроме классических исследуются задачи, трактуемые как обратные. Имеется много публикаций, посвященных таким задачам (некоторые из них описаны в [1, 2]). Основное внимание в обратных задачах уделяется вопросам их разрешимости в рамках классического подхода. В задачах математической интерпретации, рассматриваемых в настоящей статье, акцент делается на том, как по не-

полным данным, получаемым при допустимых измерениях в виде (1), расширить и уточнить их с использованием локальных связей, представляемых уравнениями (3). Данные считаются заданными и не могут изменяться, поскольку именно их необходимо интерпретировать. Модель допускает различные приближенные представления и упрощения, позволяющие согласовывать ее с имеющимися данными так, чтобы задача интерпретации имела практически пригодное решение. Более того, адекватными могут считаться описания разной структуры или в рамках одной структуры за счет наличия в них неопределенных параметров, которые представляют множество моделей. В последнем случае задача интерпретации сводится к расширению знаний о переменных состоянии системы и определению априори неизвестных параметров модели. Кроме того, допускается использование незамкнутых моделей относительно всех неизвестных переменных. Эти факторы обуславливают трактование задачи интерпретации как некорректно поставленной, поскольку ее корректность является исключением, а не правилом. Обычно имеем недоопределенную или переопределенную ситуации. При наличии погрешностей в данных классические решения не имеют смысла, поэтому находится обобщенное приближенное решение с использованием процедур регуляризации, т.е. определяется регуляризованное решение.

ЛИНЕЙНАЯ СИСТЕМА С СОСРЕДОТОЧЕННЫМИ ПАРАМЕТРАМИ

Рассматривается простейший случай задачи интерпретации системы с сосредоточенными параметрами, ее уравнение наблюдения и математическая модель линейны. Вначале исследуем непрерывную систему с непрерывным и дискретным измерением, затем для дискретного случая опишем подход к решению задачи интерпретации.

Случай непрерывных данных. Рассмотрим простейшую задачу интерпретации с уравнением наблюдения (2) для линейной стационарной динамической системы с математической моделью вида

$$\frac{dx}{dt} = Ax + Bu, \quad (6)$$

где A и B — матрицы размера $N \times N$ и $N \times R$ соответственно.

Пусть измерения проводились на интервале $[0, T]$. Если воспользоваться формулой Коши, представляющей общее решение (6), то, следуя [3], можно получить с учетом (2) соотношения вход–выход. Для простоты полагаем $D \equiv 0$, в результате имеем

$$y_m(t) = \sum_{p=1}^P [(c_{mp}^c \cos \beta_p t + c_{mp}^s \sin \beta_p t) x_{p_0}^c + (c_{mp}^s \cos \beta_p t - c_{mp}^c \sin \beta_p t) x_{p_0}^s] e^{-\alpha_p t} + \sum_{r=1}^R \sum_{p=1}^P \int_0^t [f_{mp}^{rc} \cos \beta_p(t-\theta) + f_{mp}^{rs} \sin \beta_p(t-\theta)] e^{-\alpha_p(t-\theta)} u_r(\theta) d\theta, \quad m = \overline{1, M}, \quad (7)$$

где $\alpha_p \geq 0$, $\beta_p \geq 0$ (система предполагается устойчивой); $f_{mp}^{rc} = c_{mp}^c b_{rp}^c + c_{mp}^s b_{rp}^s$; $f_{mp}^{rs} = c_{mp}^s b_{rp}^c - c_{mp}^c b_{rp}^s$; $x_{p_0}^c$, $x_{p_0}^s$, c_{mp}^c , c_{mp}^s , b_{rp}^c , b_{rp}^s — начальное состояние системы, элементы матриц C и B соответствуют некоторой жордановой реализации [3]. Здесь для упрощения рассмотрен случай отсутствия кратных корней у матрицы A . В формуле (7) использовалась унифицированная форма записи, объединяющей действительные и комплексно-сопряженные собственные значения. Действительным корням соответствуют параметры $\beta_p = 0$, $c_{mp}^s = b_{rp}^s = x_{p_0}^s = 0$. Кроме того, в матрицах A, B, C следует удалить некоторые соответствующие этим параметрам столбцы. Тогда порядок системы N определяется усло-

вием $N = P_r + 2P_{\text{Im}}$, где P_r — число действительных, а P_{Im} — комплексно-сопряженных собственных значений. Параметры $f_{mp_0}^c, f_{mp_0}^s, f_{mp}^c, f_{mp}^s$ являются инвариантами системы относительно невырожденных преобразований с сохранением жордановой формы матрицы A .

Соотношения (7) позволяют по известным выходным переменным $y_m(t)$ и входным $u_r(t)$ определять начальное состояние $x_{p_0}^c, x_{p_0}^s$ ($p = \overline{1, P}$), а следовательно, и интегральную кривую, являющуюся решением задачи интерпретации. При этом имеем избыточные данные для решения такой задачи. Даже для системы с одним выходом достаточно использовать значения функции $y(t)$ в $P_r + 2P_{\text{Im}}$ произвольных различных точках, чтобы найти решение этой задачи при условии, что измерения и вычисления точные. Если измерения содержат погрешность, задача интерпретации с вероятностью почти единица считается некорректно поставленной в смысле существования классического решения уравнения (7). Можно найти только приближенное решение, которое называют также обобщенным. Его находят из решения следующей вариационной задачи:

$$\min_{\substack{x_{p_0}^c, x_{p_0}^s \\ p=1, P}} \sum_{m=1}^M \int_0^t \left(\sum_{p=1}^P [(c_{mp}^c \cos \beta_p t + c_{mp}^s \sin \beta_p t) x_{p_0}^c + (c_{mp}^s \cos \beta_p t - c_{mp}^c \sin \beta_p t) x_{p_0}^s] e^{-\alpha_p t} - \bar{y}_m(t) \right)^2 dt, \quad (8)$$

$$\text{где } \bar{y}_m(t) = y_m(t) - \sum_{r=1}^R \sum_{p=1}^P \int_{t_0}^t [f_{mp}^{rc} \cos \beta_p(t-\theta) + f_{mp}^{rs} \sin \beta_p(t-\theta)] e^{-\alpha_p(t-\theta)} u_r(\theta) d\theta$$

есть сигнал от свободного движения системы из начального состояния, в котором она находится в момент $t=0$. Необходимое условие экстремума (8) запишем в виде

$$\sum_{p=1}^P (a_{pq_1} x_{p_0}^c + a_{pq_2} x_{p_0}^s) = d_{q_1},$$

$$\sum_{p=1}^P (a_{pq_3} x_{p_0}^c + a_{pq_4} x_{p_0}^s) = d_{q_2}, \quad q = \overline{1, P}, \quad (9)$$

$$\text{где } a_{pq_1} = \sum_{m=1}^M \int_0^T (c_{mp}^c \cos \beta_p t + c_{mp}^s \sin \beta_p t) (c_{mq}^c \cos \beta_q t - c_{mq}^s \sin \beta_q t) e^{-(\alpha_p + \alpha_q)t} dt,$$

$$a_{pq_2} = \sum_{m=1}^M \int_0^T (c_{mp}^s \cos \beta_p t - c_{mp}^c \sin \beta_p t) (c_{mq}^c \cos \beta_q t + c_{mq}^s \sin \beta_q t) e^{-(\alpha_p + \alpha_q)t} dt,$$

$$d_{q_1} = \sum_{m=1}^M \int_0^T \bar{y}_m(t) (c_{mq}^c \cos \beta_q t + c_{mq}^s \sin \beta_q t) e^{-\alpha_q t} dt,$$

$$a_{pq_3} = \sum_{m=1}^M \int_0^T (c_{mp}^c \cos \beta_p t + c_{mp}^s \sin \beta_p t) (c_{mq}^s \cos \beta_q t + c_{mq}^c \sin \beta_q t) e^{-(\alpha_p + \alpha_q)t} dt,$$

$$a_{pq_4} = \sum_{m=1}^M \int_0^T (c_{mp}^s \cos \beta_p t - c_{mp}^c \sin \beta_p t) (c_{mq}^s \cos \beta_q t - c_{mq}^c \sin \beta_q t) e^{-(\alpha_p + \alpha_q)t} dt,$$

$$d_{q_2} = \sum_{m=1}^M \int_0^T \bar{y}_m(t) (c_{mq}^s \cos \beta_q t - c_{mq}^c \sin \beta_q t) e^{-\alpha_q t} dt.$$

В результате имеем $2P$ линейных уравнений для нахождения $x_{p_0}^c, x_{p_0}^s, p = \overline{1, P}$.

При наличии действительных собственных значений таких уравнений имеем $N = P_r + 2P_{im}$. Если $c_{mq}^c = c_{mq}^s = 0$ для всех $m = \overline{1, M}$, то сигнал от соответствующей моды разложения (7) не содержится в измерениях $y_m(t)$, т.е. она ненаблюдаема и из (9) не может быть найдено ее начальное состояние $x_{q_0}^c, x_{q_0}^s$. Поэтому задача интерпретации разрешима, если все ее моды наблюдаемы хотя бы на одном выходе. Наличие шумов измерений при определенных условиях приводит к большей чувствительности решений к погрешностям в данных. Применительно к (9) эти условия сводятся к следующему. Имеем большой разброс решений для различных реализаций ошибок измерений, если система плохо наблюдаема, что соответствует малым значениям c_{mq}^c, c_{mq}^s при любом m . В этом случае система (9) плохо обусловлена. Также она становится близкой к вырожденной, если имеются разные, но мало отличающиеся собственные значения. Особенно это характерно для очень больших значений N . Тогда собственные значения на комплексной плоскости представляются в виде кластеров с достаточно близкими корнями. В противном случае собственные значения распределяются на удалении одно от другого, что при обезразмеривании для больших N может привести к наличию разнотемповых процессов в системе. В результате соотношение (6) становится эквивалентным системе с малым параметром при некоторых производных. Тогда вблизи $t = 0$ появляются погранслойные решения, для которых на остальном интервале наблюдения отношение сигнал/шум становится малой величиной. Именно этот фактор обуславливает большую чувствительность решения к погрешностям.

Случай дискретных данных. Часто модель непрерывна, а данные измерений дискретны. Пусть имеем последовательность таких измерений $\{y(t_i)\} = \{y_i\}$, $i = \overline{1, I}$. Тогда вариационная задача сводится к обычному методу наименьших квадратов, результатом которого является система уравнений вида (9) для определения $x_{p_0}^c, x_{p_0}^s$, в которой соответствующие коэффициенты представлены соотношениями

$$a_{pq_1} = \sum_{m=1}^M \sum_{i=1}^I (c_{mp}^c \cos \beta_p t_i + c_{mp}^s \sin \beta_p t_i)(c_{mq}^c \cos \beta_q t_i - c_{mq}^s \sin \beta_q t_i) e^{-(\alpha_p + \alpha_q)t_i},$$

$$a_{pq_2} = \sum_{m=1}^M \sum_{i=1}^I (c_{mp}^s \cos \beta_p t_i - c_{mp}^c \sin \beta_p t_i)(c_{mq}^c \cos \beta_q t_i + c_{mq}^s \sin \beta_q t_i) e^{-(\alpha_p + \alpha_q)t_i},$$

$$d_{q_1} = \sum_{m=1}^M \sum_{i=1}^I \bar{y}_m(t_i)(c_{mq}^c \cos \beta_q t_i + c_{mq}^s \sin \beta_q t_i) e^{-\alpha_q t_i},$$

$$a_{pq_3} = \sum_{m=1}^M \sum_{i=1}^I (c_{mp}^c \cos \beta_p t_i + c_{mp}^s \sin \beta_p t_i)(c_{mq}^s \cos \beta_q t_i + c_{mq}^c \sin \beta_q t_i) e^{-(\alpha_p + \alpha_q)t_i},$$

$$a_{pq_4} = \sum_{m=1}^M \sum_{i=1}^I (c_{mp}^s \cos \beta_p t_i - c_{mp}^c \sin \beta_p t_i)(c_{mq}^s \cos \beta_q t_i - c_{mq}^c \sin \beta_q t_i) e^{-(\alpha_p + \alpha_q)t_i},$$

$$d_{q_2} = \sum_{m=1}^M \sum_{i=1}^I \bar{y}_m(t_i)(c_{mq}^s \cos \beta_q t_i - c_{mq}^c \sin \beta_q t_i) e^{-\alpha_q t_i}.$$

Условия, при которых непрерывная задача является некорректно поставленной, распространяются на случай с дискретными измерениями. Кроме того, при $M \cdot I < P$ система (9) становится вырожденной, т.е. данных недостаточно для решения задачи их интерпретации. Таким образом, в зависимости от свойств и особенностей матриц A, C линейной системы, а также числа приближенных данных, если они дискретны, рассматриваемая задача интерпретации может оказаться некорректно поставленной. В таких случаях необходимы процедуры регуляризации, позволяющие находить приближенные устойчивые к погрешностям решения. Если некорректность определяется структурными особенностями модели и уравнения наблюдения, одной из таких процедур может быть редукция порядка модели с построением более простого аппроксимирующего описания системы без этих особенностей, у которой при любых входных воздействиях точность выходных переменных согласована с погрешностями измерений. Способ построения аппроксимирующей модели более низкого порядка при знании исходного описания (2), (6) подробно изложен в работе [4]. Здесь рассмотрим только вопрос выбора порядка этой модели.

Пусть имеем непрерывные измерения вектора y . На интервале $[0, T]$ сформируем из него дискретную последовательность $\{y_i\}$ с некоторым достаточно малым шагом Δ . Выберем его так, чтобы для самого быстрого переходного процесса, определяемого $\max_p \{\alpha_p\}, \max_p \{\beta_p\}$, число дискретных измерений было

не менее десяти. Каждая m -я компонента y_{mi} содержит погрешность ξ_{mi} , при любых реализациях и произвольном i удовлетворяющую

$$|\xi_{mi}| \leq \varepsilon_m, \quad m = \overline{1, M}. \quad (10)$$

Из каждого y_{mi} вычитаем в соответствующих точках точное значение сигнала вынужденного движения. В результате получим последовательность $\{\bar{y}_i\}$, характеризующую свободное движение. Из нее формируем ганкелеву матрицу

$$Y = \begin{bmatrix} \bar{y}_1 & \bar{y}_2 & \dots & \bar{y}_{\bar{K}} \\ \bar{y}_2 & \bar{y}_3 & \dots & \bar{y}_{\bar{K}+1} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{y}_{N_1} & \bar{y}_{N_1+1} & \dots & \bar{y}_{\bar{K}+N_1-1} \end{bmatrix}, \quad (11)$$

где \bar{y}_i — вектор размерности M .

Значение N_1 должно быть равным или большим N исходной модели. Значение \bar{K} с помощью шага дискретизации выбирается таким, чтобы $\bar{K} > N$ и точка $\bar{K} + N_1 - 1$ была равной точке T или близкой к ней. По аналогии с (11) запишем матрицу погрешностей

$$\Xi = \begin{bmatrix} \xi_1 & \xi_2 & \dots & \xi_{\bar{K}} \\ \xi_2 & \xi_3 & \dots & \xi_{\bar{K}+1} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{N_1} & \xi_{N_1+1} & \dots & \xi_{\bar{K}+N_1-1} \end{bmatrix}. \quad (12)$$

Примем следующее соглашение. Пусть имеем матрицы A и B размера $m \times n$. Тогда полагаем

$$B = |A| \Rightarrow b_{ij} = |a_{ij}| \quad \forall i \in \overline{1, m}, \quad j \in \overline{1, n},$$

$$B \leq |A| \Rightarrow b_{ij} \leq a_{ij} \quad \forall i \in \overline{1, m}, \quad j \in \overline{1, n},$$

где $|A|$ — матрица абсолютных значений. Это позволяет с учетом (10) записать матричное неравенство в виде

$$\Xi \leq E, \quad (13)$$

где E — матрица, в которой элементы ξ_{mi} в матрице (12) заменили значениями ε_m из (10). Выполнив сингулярное разложение (SVD) матрицы Y , получим

$$Y = Q\Sigma V^T, \quad (14)$$

где Q и V — ортогональные матрицы, Σ — диагональная матрица с сингулярными числами на диагонали, расположенными в невозрастающем порядке. Выполним разбиение на блоки всех матриц, входящих в (14):

$$Q = [Q_1, Q_2], \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}, \quad V_1 = [V_1, V_2]. \quad (15)$$

Размерности блоков выбираются так, чтобы

$$Y = Y_1 + Y_2 = Q_1 \Sigma_1 V_1^T + Q_2 \Sigma_2 V_2^T. \quad (16)$$

Тогда матрица Y_1 имеет неполный ранг, определяющий размерность редуцированной модели [4]. Следует так выбрать размерность Σ_1 , чтобы аппроксимирующая модель пониженного порядка по точности согласовывалась с погрешностью измерений. Установить это можно по поведению сингулярных чисел матрицы Σ . Наиболее типичны две ситуации. В первой все сингулярные числа, по крайней мере в области малых значений, близких к ε_m , строго убывают, а во второй — начиная с некоторого малого числа практически не уменьшаются. В первом случае местоположение первого сингулярного числа, по величине ставшего меньше или равным $\max_m(\varepsilon_m)$, определяет нулевое при-

ближение размерности аппроксимирующей модели, т.е. размерность первых блоков в (15). Для этой размерности проверяется выполнимость условия

$$|Y_2| = |Q_2 \Sigma_2 V_2^T| \leq \Xi. \quad (17)$$

Если оно выполняется строго или не выполняется, то уменьшаем или увеличиваем на единицу размерность первого блока и снова проверяем выполнимость (17). Искомой размерностью будет такая, для которой условие (17) выполняется, а при уменьшении ее на единицу неравенство меняет знак на противоположный. Во втором случае порядок редуцированной модели определяется сингулярным числом, после которого все последующие числа практически не уменьшаются.

Параметры модели такого редуцированного порядка определяются одним из методов, описанных в [4].

Упрощенный способ построения редуцированной модели. В определенных случаях для построения аппроксимирующей модели пониженного порядка можно применить упрощенный способ. Его суть в том, что размерность и собственные значения редуцированной модели находятся эмпирически, а остальные параметры — из квадратичной вариационной задачи, которая сводится к решению линейных уравнений. Опишем здесь первую часть задачи, решение второй подробно изложено в [4].

Для построения редуцированной модели полагаем начальные условия нулевыми. Тогда понижение порядка осуществляется отдельно по каждой импульсной переходной функции, являющейся элементом переходной матрицы, т.е. для скалярного случая, вытекающего из (7), имеем

$$y_m(t) = \int_0^t h_{mr}(t-\theta) u_r(\theta) d\theta, \quad (18)$$

где $h_{mr}(\tau) = \sum_{p=1}^{P_{mr}} [f_{mp}^{rc} \cos \beta_p \tau + f_{mp}^{rs} \sin \beta_p \tau] e^{-\alpha_p \tau}$ и $P_{mr} \leq P$.

Импульсная переходная функция редуцированной модели \hat{h}_{mr} имеет пони-

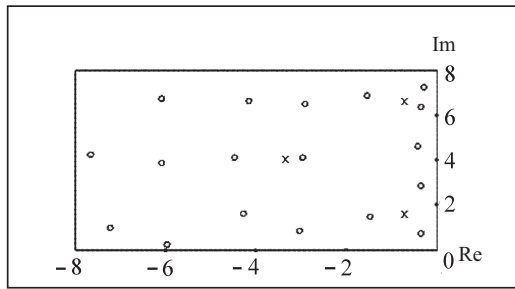


Рис. 1

ты, при больших P (более 20) качество аппроксимации улучшается до размерностей $4 \div 6$, после чего дальнейшее ее увеличение вплоть до истинной размерности слабо влияет на взаимное рассогласование импульсных функций. Согласно этому выбираются собственные значения исходной системы с наибольшими значениями параметров f_{mp}^{rc} и f_{mp}^{rs} , которые считаем центрами притяжения. Из близко расположенных к ним других собственных значений формируем кластеры области притяжения. При этом некоторые собственные значения с очень малыми f_{mp}^{rc} , f_{mp}^{rs} или большими α_p , β_p могут исключаться, т.е. не входить в кластеры. Для каждого кластера находим центры тяжести, считая, что в каждой точке с координатами (α_p, β_p) сосредоточен вес $\sqrt{(f_{mp}^{rc})^2 + (f_{mp}^{rs})^2}$. Координаты центра тяжести и будут собственными значениями аппроксимирующей системы. Зная эти значения, достаточно просто на основе результатов [4] вычисляем оценки \hat{f}_{mp}^{rc} и \hat{f}_{mp}^{rs} . По данным оценкам можно установить, насколько корректно выбраны центры притяжения. Если среди \hat{f}_{mp}^{rc} и \hat{f}_{mp}^{rs} имеются малые значения, то можно скорректировать соответствующие центры тяжести, изменив конфигурацию соответствующих кластеров. На рис. 1 для исходной (кружки) и аппроксимирующей (крестики) моделей представлены собственные значения, расположенные в верхней части устойчивой области комплексной плоскости (нижняя часть симметрична).

Описанную процедуру выбора размерности и собственных значений аппроксимирующей модели можно использовать как начальное приближение при нахождении итеративным способом минимизационной задачи, к которой сводится редукция модели [4].

НАБЛЮДАТЕЛЬ ЛУЕНБЕРГЕРА И ФИЛЬТР КАЛМАНА

Оригинальный способ решения рассматриваемой задачи был предложен Луенбергером [5], который имплантировал уравнение наблюдения в уравнение динамики:

$$\frac{dx}{dt} = Ax + Bu + L(y - Cx), \quad (19)$$

где L — матрица $N \times N$, позволяющая для наблюдаемой системы формировать любые собственные значения матрицы $A - LC$. Выбрав их так, чтобы все собственные значения обеспечивали устойчивость (19) с быстрым затуханием переходного процесса, получим асимптотическое приближение к искомому решению, например, при нулевом начальном значении $x_0 = x(0)$. На большей части интервала $[0, T]$ получим близкую к точной оценку, кроме окрестности точки 0, где имеем погранслоное решение. Однако при приближенном $y(t)$ такое сингулярно выраженное решение чувствительно к погрешностям, т.е. невозможно неограниченно уменьшать толщину погранслоя — должно быть согласование его толщины с уровнем шумов измерений. В связи с этим Кал-

женную размерность $\hat{P}_{mr} \leq P_{mr}$ и параметры $\hat{\alpha}_p$, $\hat{\beta}_p$, \hat{f}_{mp}^{rc} , \hat{f}_{mp}^{rs} . Приведем исходную систему к безразмерному виду, выбрав в качестве характерного времени длительность интервала наблюдения T . Тогда $t \in [0, 1]$, а параметры α_p , β_p преобразуются в безразмерные. Вычислим их и разместим на комплексной плоскости. Как показывают вычислительные эксперименты,

ман разработал фильтр, в котором матрица L однозначно определяется характеристиками помех, включая возмущения на входе, обеспечивающий минимум среднеквадратической ошибки оценки состояния. Как показали дальнейшие исследования, фильтр Калмана не обладает робастностью по отношению к вариациям параметров и внешних условий. Особенно это проявляется при плохой наблюдаемости и больших размерностях уравнений динамики. В таких случаях его эффективность существенно снижается, вплоть до получения неустойчивых решений. Другими словами, задача интерпретации с использованием фильтра Калмана также может считаться некорректно поставленной. В отличие от ранее рассмотренных вариантов установить непосредственно по уравнению (19), в каком случае задача становится некорректной, пока не удалось. В этом направлении необходимы дополнительные исследования. Полагаем, что условия некорректности остаются такими же, как в предыдущих разделах, а именно зависящими от расположения собственных значений на комплексной плоскости, большой размерности и плохой наблюдаемости.

ВАРИАЦИОННЫЙ СПОСОБ РЕШЕНИЯ

Такой способ вполне естествен, если задача интерпретации является переопределенной по количеству исходных данных. Тогда в качестве ее решения принимается элемент $x(t)$, минимизирующий аддитивный функционал, составленный из невязок уравнений, описывающих динамику системы и уравнений наблюдения, т.е. из решения вариационной задачи, формулируемой с помощью норм соответствующих функциональных пространств. Запишем ее для общего случая, когда уравнения наблюдения (1) и уравнения динамики, которые имеют вид

$$\frac{dx}{dt} = f(t, x, u), \quad (20)$$

нелинейны. В такой ситуации наиболее употребительны пространства с нормой L_2 . Тогда вариационную задачу запишем как

$$J[x(\cdot)] = \int_0^T \left[\left\langle \frac{dx}{dt} - f, \Omega_1 \left(\frac{dx}{dt} - f \right) \right\rangle + \left\langle \tilde{y} - h, \Omega_2 (\tilde{y} - h) \right\rangle \right] dt \rightarrow \min, \quad (21)$$

где \tilde{y} — выходные сигналы, измеряемые с погрешностью, а Ω_1 и Ω_2 — положительные диагональные матрицы.

Фактически функционал $J[x(\cdot)]$ является взвешенной аддитивной суммой двух типов невязок: невыполнимость дифференциальных уравнений динамики и функциональных уравнений наблюдения. Матрицы Ω_1 и Ω_2 определяют в (21) вес тех или иных невязок, а значения их элементов должны согласовываться со значениями погрешностей измерения и возмущениями в уравнении динамики, если таковые имеются. Если задача оценивания $x(t)$ для каждого t разрешима и при этом решение единственное, то в случае отсутствия погрешностей в исходных данных задача минимизации (21) при любых Ω_1 и Ω_2 имеет абсолютный минимум на этом решении. При наличии в данных погрешностей можно с помощью весовых матриц Ω_1 и Ω_2 регулировать невязки в получаемых решениях, добиваясь для них выполнения условий согласования с погрешностью.

Существование минимизирующего элемента экстремальной задачи (21) в пространстве Соболева для ограниченного снизу квадратичного функционала с непрерывной частной производной по $\frac{dx}{dt}$ и непрерывными по своим аргументам функциями f и h следует из теоремы Тонелли [6]. Более того, задача (21) есть не что иное, как задача Больца с нулевым терминантом, т.е. задача классического вариационного исчисления (без ограничений), для которой можно записать необходимые условия экстремума в виде уравнений Эйлера и условия трансверсальности.

Для функционала (21) запишем их в виде

$$\begin{aligned} \frac{d^2x}{dt^2} + \left(\Omega_1^{-1} \left(\frac{\partial f}{\partial x} \right)^T \Omega_1 - \frac{\partial f}{\partial x} \right) \frac{dx}{dt} - \Omega_1^{-1} \left(\frac{\partial f}{\partial x} \right)^T \Omega_1 f - \\ - \frac{\partial f}{\partial t} - \frac{\partial f}{\partial u} \frac{du}{dt} + \Omega_1^{-1} \left(\frac{\partial h}{\partial x} \right)^T \Omega_2 (\tilde{y} - h) = 0 \end{aligned} \quad (22)$$

с условием трансверсальности

$$\frac{dx}{dt} - f(t, x, u) = 0, \quad t = 0, \quad t = T. \quad (23)$$

В результате получаем нелинейную двухточечную краевую задачу. Однозначная ее разрешимость дает единственный минимизирующий элемент задачи (21), который примем за оценку вектора состояния на интервале наблюдения. Для общего случая систем нелинейных краевых задач вида (22), (23) доказательство единственности достаточно сложно. Полагаем, что условия единственности решения этой задачи вытекают из локальных условий полной наблюдаемости и $(n-1)$ -кратной дифференцируемости вектора наблюдений $\tilde{y}(t)$. Согласно [7] вектор $\hat{x}(t)$ однозначно определяется, если матрица Якоби на всех финитных траекториях интервала наблюдения, реализуемых допустимыми входными воздействиями, полноранговая. При этом матрица Якоби формируется из функций h и оператора Ляпунова задачи интерпретации.

Данная задача относится к классу обратных, для которых обратный оператор, записанный в виде (22), (23), может не быть вполне непрерывным. Тогда необходимо исследовать устойчивость получаемого по приближенным данным решения, т.е. его чувствительность к имеющимся погрешностям в исходных данных. Речь идет о практической пригодности получаемого решения, или, другими словами, его робастности. Это можно установить, оценивая свойство непрерывности обратного оператора (определяемого двухточечной краевой задачей) от исходных данных $y(t)$ и $u(t)$. В общем случае осуществить это сложно, так же, как доказать однозначную разрешимость задачи. Однако после применения конечно-разностной аппроксимации и итеративной процедуры решения задачи можно сформулировать конструктивные критерии для оценки устойчивости получаемых решений.

Итеративный метод решения. Пусть функции $f(\cdot)$ и $h(\cdot)$ в (1) таковы, что допускается их представление в виде

$$\begin{aligned} f(t, x, u) &= Ax + Bu + f^{NL}(t, x, u), \\ h(t, x) &= Hx + h^{NL}(t, x), \end{aligned} \quad (24)$$

где A, B, H — постоянные матрицы соответствующих размеров.

Разложение (24) допустимо, если, например, система (1) автономна и выделяемая линейная часть аппроксимирует ее движение вблизи некоторого равновесного состояния (x_e, u_e) . Тогда задачу (22), (23) можно преобразовать в систему дифференциальных уравнений

$$\frac{dx}{dt} = z,$$

$$\frac{dz}{dt} = A_1x + A_2z + h^0 + h^{NL} \quad (25)$$

с краевыми условиями

$$z - Ax = Bu + f^{NL}, \quad t = 0, \quad t = T. \quad (26)$$

В (25) приняты следующие обозначения:

$$A_1 = \Omega_1^{-1}(A^T \Omega_1 A - H^T \Omega_2 H), \quad A_2 = A - \Omega_1^{-1} A^T \Omega_1,$$

$$h^0 = \Omega_1^{-1}(A^T \Omega_1 B u - H^T \Omega_2 \tilde{y}) + B u,$$

$$h^{NL} = \Omega_1^{-1} \left[A^T \Omega_1 f^{NL} + \frac{\partial f^{NL}}{\partial x} \Omega_1 (A x + B u + f^{NL}) - \frac{\partial (h^{NL})^T}{\partial x} \Omega_2 (\tilde{y} - H x - h^{NL}) \right] -$$

$$- \left(\Omega_1^{-1} \frac{\partial f^{NL}}{\partial x} \Omega_1 - \frac{\partial f^{NL}}{\partial x} \right) z + \frac{\partial f^{NL}}{\partial t} + \frac{\partial f^{NL}}{\partial u} \frac{du}{dt}.$$

Аппроксимирующая (25), (26) простейшая конечно-разностная модель записывается как разностные уравнения

$$x_{k+1} = x_k + \Delta z_k, \quad k \in \overline{0, N},$$

$$z_{k+1} = \Delta A_1 x_k + (E + \Delta A_2) z_k + h_k^0 + h_k^{NL} \quad (27)$$

с краевыми условиями

$$z_0 - A x_0 = B u_0 + f_0^{NL},$$

$$z_N - A x_N = B u_N + f_N^{NL}, \quad (28)$$

где Δ — шаг квантования по времени, k — дискретное время, $T = \Delta N$, h_k^0 , h_k^{NL} , u_k , f_k^{NL} — значения соответствующих функций в дискретных точках.

Полученная задача (27), (28) нелинейна. Для ее решения применим итерационную схему, суть которой в следующем. Действуя формально, с помощью рекуррентных формул запишем уравнения, связывающие начальное состояние (x_0, z_0) со значением переменных в момент времени $k+1$:

$$x_{k+1} = P_k^1 x_0 + Q_k^1 z_0 + S_k^1,$$

$$z_{k+1} = P_k^2 x_0 + Q_k^2 z_0 + S_k^2, \quad k \in \overline{0, N-1}, \quad (29)$$

где матрицы P_k^1 , P_k^2 , Q_k^1 , Q_k^2 и векторы S_k^1 , S_k^2 определяются рекуррентными соотношениями

$$P_{k+1}^1 = P_k^1 + \Delta P_k^2,$$

$$P_{k+1}^2 = \Delta A_1 P_k^2 + (E + \Delta A_2) P_k^2,$$

$$Q_{k+1}^1 = Q_k^1 + \Delta Q_k^2,$$

$$Q_{k+1}^2 = \Delta A_1 Q_k^2 + (E + \Delta A_2) Q_k^2,$$

$$S_{k+1}^1 = S_k^1 + \Delta S_k^2,$$

$$S_{k+1}^2 = \Delta A_1 S_k^2 + (E + \Delta A_2) S_k^2 + h_{k+1}^0 + h_{k+1}^{NL}, \quad k \in \overline{0, N-1}.$$

При этом имеем

$$P_0^1 = E, \quad Q_0^1 = \Delta E, \quad S_0^1 = 0,$$

$$P_0^2 = \Delta A_1, \quad Q_0^2 = E + \Delta A_2, \quad S_0^2 = h_0^0 + h_0^{NL}.$$

Фактически формула (29) — дискретная аппроксимация нелинейных интегральных уравнений, которые можно получить из (25) на основе формулы Коши с ядром, определяемым линейной частью в (25).

Далее, действуя в рамках того же формализма, краевые условия (28) можно привести к виду

$$z_0 - Ax_0 = Bu_0 + f_0^{NL},$$

$$(Q_{N-1}^2 - AQ_{N-1}^1)z_0 + (P_{N-1}^2 - AP_{N-1}^1)x_0 = Bu_N + AS_{N-1}^1 + f_N^{NL}. \quad (30)$$

Левая часть (30) состоит из матриц с постоянными коэффициентами. Нелинейности вошли в правую часть через функции f^{NL} и h^{NL} . Если их положить равными нулю, то (30) преобразуется в систему уравнений для оценивания неизвестного начального состояния (x_0, z_0) по измеряемым (в большинстве случаев приближенно) входным и выходным переменным в линейных стационарных системах, т.е. двухточечная краевая задача сводится к системе линейных алгебраических уравнений. При приближенных исходных данных решение устойчиво, если квадратная матрица в левой части (30) не является плохо обусловленной. Более того, условие наблюдаемости в данном случае эквивалентно невырожденности этой матрицы (минимизирующий элемент единственный), а ее обусловленность определяет чувствительность решения к погрешностям в исходных данных. Число обусловленности данной матрицы возрастает с увеличением длины интервала T , причем независимо от того, является система устойчивой, нейтральной или неустойчивой. Поэтому имеет значение вопрос выбора подходящего T . Как показали исследования, значения матриц Ω_1 и Ω_2 слабо влияют на обусловленность основной матрицы, а точность получаемых оценок зависит от выбора элементов этих матриц. Чем больше диагональные элементы матрицы Ω_2 в сравнении с Ω_1 , тем меньше невязка в уравнениях наблюдения, и, наоборот, чем меньше Ω_2 по отношению к Ω_1 , тем точнее выполняется уравнение движения. Следовательно, выбирать эти матрицы необходимо так, чтобы значения их элементов были согласованы с невязками, определяемыми погрешностями. Здесь прослеживается аналогия с выбором оптимальных параметров фильтра Калмана по известным ковариационным матрицам шумов и возмущений.

Описанные свойства позволяют реализовать следующий итерационный метод решения задачи интерпретации в корректном нелинейном случае на основе (30).

Итерация 0. Полагаем в (29), (30) $f^{NL} = g^{NL} = 0$ и решаем линейную задачу оценивания. Определяем во всех точках интервала $[0, T]$ функции x_k и z_k .

Итерация 1. Полученные на 0-й итерации оценки вместе с временными функциями входного и выходного сигналов подставляем в f^{NL} и g^{NL} . В результате находим S_k^1 и S_k^2 по известной временной функции g_k^{NL} , а также значения f_0^{NL} и f_N^{NL} . Решаем ту же линейную задачу с уточненной правой частью, т.е. при найденных f_0^{NL} , S_{N-1}^1 , f_N^{NL} как функциях времени (k). Получаем новую уточненную оценку x_0, z_0 .

Итерация 2. По уточненным x_0, z_0 находим x_k и z_k из (28). По аналогии с предыдущей итерацией определяем f^{NL} и g^{NL} как функции времени (k). Находим решение линейной задачи (30) с правой частью, определяемой уточненными f^{NL} и g^{NL} .

Все последующие итерации реализуются аналогично. Итерации прекращаются, как только все невязки станут равными или меньшими погрешности. При правильном выборе матриц Ω_1 и Ω_2 это условие должно выполняться одновременно по всем уравнениям невязок.

Замечание 1. Описанные итерации сходятся, если N такое, что основная матрица системы (30) невырожденная (выполняются условия наблюдаемости) и не является плохо обусловленной (выполняются условия устойчивости), т.е. T не должно быть очень малым или очень большим.

Вычислительные эксперименты показывают, что для наблюдаемой системы с хорошей обусловленностью матрицы в (30) нестрогие уравнения невязок выполняются за небольшое число итераций.

В случае, если задача (21) некорректно поставлена или предполагаем, что она такова, следует воспользоваться одним из методов их решения, описанных в [8]. В рассмотрение вводится стабилизатор $\Lambda[x]$, который в данном случае запишем как

$$\Lambda[x] = \int_0^T \langle x, \Omega_3 x \rangle dt,$$

где Ω_3 — положительно-определенная квадратная матрица, например единичная.

Определяем Λ -нормальное решение задачи (21), которое на основе лемм о регуляризации сводится к построению одним из методов, описанных в [8], регулярных минимизирующих последовательностей, сходящихся к этому решению. При неточных исходных данных строится приближенное устойчивое решение, согласованное по точности с погрешностями в ситуации, когда задача (21) поставлена некорректно. В общем случае не гарантируется единственности решения, что проветряется варьированием начальной точки минимизирующей последовательности.

Вычислительные эксперименты. Эффективность метода проверялась численным моделированием. Рассматривались линейный и нелинейный случаи. В линейном случае решалась и исследовалась модельная задача, в нелинейном — задача оценивания полного вектора состояния по измеряемым угловым координатам (угловые скорости не измерялись) вращательного движения космического аппарата вокруг центра масс. При корректной постановке и точных данных в обоих случаях получено точное решение для произвольных положительно-определенных диагональных матриц Ω_1 и Ω_2 . В линейном случае собственные значения приводились к безразмерному виду с характерным временем T . При варьировании T расположение собственных значений изменялось так, что при малом и большом T решение становилось некорректным, в результате чего возрастала его чувствительность к погрешностям. В области, в которой задача оставалась корректной, было получено решение, согласованное по точности с величиной ε , ограничивающей в каждой точке наблюдения величину погрешности, что установлено методом Монте-Карло. В рассматриваемой нелинейной задаче ориентации космического аппарата описанная итерационная схема работала эффективно. Методом Монте-Карло показана робастность метода по отношению к вариациям погрешности в пределах, допустимых ограничением ε . Для конкретной заданной системы существует интервал значений T , при которых решение было наилучшим. При малых ε согласованное по точности решение требовало больше итераций. Более того, существует такое ε , начиная с которого согласованное по точности с погрешностью решение получалось при нулевой итерации, т.е. приближенная линейная модель давала приемлемый результат. Более подробно результаты численного моделирования представлены в работах [9, 10].

ЗАКЛЮЧЕНИЕ

В данной работе задачи интерпретации данных с использованием математических моделей, адекватных рассматриваемым процессам, трактуются как некорректно поставленные, поскольку выполнить условия, при которых они имеют классические решения, практически невозможно. Как правило, имеется либо недостаточное, либо избыточное число данных, что при наличии в них погрешностей ставит под вопрос разрешимость задачи их интерпретации. Более того, при определенных реальных условиях проявляется чувствительность искомых решений к имеющимся погрешностям, что затрудняет их практическое использование. Применительно к системам с сосредоточенными параметрами описаны возможные подходы, позволяющие давать близкую к реальной интерпретацию экспериментальных данных, т.е. представление в целом об исследуемых процессах по неполным данным. Для заданных данных допускаются приближенные описания

путем математического упрощения исходной модели. Во многих случаях такой подход соответствует процедуре регуляризации некорректно поставленной задачи интерпретации, преобразуя ее в приближенно разрешимую.

СПИСОК ЛИТЕРАТУРЫ

1. Prilepko A.I., Orlovsky D.G., Vasin I.A. Methods for solving inverse problems in mathematical physics. N.Y.; Basel: Marcel Dekker, 1999. 357 p.
2. Кабанихин С.И. Обратные и некорректные задачи. Новосибирск: Сибир. науч. изд-во, 2009. 457 с.
3. Губарев В.Ф. Метод итеративной идентификации многомерных систем по неточным данным. Часть 1. Теоретические основы. *Проблемы управления и информатики*. 2006. № 5. С. 16–32.
4. Губарев В.Ф., Фатенко В.В. Геометрический и вариационные методы редукции порядка модели. *Проблемы управления и информатики*. 2018. № 1. С. 38–52.
5. Андреев Ю.Н. Управление конечномерными линейными объектами. Москва: Наука, 1976. 424 с.
6. Буттацо Дж., Джаквинта М., Гильдебрант С. Одномерные вариационные задачи. Новосибирск: Науч. книга, 2002. 248 с.
7. Справочник по теории автоматического управления (под ред. Красовского А.А.). Москва: Наука, 1987. 712 с.
8. Васильев Ф.П. Методы решения экстремальных задач. Москва: Наука, 1981. 400 с.
9. Gubarev V.F., Shevchenko V.M., Zhykov A.O., Gummel A.V. State estimation for systems subjected bounded uncertainty using moving horizon approach. *Preprints of the 15th IFAC Symposium on System Identification*. Saint-Malo, France, 2009. P. 910–915.
10. Губарев В.Ф., Дарьин А.Н., Лысюченко И.А. Нелинейный оценщик состояния по данным на скользящем интервале и возможность его применения в задаче ориентации космического аппарата. *Проблемы управления и информатики*. 2011. № 1. С. 118–132.

Надійшла до редакції 25.04.2018

В.Ф. Губарев

ПРОБЛЕМА МАТЕМАТИЧНОЇ ІНТЕРПРЕТАЦІЇ ДАНИХ.

I. СИСТЕМИ ІЗ ЗОСЕРЕДЖЕНИМИ ПАРАМЕТРАМИ

Анотація. Розглянуто проблему інтерпретації даних, отриманих в експериментальних дослідженнях, як неklasичну математичну задачу, що у більшості випадків є в широкому сенсі некоректно поставленою. Використано додаткову інформацію про об'єкт у вигляді рівнянь локальних зв'язків, які визначають його замкнуту або незамкнуту математичну модель. Описано процедури регуляризації, що дозволяють знаходити практично придатні розв'язки, узгоджені з наявними даними.

Ключові слова: експериментальні дані, математична модель, інтерпретація, некоректність, регуляризація, редукція моделі, узагальнений розв'язок, варіаційний метод.

V.F. Gubarev

PROBLEM OF MATHEMATICAL DATA INTERPRETATION.

I. SYSTEMS WITH LUMPED PARAMETERS

Abstract. Interpretation problem for data obtained in experimental research is considered as a nonclassical mathematical problem, which generally is ill-posed in many cases. Additional information in the form of equations of local constraints that define its closed or open mathematical model are used for this. Regularization procedures are described, which make possible to find applicable solutions consistent with available data.

Keywords: experimental data, mathematical model, interpretation, incorrectness, regularization, model order reduction, general solution, variational method.

Губарев Вячеслав Федорович,

чл.-кор. НАН України, доктор техн. наук, заведуючий відделом Інститута космічних досліджень НАН України і ГКА України, Київ, e-mail: v.f.gubarev@gmail.com.