

01;12

Использование принципа минимума информации для определения параметров модели по данным "эксперимента"

© А.В. Чижов

Физико-технический институт им. А.Ф. Иоффе РАН,
194021 Санкт-Петербург, Россия

(Поступило в Редакцию 12 ноября 1998 г.)

Описан метод восстановления параметров модели физического процесса из статистических данных эксперимента на основе принципа минимума информации. Исследованы возможности метода на примерах регулярного и хаотического процессов с использованием данных численных экспериментов с уравнением переноса и диффузии и с логистическим отображением.

Введение

При сопоставлении математических моделей и экспериментов иногда приходится решать обратную задачу определения неизвестных коэффициентов в уравнениях модели по данным измерений. При этом экспериментальные данные являются статистическими и целью становится построение наиболее правдоподобной модели явления. Такого рода задачами являются задачи спектроскопии, геологоразведки, определение коэффициентов переноса в уравнениях гидродинамики, определение турбулентной вязкости, согласование моделей биологических систем. Примеры такого рода обратных задач и методы их решения описаны в обзоре [1]. Данная работа демонстрирует принцип восстановления неизвестных параметров уравнения в частных производных, описывающих процессы переноса и диффузии, а также параметров логистического отображения. Идея восстановления состоит в использовании принципа минимума информации, подробно обсуждавшегося Г. Хакеном в работе [2], в частности, для решения задач поиска образца. В рассматриваемом случае искомым образом является уравнение, описывающее наблюдаемый процесс.

1. Экспериментальные данные

Будем рассматривать нестационарный процесс в предположении, что он описывается уравнением

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} - \eta \frac{\partial^2 u}{\partial x^2} = s \quad (1)$$

с неизвестными постоянными коэффициентами a, η, s .

Предполагается также, что в эксперименте мы можем измерять значения величины u в трех близких точках области Ω . А именно каждое измерение, обозначенное индексом n , подразумевает определение величин $u_0^n = u(t_n, x) + \xi_1^n$, $u_-^n = u(t_n, x - \Delta x) + \xi_2^n$, $u_+^n = u(t_n, x + \Delta x) + \xi_3^n$ в момент t_n и повторно в момент $t_n + \tau$ в центральной точке $\tilde{u}^n = u(t_n + \tau, x) + \xi_4^n$ со случайными ошибками ξ_i^n ($i = 1, 4$), причем $|\xi_i^n| < \varepsilon$. Значения $\Delta x, \tau$

заданы. Точки на плоскости (x, t) , в которых проводятся измерения, составляют четырехточечный шаблон, показанный на рис. 1.

Далее, вводятся вспомогательные величины

$$q_1^n = \tilde{u}^n - u_0^n, \quad q_2^n = u_-^n - u_0^n, \quad q_3^n = u_+^n - u_0^n. \quad (2)$$

При обработке данных серии из N измерений вычисляются статистические моменты, которые составляют вектор \bar{q} и матрицу Q

$$\bar{q} = \begin{pmatrix} \overline{q_1} \\ \overline{q_2} \\ \overline{q_3} \end{pmatrix}, \quad Q = \begin{pmatrix} \overline{q_1^2} & \overline{q_1 q_2} & \overline{q_1 q_3} \\ \overline{q_1 q_2} & \overline{q_2^2} & \overline{q_2 q_3} \\ \overline{q_1 q_3} & \overline{q_2 q_3} & \overline{q_3^2} \end{pmatrix}, \quad (3)$$

где надчеркивание обозначает осреднение по серии экспериментов

$$\bar{f}_j = \frac{1}{N} \sum_{n=1}^N f_j^n.$$

Эти данные являются исходными для задачи определения коэффициентов a, η, s .

2. Функция распределения

Основной этап решения задачи состоит в восстановлении совместной функции распределения вероятности величин q_1, q_2, q_3 , т.е. $P(q_1, q_2, q_3)$ такой, что

$$\iiint f_j(q_1, q_2, q_3) P(q_1, q_2, q_3) dq_1 dq_2 dq_3 = \bar{f}_j; \quad (4)$$

$$j = 1, 9,$$

где f_j — величины, средние значения которых \bar{f}_j составляют вектор \bar{q} и матрицу Q .

Условию (4) удовлетворяет множество функций $P(\mathbf{q})$. Чтобы при определении вероятности $P(\mathbf{q})$ внести минимум информации кроме той, которая содержится в соотношениях (4), мы из всех таких функций выберем

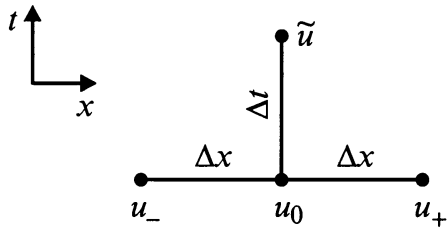


Рис. 1.

ту, которая содержит минимум информации о q , т.е. минимизирует функционал

$$\int P(\mathbf{q}) \ln P(\mathbf{q}) d\mathbf{q} = \min. \quad (5)$$

Можно показать [1], что такая плотность распределения имеет вид

$$P(\mathbf{q}) = c \exp \left\{ -\frac{1}{2} (\mathbf{q} - \bar{\mathbf{q}})^T G (\mathbf{q} - \bar{\mathbf{q}}) \right\} = c \exp \{V(\mathbf{q})\}, \quad (6)$$

где c — постоянный коэффициент, $G = (Q - \bar{\mathbf{q}} \bar{\mathbf{q}}^T)^{-1}$.

3. Восстановление численной схемы

Следующим этапом является составление численной схемы с заданным шаблоном (рис. 1), которая описывает процесс $u(t, x)$. Для этого мы должны определить наиболее вероятное значение величины q_1 , отвечающей за изменение поля $u(t, x)$ при переходе на новый временной слой, при заданных значениях q_2, q_3 на старом временном слое, т.е. мы должны потребовать, чтобы $\partial P / \partial q_1 = 0$. Подставляя выражение (6), получаем

$$\begin{aligned} \frac{\partial P}{\partial q_1} = & (G_{11}(q_1 - \bar{q}_1) + G_{12}(q_2 - \bar{q}_2) \\ & + G_{13}(q_3 - \bar{q}_3)) c \exp\{V\} = 0, \end{aligned}$$

т.е. $q_1 = a_1 q_2 + a_2 q_3 + a_3$, где $a_1 = -G_{12}/G_{11}$, $a_2 = -G_{13}/G_{11}$, $a_3 = \bar{q}_1 - a_1 \bar{q}_2 - a_2 \bar{q}_3$, или

$$\tilde{u} - u_0 = a_1(u_- - u_0) + a_2(u_+ - u_0) + a_3. \quad (7)$$

Подставляя разложение Тейлора для функции u около центрального узла шаблона, получаем

$$\begin{aligned} \frac{\partial u}{\partial t} + (a_1 - a_2) \frac{\Delta x}{\tau} \frac{\partial u}{\partial x} \\ = \frac{\Delta x^2}{2\tau} (a_1 + a_2) \frac{\partial^2 u}{\partial x^2} + \frac{a_3}{\tau} + O(\Delta x^3, \tau). \quad (8) \end{aligned}$$

Поскольку уравнение (1) является нулевым дифференциальным приближением для схемы (7), то, сравнивая (8) и (1), мы можем определить

$$a = \frac{\Delta x}{\tau} (a_1 - a_2), \quad \eta = \frac{\Delta x^2}{2\tau} (a_1 + a_2), \quad s = \frac{a_3}{\tau}.$$

Заметим, что к такому же результату можно прийти аналогично тому, как это делается в гл. 9 монографии [2]. Для этого в предположении о том, что процесс является марковским, определяется условная вероятность $P(q_1|q_2, q_3)$. Затем вычисляются дрейфовые коэффициенты для уравнения Фоккера–Планка, соответствующего условной вероятности. Эти коэффициенты определяют уравнение Ланжевена и, следовательно, искомые параметры, a, η, s .

4. Примеры

1. Восстановление параметров a, η, s проводилось для процесса, описываемого уравнением (1) с точными значениями коэффициентов $a = 1.3$, $\eta = 0.4$, $s = 2$, с нулевым начальным условием внутри области Ω и с граничными условиями

$$u|_{x=0} = \sin 5t + t, \quad u|_{x=1} = 1.$$

На шаблоне с $\Delta x = 0.033$, $\tau = 0.001$ проводилось $n = 400$ измерений со случайными ошибками,

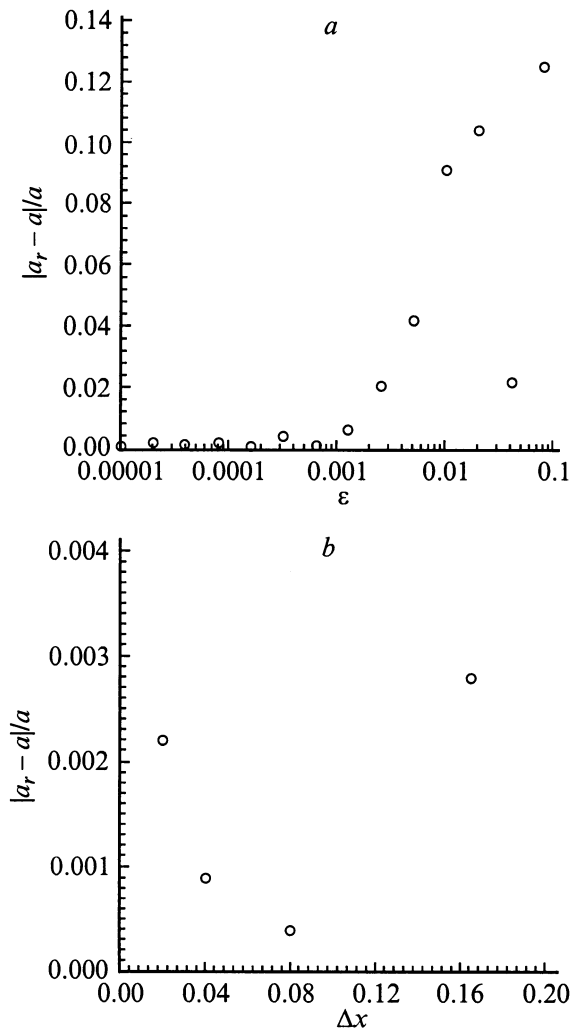


Рис. 2.

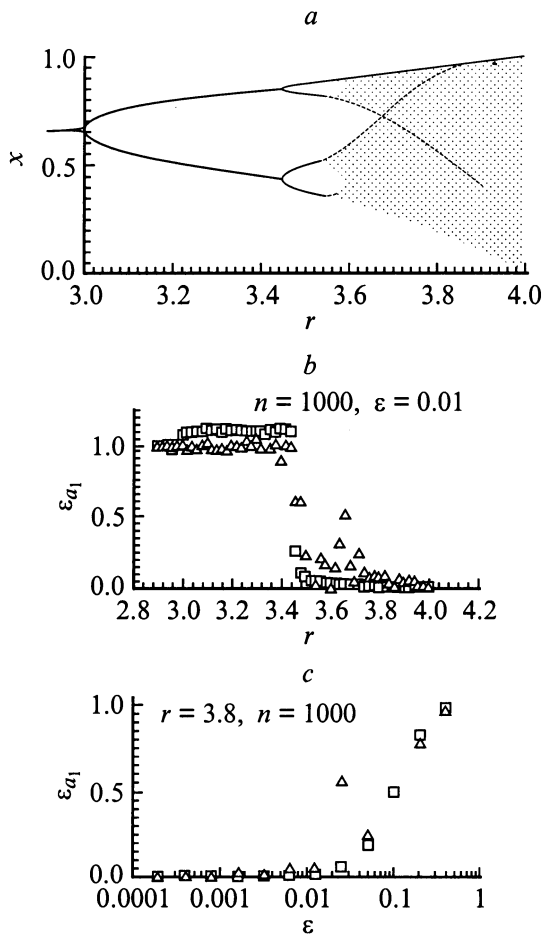


Рис. 3.

амплитуда которых не превышала 0.001 (для сравнения $\overline{q_1} = 0.0039$). Ошибки распределены по Гауссу с разбросом $\sigma = \varepsilon/\sqrt{3}$. Получились восстановленные значения $a_r = 1.2866$, $\eta_r = 0.3996$, $s_r = 2.031$.

На рис. 2 приведены зависимости относительной точности восстановленных значений от амплитуды шума $\varepsilon(a)$ от величины шаблона измерений $\Delta x(b)$ (при этом значение временного параметра τ таково, что оставалось постоянным "число Куранта" шаблона $a\tau/\Delta x + \eta\tau/\Delta x^2 = 0.8$) и от количества измерений. Прочие параметры равны значениям из раздела 1. Точность определения коэффициента a_r увеличивается при уменьшении погрешности измерений ε , при увеличении количества измерений N . Малая точность при уменьшении пространственного размера шаблона объясняется тем, что за соответствующий малый временной шаг τ изменение u становится меньше амплитуды шума.

2. Рассматривался установившийся процесс, описываемый существенно нелинейным логистическим отображением,

$$x_{i+1} = rx_i(1 - x_i), \quad 0 < r \leq 4, \quad (9)$$

бифуркационная диаграмма для которого приведена на рис. 3,а. Полагалась известной формула отображения

вида $x_{i+1} = a_1x_i + a_2x_i^2 + a_3$ с неизвестными коэффициентами a_1, a_2, a_3 . Для определения этих коэффициентов применялся принцип минимума информации (5) аналогично тому, как это делалось в случае с уравнением (1) (здесь измеряемыми величинами являлись $q_1^n = x_{i+1} + \xi_1^n$, $q_2^n = x_i + \xi_2^n$, $q_3^n = x_i^2 + \xi_3^n$ со случайными ошибками $\xi_j^n: |\xi_j^n| < \varepsilon$).

На рис. 3,б квадратиками представлена зависимость относительной точности восстановления первого коэффициента $\varepsilon_r = |a_1 - r|/r$ от параметра r отображения (9) при амплитуде шума $\varepsilon = 0.01$. Видно, что восстановление успешно, когда установившееся решение достаточно нетривиально, т.е. при $r \geq 3.5$, в частности, когда решением является странный аттрактор. Влияние шума проиллюстрировано на рис. 3,с. Для сравнения треугольниками показаны точки аналогичных расчетов по методу наименьших квадратов, предлагаемого для решения подобных задач в [3]. При достаточно большой амплитуде шума проявляются преимущества применения принципа минимума информации. Так, он позволяет восстановить коэффициенты в диапазоне параметра $r \approx 3.5-3.8$.

Заключение

Таким образом, показана принципиальная возможность восстановления неизвестных параметров уравнений из статистических данных эксперимента с помощью принципа минимума информации. Восстановление оказалось успешным как для процесса без нелинейных эффектов (1), так и для существенно нелинейных процессов (9) вплоть до детерминированного хаоса. Выявлены преимущества метода по сравнению с методом наименьших квадратов, к тому же, он имеет большее физическое обоснование [2]. Следовательно, метод может быть применен для адаптации общих математических моделей к условиям изучаемого явления с помощью экспериментальных данных.

Список литературы

- [1] Турчин В.Ф., Козлов В.П., Малкевич М.С. // УФН. 1970. Т. 102. Вып. 3. С. 345–386.
- [2] Хакен Г. Информация и самоорганизация. М.: Мир, 1991. 240 с.
- [3] Baker G.L., Gollub J.P., Blackburn J.A. // Chaos. 1996. Vol. 6. N 4. P. 528–533.