Глава 8. Модельные уравнения: оценка параметров

Движения и процессы, наблюдаемые в природе, исключительно разнообразны и сложны, поэтому возможности их моделирования явными функциями времени весьма ограничены. Значительно большими потенциальными возможностями обладают уравнения – разностные и дифференциальные (пп. 3.3, 3.5, 3.6). Даже простое одномерное отображение с квадратичным максимумом способно демонстрировать хаотическое поведение (п. 3.6.2). Такие модельные уравнения, *в отличие от явных временных зависимостей*, проблемы восстановления которых обсуждались в предыдущей главе, описывают *зависимость будущего состояния объекта от текущего* или *скорости изменения состояния от самого состояния*. Но технология построения этих более сложных моделей (оценка параметров, подбор аппроксимирующих функций) в основном такая же. Простой пример: построение одномерного модельного отображения $\eta_{n+1} = f(\eta_n, \mathbf{c})$ отличается от получения явной временной зависимости $\eta = f(t, \mathbf{c})$ только тем, что нужно провести кривую через экспериментальные точки на плоскости (η_n, η_{n+1}) (рис.8.1,а-в), а не на плоскости (t, η) (рис.7.1). В несколько более сложной задаче построения модельных ОДУ $d\mathbf{x}/dt = \mathbf{f}(\mathbf{x}, \mathbf{c})$ сначала численным дифференцированием получают временной ряд производных dx_k/dt (k = 1, …, D, где D – размерность модели), а затем обычным образом аппроксимируют зависимость каждой производной dx_k/dt от $\mathbf{x}$. Поскольку модельные уравнения могут быть многомерны, в задаче появляется своя специфика.

Долгое время для эмпирического моделирования сложных процессов использовались *линейные разностные уравнения*, в которые для обеспечения нерегулярности вводился шум (п. 4.4). Идея впервые была предложена в 1927 году [338] и оказалась очень плодотворной, на 50 лет вперед определив основной инструмент для описания сложного поведения – модели авторегрессии – скользящего среднего, см. лабораторную работу в [39].

Тот факт, что простые низкоразмерные модели в виде нелинейных отображений или дифференциальных уравнений могут демонстрировать сложные колебания даже в отсутствие шума, был осознан только в 1960-70-х годах. Это дало новый импульс развитию методов эмпирического моделирования, а появление мощной и доступной вычислительной техники обеспечило практическую реализацию идей.

В данной главе мы рассмотрим ситуацию, когда наблюдаемый временной ряд $\eta_i = h(\mathbf{x}(t_i)), i = 1,...,N$, задается итерациями отображения $\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{x}_n, \mathbf{c})$ или решением обыкновенного дифференциального уравнения $d\mathbf{x}/dt = \mathbf{f}(\mathbf{x}, \mathbf{c})$, структура которых *полностью известна*. Задача

состоит в *оценке параметров* **c** по наблюдаемым данным. Эту ситуацию мы договорились называть «прозрачный ящик» (п. 5.2). Для большей реалистичности будем добавлять в систему динамический и/или измерительный шум.

Такая постановка встречается в различных приложениях и привлекает серьезное внимание. Выделяются две основные проблемы, интересующие исследователей, которые мы рассмотрим поочередно ниже:

1) *получение оценок параметров с необходимой точностью*, что важно, если по условиям эксперимента параметры не могут быть измерены непосредственно. При этом процедура моделирования выступает в роли «измерительного прибора» [246, 272, 50, 248, 239, 287, 158, 316, 41] (п.8.1);

2) *оценка параметров в ситуации дефицита данных*, когда по имеющемуся ряду (возможно, векторной) наблюдаемой η не удается сформировать ряды всех динамических переменных модели $x_k, k = 1,...,D$, т.е. некоторые переменные являются «скрытыми» [187, 197, 283, 334, 41] (п. 8.2).

8.1. Оценки параметров и их точность

Проиллюстрируем варианты постановки задачи и методик ее решения на примере оценки параметра эталонного нелинейного отображения по временному ряду, представляющему собой его зашумленное решение. В качестве объекта используем квадратичное отображение в хаотическом режиме, считая неизвестным его единственный параметр $c = c_0$:

$$x_{n+1} = f(x_n, c_0) + \xi_n = 1 - c_0 x_n^2 + \xi_n, \ \eta_n = x_n + \zeta_n, \qquad (8.1)$$

где ξ_n, ζ_n – случайные процессы, первый из которых есть динамический шум (влияет на динамику), а второй – измерительный шум (влияет только на наблюдаемые значения).

В отсутствие обоих шумов имеем $\eta_n = x_n$, и экспериментальные точки на плоскости x_n, x_{n+1} лежат точно на искомой параболе (рис.8.1,а). Задача определения *c* сводится к алгебраическому уравнению, решение которого имеет вид $\hat{c} = (1 - x_{n+1}) / x_n^2$. При этом достаточно использовать любые два измерения x_n, x_{n+1} при $x_n \neq 0$. В результате модель совпадает с объектом с точностью до погрешностей вычислений.

При наличии шума в динамике или измерениях вместо точных значений ищутся *оценки* параметров. Наиболее известные методы оценивания описаны в п. 7.1.1. Рассмотрим, в чем отличие их применения при данной постановке задачи.

8.1.1. Динамический шум

Начнем с ситуации, когда в системе (8.1) есть только *динамический* шум. Ограничимся наиболее часто рассматриваемым случаем: ξ_n – последовательность случайных величин, которые статистически независимы и одинаково распределены с плотностью $p(\xi)$. Для оценки параметров **c** часто используют метод максимального правдоподобия (пп. 2.2.1.7, 7.1.1.3, 7.1.2, 7.1.5), который наиболее эффективен при достаточно общих условиях [81, 287]. Функция правдоподобия (см. (2.26) и (7.10)) принимает вид:

$$\ln L(c) \equiv \ln p(\eta_1, \eta_2, ..., \eta_N | c) \approx \sum_{n=1}^{N-1} \ln p_\xi \left(\eta_{n+1} - f(\eta_n, c)\right). \qquad (8.2)$$

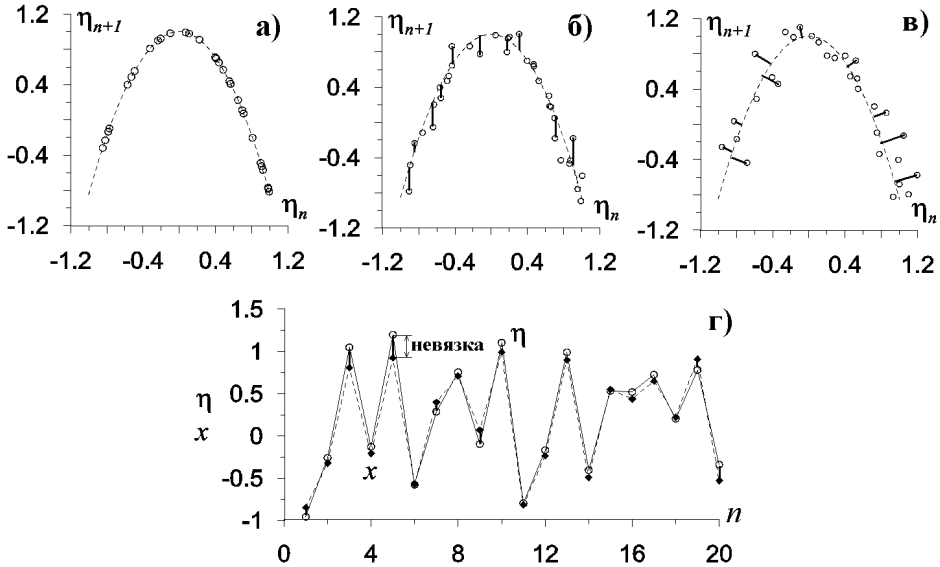

Рис.8.1. Оценка параметров на примере квадратичного отображения (8.1) при c_0 = 1.85, кружки – наблюдаемые значения: а) нет шума, пунктир – исходная парабола; б) только динамический шум, пунктир – модельная парабола, полученная минимизацией среднего квадрата вертикальных расстояний (некоторые показаны жирными линиями); в) только измерительный шум, пунктир – модельная парабола, полученная минимизацией среднего квадрата ортогональных расстояний; г) только измерительный шум, ромбики – реализация модели, наиболее близкая к наблюдаемому ряду в смысле (8.4)

Для применения метода необходимо знать закон распределения $p(\xi)$, что редко имеет место. Чаще всего предполагают нормально распределенный шум; при этом очевидно, что максимизация (8.2) эквивалентна так называемому *обычному методу наименьших квадратов*, т.е. минимизации

$$S(c) = \sum_{n=1}^{N-1} \left(\eta_{n+1} - f(\eta_n, c)\right)^2 \to \min. \qquad (8.3)$$

Это означает, что график модельной функции на плоскости (η_n, η_{n+1}) должен пройти так, чтобы сумма квадратов *вертикальных расстояний* от экспериментальных точек до него была минимальна (рис.8.1,б).[1]

Как правило, ошибка оценки $\hat{c}$ уменьшается с ростом длины ряда *N*. В рассматриваемой сейчас постановке задачи ММП и обычный МНК дают асимптотически несмещенные и состоятельные оценки. Можно показать, что дисперсия оценок убывает пропорционально N^{-1}, аналогично задачам п. 7.1.2.1. Причину такого закона спадания можно описать на том же языке: слагаемые в (8.3) стационарны по *i*, т.е. частные информационные количества Фишера (п. 7.1.2.2) ограничены.

Заметим, что если функция *f* линейна по *x*, то модель (8.1) – это линейная модель авторегрессии 1-го порядка. Более общие модели авторегрессии – скользящего среднего включают зависимость x_{n+1} от предыдущих значений *x* и ξ, см. (4.13) в п. 4.4, [47] и лабораторную работу [39].

8.1.2. Измерительный шум

Если присутствует только измерительный шум ($\eta_n = x_n + \zeta_n$), то, как мы увидим ниже, задача оценивания усложняется. Это обстоятельство связано с тем, что для искомой зависимости x_{n+1} от x_n «зашумлены» и наблюдаемые значения «независимой» переменной x_n (см. п.2.3.1.3).

8.1.2.1. Смещенность оценок, полученных обычным МНК. Она имеет место при сколь угодно длинном ряде, так как метод (8.3) рассчитан только на динамический шум. Покажем это на примере (8.1). Имеем

$$S(c) = \sum_{i=1}^{N-1} \left(\eta_{i+1} - f(\eta_i, c)\right)^2 = \sum_{i=1}^{N-1} \left(x_{i+1} + \xi_{i+1} - 1 + c\left(x_i + \xi_i\right)^2\right)^2 =$$
$$= \sum_{i=1}^{N-1} \left(c x_i^2 - c_0 x_i^2 + \xi_{i+1} + 2c x_i \xi_i + c \xi_i^2\right)^2.$$

Найдем минимум *S* по *c* из условия $\partial S / \partial c = 0$, которое приводится к виду:

---

[1] МНК часто дает приемлемую точность оценок, даже если шум не нормальный, т.е. метод имеет самостоятельное значение (пп. 7.1.1.4, 7.1.3). Хотя допустимо применять и другие методы, например, метод наименьших модулей, МНК гораздо удобнее в реализации (п. 7.1.5). Существенная техническая проблема появляется, если «рельеф» целевой функции (8.3) содержит много локальных минимумов, что часто имеет место, когда *f* нелинейна по *c*. Тогда задача оптимизации решается итерационным путем, начиная с некоторой стартовой «догадки» относительно искомых параметров [73]. Будет ли найден глобальный экстремум, зависит от «удачи» при подборе стартовой догадки, ее близости к истинному значению параметра. В примере (8.1) *f* линейна по *c*, поэтому целевая функция *S* квадратична по *c* и имеет единственный глобальный минимум, который легко находится решением линейного алгебраического уравнения.

$$\sum_{i=1}^{N-1}\left(cx_i^2 - c_0x_i^2 + \xi_{i+1} + 2cx_i\xi_{i+1} + c\xi_i^2\right)\cdot\left(x_i^2 + 2x_i\xi_{i+1} + \xi_i^2\right) = 0\,.$$

Решая это уравнение, получим оценку

$$\hat{c} = \frac{c_0\left(\sum_{i=1}^{N-1}x_i^4 + 2\sum_{i=1}^{N-1}x_i^3\xi_i + \sum_{i=1}^{N-1}x_i^2\xi_i^2\right) - \sum_{i=1}^{N-1}x_i^2\xi_{i+1} - 2\sum_{i=1}^{N-1}x_i\xi_i\xi_{i+1} - \sum_{i=1}^{N-1}\xi_i^2\xi_{i+1}}{\sum_{i=1}^{N-1}x_i^4 + 6\sum_{i=1}^{N-1}x_i^2\xi_i^2 + \sum_{i=1}^{N-1}x_i^4 + 4\sum_{i=1}^{N-1}x_i^3\xi_i + 4\sum_{i=1}^{N-1}x_i\xi_i^3}\,.$$

Устремим теперь N к бесконечности, учтем независимость ξ_i от ξ_{i+1} и от x_i, заменим суммы типа $\sum_{i=1}^{N-1}x_i^4$ (усреднение по времени) на интегралы типа $\int_{-\infty}^{\infty}\mu(x,c_0)x^4dx \equiv \left\langle x^4\right\rangle$ (усреднение по ансамблю). Здесь $\mu(x,c_0)$ – инвариантная мера (плотность распределения точек на аттракторе) для отображения (8.1). При $c_0 = 2$ она может быть найдена аналитически: $\mu(x,2) = 1/\pi\left(1-x^2\right), -1 < x < 1$. Отсюда получаем $\left\langle x^2\right\rangle = 1/2$, $\left\langle x^4\right\rangle = 3/8$, $\left\langle x^n\right\rangle = 0$ для нечетных n. В итоге при $c_0 = 2$ получим асимптотическое (для $N \to \infty$) значение оценки $\hat{c} = c_0\,(4\sigma_\xi^2 + 3)\big/(8\left\langle\xi^4\right\rangle + 24\sigma_\xi^2 + 3)$.

Это оценка смещенная: она занижена, т.к. знаменатель больше числителя. На рис.8.2 показана зависимость этого асимптотического значения от отношения шум/сигнал для нормально распределенного шума. Оно практически не отличается от истинного значения только при малых шумах (смещение меньше 1%, если уровень шума меньше 0.05). Смещение оценки тем больше, чем больше уровень шума. Аналогичные свойства наблюдаются и при других распределениях шума [272, 50]. Но поскольку обычный МНК прост в реализации и может легко использоваться в случае большого числа оцениваемых параметров, в отличие от других методов, его часто применяют на практике, предполагая малый измерительный шум.

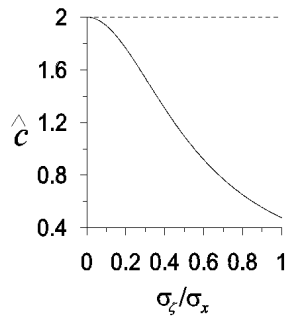

Рис.8.2. Оценки параметра c в (8.1), полученные с помощью обычного МНК, при $N = \infty$, $c_0 = 2$ и дисперсии шума σ_ζ^2. По оси абсцисс отложено отношение шум/сигнал, измеряемое здесь как отношение их среднеквадратичных

8.1.2.2. Как повысить точность оценок в случае значительного измерительного шума? Отчасти это удается при использовании так

называемого метода *полных наименьших квадратов* [246], когда минимизируется сумма квадратов ортогональных расстояний от точек (η_n, η_{n+1}) до графика $f(x_n, c)$, рис.8.1,в. При этом учитывается то обстоятельство, что отклонение наблюдаемых точек с координатами (η_n, η_{n+1}) от графика искомой функции $f(x_n, c_0)$ вызвано действием шума на обе координаты. Поэтому отклонение может происходить в любом направлении, а не только по вертикали. Использование именно ортогональных расстояний обосновано в [246] как приближенный вариант ММП.

Однако смещенность оценок при использовании метода полных наименьших квадратов полностью не устраняется (особенно при очень больших шумах), т.к. это только приближение к ММП. Казалось бы, выход состоит в том, чтобы «честно» записать функцию правдоподобия для новой ситуации с учетом способа вхождения шумов. При нормальном шуме задача сводится к минимизации суммы квадратов отклонений *реализации* модели от *наблюдаемого ряда* (рис.8.1,г):

$$S(c, x_1) = \sum_{n=0}^{N-1} \left(\eta_{n+1} - f^{(n)}(x_1, c)\right)^2 \to \min, \qquad (8.4)$$

где $f^{(n)}$ – n-я итерация отображения $x_{n+1} = f(x_n, c)$, $f^{(0)}(x, c) = x$, и в число оцениваемых величин включено начальное состояние модели x_1.

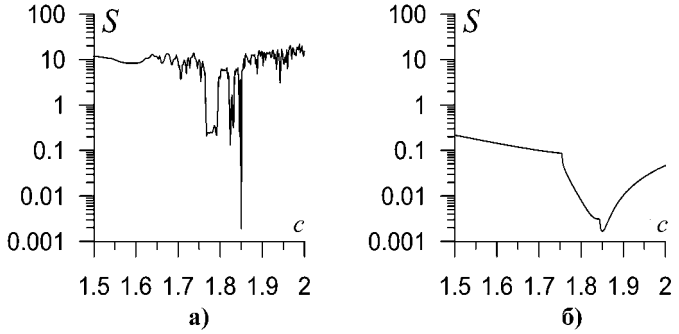

Рис.8.3. Целевые функции для квадратичного отображения (8.1) при N = 20, $c_0 = 1.85$, $x_1 = 0.3$: слева – целевая функция для итераций в прямом времени (8.4), справа – целевая функция для итераций в обратном времени (8.5)

Траектория хаотической системы очень чувствительна к начальным условиям и параметрам. Поэтому дисперсия оценок (8.4) убывает в таком случае очень быстро с ростом N, иногда даже экспоненциально [287, 239]. Это желательное свойство, но оно имеет место на практике только при условии, что всегда удается находить глобальный минимум (8.4). Даже при умеренном N график функции S в случае хаотической системы становится сильно «изрезанным» (рис.8.3,а), так что найти глобальный минимум численными методами [73] практически невозможно. Для этого потребовалось бы очень «удачное» задание стартовых догадок для c и x_1. Об асимптотических свойствах оценок тоже говорить трудно, т.к. в пределе $N \to \infty$ целевая

функция становится негладкой. Поэтому разрабатываются модификации ММП в приложении к задаче оценки параметров модели по хаотическому ряду [287, 316].

Так, предлагалось делить исходный ряд на сегменты небольшой длины L, по каждому из которых можно найти минимум (8.4), и усреднять полученные оценки (кусочный метод). Это целесообразный подход, но итоговая оценка может оставаться асимптотически смещенной, а ее дисперсия убывает опять только как N^{-1}. Некоторые варианты улучшения свойств оценок описаны ниже (п. 8.2).

Здесь мы отметим только специальный подход, предложенный в [158] для случая одномерных отображений. Он основан на использовании того свойства, что единственный ляпуновский показатель одномерного хаотического отображения в обратном времени становится отрицательным и траектория отображения не столь чувствительна к параметрам и «начальному» условию. Поэтому минимизируется величина

$$S(c,x_1)=\sum_{n=0}^{N-1}\left(\eta_{N-n}-f^{(-n)}(x_N,c)\right)^2\to\min, \qquad (8.5)$$

где $f^{(-n)}$ – n-ая итерация отображения $x_{n+1}=f(x_n,c)$ в обратном времени. Для сколь угодно длинного ряда график целевой функции выглядит достаточно плавным (рис.8.3,б), и найти глобальный минимум не сложно. При малых и умеренных уровнях шума (σ_ζ/σ_x до 0.05-0.15) погрешности метода (8.5) оказываются меньше, чем для кусочного. Причем для малых шумов (8.5) дает асимптотически несмещенные оценки и дисперсия убывает в типичном случае как N^{-2}. Эта высокая скорость обусловлена возвратами траектории в окрестность экстремума функции f [316].

8.2. Скрытые переменные

Когда уровень измерительного шума значителен, переменную состояния x часто считают «скрытой», т.к. ее значения, строго говоря, не известны. «Еще более скрытыми» являются переменные, зашумленные значения которых нельзя ни непосредственно измерить, ни вычислить из рядов наблюдаемых величин, что часто имеет место на практике. В этом случае получение оценок параметров гораздо более проблематично, чем в задачах п. 8.1. Но если удается успешно провести эту процедуру, то, как ее побочный продукт, появляется дополнительная возможность – получить временные ряды скрытых переменных. Тогда процедура моделирования опять выступает как измерительный прибор, но уже и в отношении динамических переменных, а не только параметров.

8.2.1. Методы

Мы проиллюстрируем их на примере оценки параметров обыкновенных дифференциальных уравнений без динамического шума. Объект – классическая хаотическая система (система Лоренца):

$$\begin{aligned} \dot{x}_1 &= c_1(x_2 - x_1), \\ \dot{x}_2 &= -x_2 + x_1(c_3 - x_3), \\ \dot{x}_3 &= -c_2 x_3 + x_1 x_2, \end{aligned} \tag{8.6}$$

с параметрами $c_1 = 10, c_2 = 8/3, c_3 = 46$. Наблюдается только зашумленная реализация величины x_1: $\eta_n = x_1(t_n) + \zeta_n$, переменные x_2 и x_3 – скрытые. Модель имеет вид (8.6), все три параметра c_i считаются неизвестными.

8.2.1.1. Метод подгонки начального условия. Все методики оценки основаны так или иначе на идее типа (8.4), т.е. подбираются такое начальное состояние модели и параметры, чтобы ее реализация была близка к наблюдаемому ряду в смысле наименьших квадратов. Непосредственное решение задачи типа (8.4) называют методом подгонки начального условия [238, 334]. Согласно сказанному в п.8.1.1.2, он неприменим для длинных хаотических рядов. Простое деление ряда на сегменты с последующим усреднением оценок дает низкую точность итоговой оценки, а итерации в обратном времени для многомерной диссипативной системы не пригодны.

8.2.1.2. Метод множественной стрельбы. Отчасти обойти трудности и использовать более длинные ряды позволяет алгоритм Бока [196, 187], который называют еще методом множественной стрельбы, т.к. при этом от решения задачи Коши для получения траектории модели на всем интервале наблюдения переходят к решению набора краевых задач. А именно, разбивают исходный ряд $\{\eta_1, \eta_2, ..., \eta_N\}$ на L неперекрывающихся сегментов длины M и рассматривают начальные состояния модели $\mathbf{x}^{(i)}$ на каждом их них (т.е. в моменты времени $t_{(i-1)M+1}$, $i = 1,...,L$) как оцениваемые величины, но *не свободные*. Решается задача условной минимизации, которая в случае скалярной наблюдаемой $\eta = h(\mathbf{x}) + \zeta$ (в нашем примере $\eta = x_1 + \zeta$) записывается в виде

$$\begin{aligned} &S(\mathbf{c}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(L)}) = \sum_{i=1}^{L} \sum_{n=0}^{M-1} \left[\eta(t_{(i-1)M+n}) - h\left(\mathbf{x}(t_{(i-1)M+n}, \mathbf{x}^{(i)}, \mathbf{c})\right)\right]^2 \to \min, \\ &\mathbf{x}(t_M, \mathbf{x}^{(i)}, \mathbf{c}) = \mathbf{x}^{(i+1)}, \ i = 1,...,L-1. \end{aligned} \tag{8.7}$$

Здесь выражение $\mathbf{x}(t, \mathbf{x}^{(i)}, \mathbf{c})$ задает траекторию модели (решение модельных уравнений), т.е. состояние модели $\mathbf{x}$ в момент времени t при

начальном условии $\mathbf{x}^{(i)}$ и параметрах $\mathbf{c}$. Первая строка означает минимизацию отклонения реализации модели от ряда для всего интервала наблюдения, вторая – «сшивание» сегментов, чтобы в итоге получить непрерывную траекторию модели на всем интервале наблюдения. «Сшивание» накладывает $L-1$ условие типа равенства на L искомых векторов $\mathbf{x}^{(i)}$, т.е. только один из них можно считать «свободным параметром задачи».

Как обычно, задача решается численными итерационными методами, начиная со стартовых догадок для искомых величин $\mathbf{c}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(L)}$. Стартовым догадкам соответствует, как правило, разрывная траектория модели, состоящая из L совсем не стыкующихся друг с другом кусков (рис.8.4,б, верхняя панель). Подобная ситуация имеет место и для промежуточных значений искомых величин в процессе работы итерационного метода минимизации, но «нестыковки» должны становиться все меньше и меньше, если метод сходится при заданных стартовых догадках. Это «временное» допущение разрывности траектории модели отличает алгоритм Бока (рис.8.4,б) от метода подгонки начального условия (рис.8.4,а) и обеспечивает большую гибкость первого.

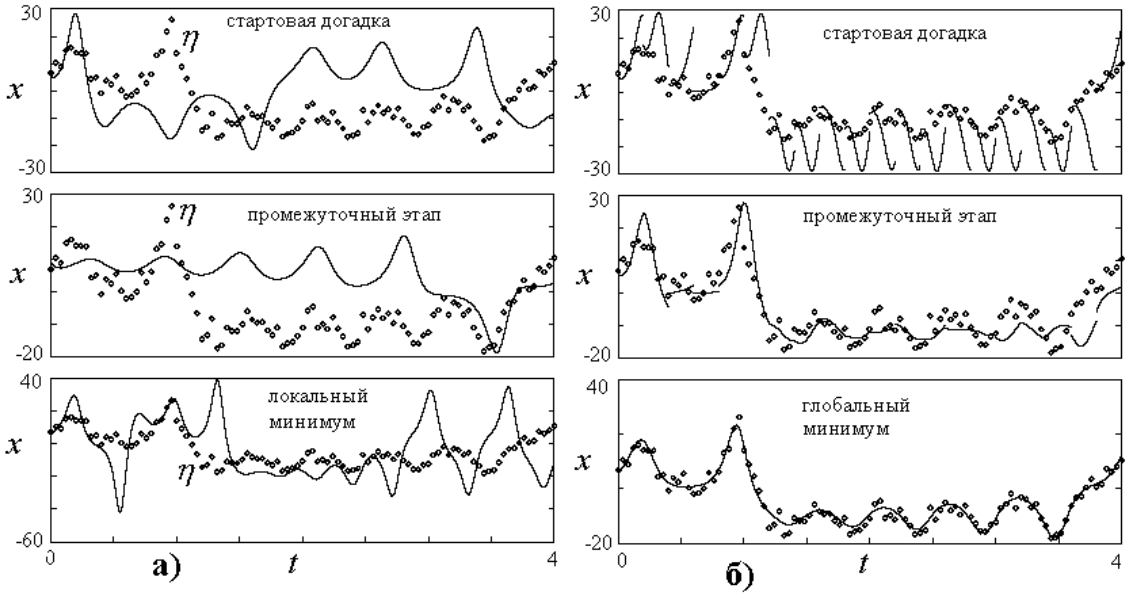

Рис.8.4. Оценка параметров по хаотической реализации координаты $x = x_1$ системы Лоренца (8.6), $N = 100$ точек, выборочный интервал 0.04, нормальный измерительный шум со стандартным отклонением 0.2 от стандартного отклонения сигнала [238]: а) метод подгонки начального условия. Реализации наблюдаемой (ромбики) и соответствующей модельной переменной. Процесс подгонки сходится к локальному минимуму, где траектория модели и оценки параметров сильно отличаются от истинных; б) метод множественной стрельбы. Процесс подгонки сходится к глобальному минимуму, где траектория модели и оценки параметров близки к истинным

На рис.8.4 представлен пример применения методов для оценки параметров системы (8.6). Метод множественной стрельбы «находит» глобальный минимум, тогда как метод подгонки начального условия останавливается в локальном. Это весьма типичная ситуация.

Однако и метод множественной стрельбы не гарантирует нахождение именно глобального минимума, а только смягчает требования к удачности стартовых догадок для искомых величин [41]. Для еще более длинного ряда и он станет неэффективным, т.к. требование близости хаотической траектории модели к наблюдаемому ряду на длинном интервале приведет вновь к слишком жестким требованиям на удачность стартовых догадок.

8.2.1.3. Модификация метода множественной стрельбы. Как показано в [41], снизить остроту проблемы можно, допустив разрывность итоговой траектории модели в некоторые моменты времени на интервале наблюдения, т.е. отказавшись от некоторых равенств во второй строке (8.7). При этом легче отыскать глобальный минимум целевой функции S в (8.7).

Такая модификация позволяет использовать сколь угодно длинные хаотические ряды, а расплата состоит в том, что иногда модель с неадекватной структурой может быть признана «хорошей» за счет способности воспроизвести короткий участок ряда. Поэтому требуется осторожный выбор числа и размера сегментов непрерывности траектории модели. Отметим также дополнительную трудность ситуации со скрытыми переменными: кроме удачных стартовых догадок для параметров $\mathbf{c}$ очень важно найти и удачные стартовые догадки для скрытых переменных – компонент $\mathbf{x}^{(i)}$ – в отличие от чрезмерно оптимистичных ранних утверждений [187]. Часто приходится действовать наугад, но полезную информацию можно получить из предварительного изучения свойств модельных уравнений при пробных значениях параметров [41].

8.2.1.4. Заключительные комментарии. Существуют и развиваются способы оценки параметров и скрытых переменных, пригодные в случае одновременного наличия динамического и измерительного шумов. Они основаны на байесовском подходе [274, 217, 198] и использовании модифицированного фильтра Калмана [307, 334]. На эту широкую область исследований мы обращаем внимание читателя, но рассмотреть не имеем возможности.

Верификация моделей в рассмотренных в данной главе задачах о прозрачном ящике проводится с помощью двух основных подходов. Во-первых, это анализ остатков (невязок) модели, т.е. проверка их соответствия предполагаемым свойствам шумов (см. п. 7.3) [45]. Во-вторых, это расчет динамических, геометрических и топологических

характеристик аттрактора модели и их сравнение с соответствующими свойствами объекта [233], см. п. 10.4.

8.2.2. Что дают успехи и какая польза от неудач моделирования?

Успешная реализация описанных методов обещает получение не только оценок параметров, но и временных рядов скрытых переменных, а также сулит ряд соблазнительных приложений: проверка адекватности заложенных в модель представлений, «измерение» величин, недоступных прибору экспериментатора, восстановление утерянных или искаженных участков временной зависимости измеряемой величины. Однако и неудачи дают полезную информацию. Прокомментируем это подробнее.

На практике при исследовании реальных объектов никогда не встречается в чистом виде задача о прозрачном ящике. Исследователь только субъективно может верить в то, что пробная структура модели адекватна объекту. Поэтому даже при идеальном выполнении действий нижнего яруса схемы рис.5.1 результат моделирования может оказаться отрицательным, т.е. не удастся получить адекватную модель с заданной структурой. Тогда исследователю придется признать несправедливость его представлений о механизмах протекания процесса и вернуться к этапу выбора структуры модели. Если есть несколько альтернативных математических конструкций, то результаты моделирования по рядам могут выявить из них наиболее адекватную. Таким образом, процедура моделирования даст возможность опровергнуть или подтвердить (может быть, уточнить) содержательные представления об исследуемом объекте.

Практические примеры успехов уже имеются. В [240] подтверждена справедливость модельных представлений о работе газового лазера и получены не измеряемые непосредственно параметры скоростей перехода между энергетическими уровнями в зависимости от тока накачки. В [321] представлен показательный пример успешного моделирования и содержательного вывода о механизме биохимического сигнального процесса в клетках, с которым мы в кратком пересказе познакомим читателей.

8.2.2.1. Пример из клеточной биологии. Для многих приложений необходимо выяснить, изменение каких свойств клеток ведет к наиболее эффективному подавлению нежелательного процесса в них, и как целенаправленно воздействовать именно на эти свойства.[2] Для ответа на

[2] Так, рост раковых клеток обусловлен тем, что они производят вещества «неадекватно» окружающей обстановке. Вариант борьбы с заболеванием, который пока только в предположениях, мог бы опираться на эмпирическое моделирование, подобное описанному в п. 8.2.2.1. Здесь свойства биохимического процесса в клетках типа BaF3 [321] и возможность управления ими выяснены с опорой на экспериментальный временной ряд и привлечение биохимических соображений для формирования структуры модели.

поставленные вопросы может быть достаточно получения адекватной математической модели. Рассмотрим пример успешного моделирования сложного биологического процесса по временным рядам.

Моделируемый процесс – одна из разновидностей внутриклеточных сигнальных путей, которые дают клетке возможность производить нужные ей вещества в ответ на изменение окружающей обстановки. В частности, они обеспечивают размножение, дифференциацию и выживание клеток. Здесь рассмотрен так называемый сигнальный путь JAK-STAT [321], который преобразует внешний химический сигнал в активацию транскрипции с соответствующего гена внутри ядра клетки, см. схему на рис.8.5. Одна из самых простых математических моделей этого биохимического процесса может быть записана, исходя из закона действующих масс (обычный подход в химической кинетике), и имеет вид:

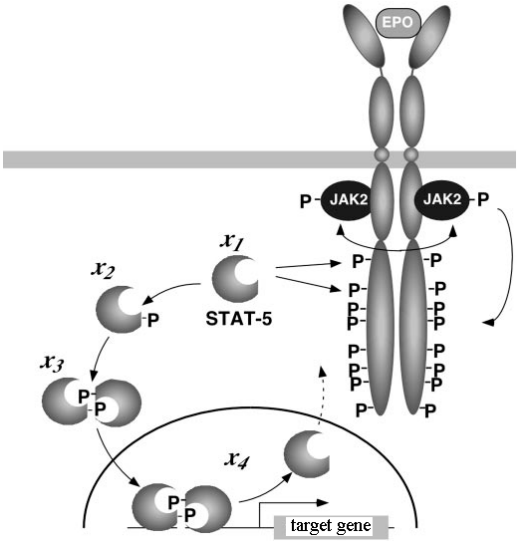

Рис.8.5. Схема биохимического сигнального процесса в клетке [321]

$$
\begin{aligned}
&dx_1/dt = -k_1 x_1 E(t),\\
&dx_2/dt = -k_2 x_2^2 + k_1 x_1 E(t),\\
&dx_3/dt = -k_3 x_3 + k_2 x_2^2/2,\\
&dx_4/dt = k_3 x_3.
\end{aligned}
\tag{8.8}
$$

Здесь k_i – параметры скоростей реакций, $E(t)$ (Epo на рис.8.5) – концентрация вещества эритропойетина в окружающей среде, в ответ на изменение которой активируются соответствующие рецепторы в мембране клетки. Рецепторы связываются с тирозинкиназами типа JAK-2, которые имеются в цитоплазме клетки. Последние вступают в реакцию с молекулами вещества STAT5, концентрация которых обозначена x_1. В результате реакции последние фосфорилируются. Получаются мономерные фосфорилированные молекулы STAT5, концентрация которых обозначена x_2. Эта реакция ведет к убыванию концентрации x_1 (первое уравнение (8.8)) и увеличению концентрации x_2 (положительное слагаемое во втором уравнении (8.8)). Мономерные молекулы, встречаясь друг с другом, димеризуются. Концентрация димеризованных молекул – x_3. Эта реакция ведет к убыванию x_2 (отрицательное слагаемое во втором

уравнении (8.8)) и увеличению x_3 (положительное слагаемое в третьем уравнении (8.8)). Димеризованные молекулы проникают в ядро, их концентрация в ядре – x_4. Этот процесс ведет к убыванию x_3 (отрицательное слагаемое в третьем уравнении (8.8)) и увеличению x_4 (четвертое уравнение (8.8)). В ядре димеризованные молекулы активируют транскрипцию гена, обозначаемого CIS (target gene на рис.8.5). В результате производится некоторый белок. После этого димеризованные молекулы распадаются на мономерные, которые *согласно гипотезе, лежащей в основе модели* (8.8), *деградируют в ядре*.

Но есть и другая версия: мономерные молекулы STAT5 *возвращаются с некоторой задержкой из ядра в цитоплазму клетки*. При таком предположении математическая модель несколько меняется и принимает вид:

$$\begin{aligned} dx_1(t)/dt &= -k_1 x_1(t)E(t) + 2k_4 x_3(t-\tau), \\ dx_2(t)/dt &= -k_2 x_2^2(t) + k_1 x_1(t)E(t), \\ dx_3(t)/dt &= -k_3 x_3(t) + k_2 x_2^2(t)/2, \\ dx_4(t)/dt &= -k_4 x_3(t-\tau) + k_3 x_3(t), \end{aligned} \tag{8.9}$$

где появляются слагаемые с запаздыванием в первом и четвертом уравнениях, τ – время запаздывания. Какая именно из двух моделей (т.е. какая из двух гипотез) более адекватна – неизвестно. Возможности наблюдения крайне ограничены, техника весьма дорогостояща и процесс измерений очень сложен [321]. В эксперименте можно было измерять с точностью до постоянного множителя только суммарную массу STAT5 в цитоплазме η_1 и суммарную массу фосфорилированного STAT5 в цитоплазме η_2:

$$\begin{aligned} \eta_1 &= k_5(x_2 + 2x_3), \\ \eta_2 &= k_6(x_1 + x_2 + 2x_3), \end{aligned} \tag{8.10}$$

где k_5, k_6 – неизвестные коэффициенты пропорциональности. Кроме этих двух наблюдаемых измерялась и величина концентрации $E(t)$, и тоже с точностью до коэффициента пропорциональности (рис.8.6,а). Подчеркнем, что все переменные моделей – скрытые, есть лишь две наблюдаемых, которые связаны с четырьмя переменными известным образом (8.10)

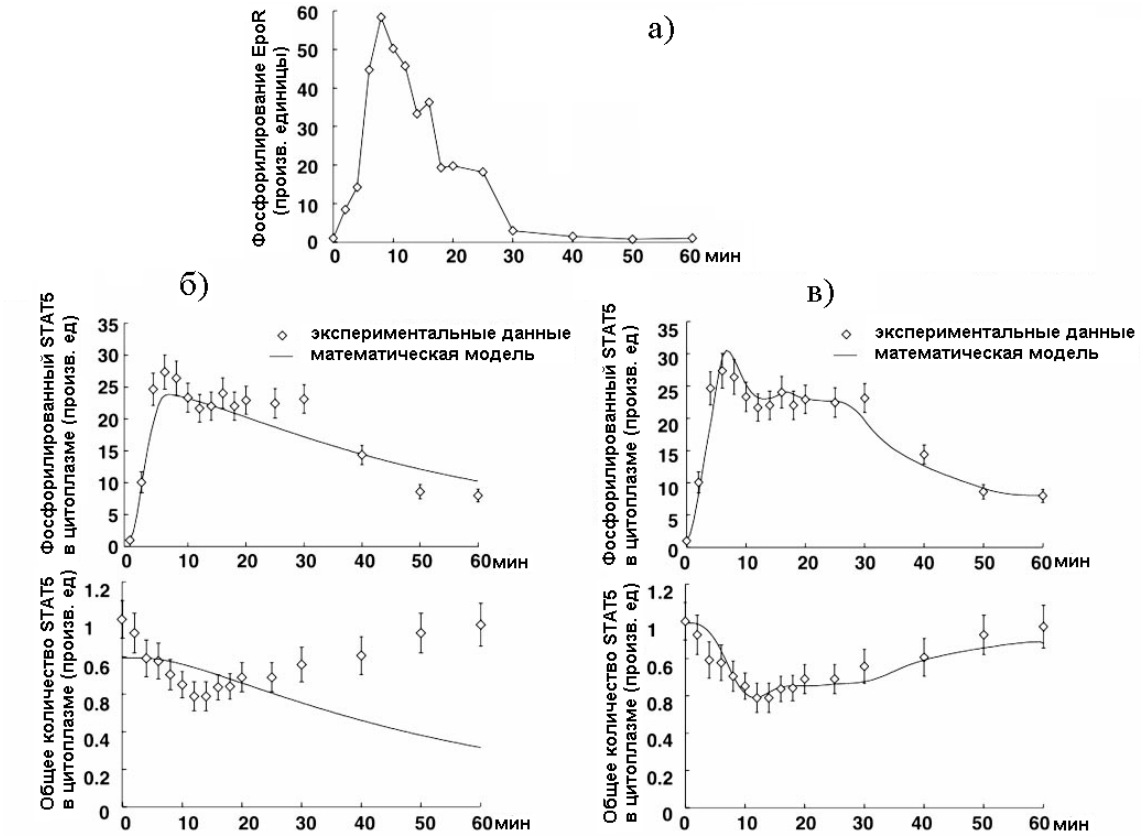

Рис.8.6. Результаты моделирования биохимического сигнального процесса в клетке [321]: а) временной ряд внешнего воздействия – изменения концентрации эритропойетина; б) результаты моделирования с помощью (8.8) по двухканальному ряду; в) результаты моделирования с помощью (8.9)

В работе [321] по описанным экспериментальным данным оценивались параметры обеих моделей (8.8) и (8.9). Была показана неадекватность первой и хорошее соответствие с экспериментом – второй. А именно, были измерены временные ряды в трех независимых экспериментах. Один из экспериментов представлен на рис.8.6. Параметры моделей оценивались по данным всех трех экспериментов вместе с помощью метода подгонки начального условия (п. 8.2.1.1). Он наиболее уместен, т.к. ряды короткие – всего по 16 точек в каждом (это отклики на «импульсное» изменение концентрации эритропойетина, рис.8.6,а), так что проблем с поиском глобального минимума целевой функции не возникает. Модель (8.8) оказалась не способной воспроизвести сигнал (рис.8.6,б), а модель (8.9) вполне адекватна в этом отношении (рис.8.6,в), оценка времени запаздывания составила $\tau \approx 6$ минут, что согласуется по порядку величины с результатами других авторов, полученными другими методами для похожих объектов.

Таким образом, только с помощью моделирования по рядам авторы смогли сделать нетривиальный вывод о том, что возврат молекул STAT5 в цитоплазму играет существенную роль в изучаемом процессе. Выяснены и

некоторые детали процесса, которые нельзя наблюдать непосредственно, например, пребывание молекул STAT5 в ядре в среднем около 6 минут.

Исследуя модель (8.9), можно предсказать, что произойдет при изменении тех или иных параметров. Например, авторы исследовали, как изменится суммарное количество произведенного клеткой вещества (оно пропорционально общему количеству STAT5, принявшему участие в процессе) при изменении разных параметров модели. Оказалось, что оно очень слабо зависит от параметров k_1, k_2 и сильно – от k_3, k_4, τ. Т.е. процессы в ядре клетки играют первостепенную роль. Согласно модели (8.9) уменьшение k_4 до нуля (подавление экспорта из клетки) ведет к уменьшению произведенного вещества на 55 %. В эксперименте с помощью добавления лептомицина B снижали экспорт из клетки (аналог параметра k_4) только на 60 %. Согласно модели это должно привести к уменьшению количества активированного STAT5 на 40 %. В эксперименте наблюдалось уменьшение на 37 %. Таким образом, модельный прогноз отлично подтвержден, что еще более повышает доверие к модели и позволяет использовать ее для более детального изучения процесса и управления им в случае необходимости. В связи с этим достижением уже более предметно можно говорить о возможностях использования эмпирических моделей для медицинских целей, упомянутых на с. 239.

Однако следует заметить, что, несмотря на указанные успехи, задача моделирования часто может оказаться технически неразрешимой даже в рассмотренной относительно простой постановке «прозрачного ящика» при адекватной структуре модели. Пока примеры успешного моделирования имели место для случаев, когда разница между числом наблюдаемых и размерностью модели не более 2-3 и число оцениваемых параметров не более 3-5 (это грубые ориентировочные оценки, в каждом конкретном случае результат зависит от вида нелинейности в модельных уравнениях). Чем больше скрытых переменных и неизвестных параметров модели, тем меньше шансов на успех и хуже свойства получаемых оценок.