

АКАДЕМИЯ НАУК СССР

УСПЕХИ ХИМИИ

ВЫПУСК 2

ФЕВРАЛЬ — 1980 г.

ТОМ XLIX

МОСКВА

ЖУРНАЛ ОСНОВАН В 1932 ГОДУ
ВЫХОДИТ 12 РАЗ В ГОД

УДК 541.124

СОВРЕМЕННОЕ СОСТОЯНИЕ ПРОБЛЕМЫ ИДЕНТИФИКАЦИИ КИНЕТИЧЕСКИХ МОДЕЛЕЙ

B. B. Кафаров, B. N. Писаренко

Изложено современное состояние проблемы идентификации кинетических моделей и основные пути ее решения. Рассмотрены принципы оценки констант в нелинейных кинетических моделях и установлены свойства получаемых оценок. Проанализированы способы определения адекватности однотипных моделей, статистической значимости их отдельных составляющих и точности оценок кинетических параметров. Приведены байесовские методы и отношения вероятностей, используемые при проверке гипотез о механизме протекания сложной химической реакции, а также методы установления причины возможных неадекватностей кинетических моделей и способы направленного построения адекватной модели. Показана важность активного планирования кинетического эксперимента для эффективного решения проблемы идентификации моделей.

Библиография — 80 ссылок.

ОГЛАВЛЕНИЕ

I. Введение	193
II. Построение системы конкурирующих гипотез о механизме реакции	194
III. Оценка констант в кинетических моделях	198
IV. Проверка адекватности кинетических моделей	201
V. Планирование прецизионных экспериментов	206
VI. Дискриминация моделей	212
VII. Планирование дискриминирующих экспериментов	218

I. ВВЕДЕНИЕ

Проблема идентификации кинетических моделей является одной из кардинальных в любой области химии — физической, органической, биологической. Это объясняется, в частности, тем, что решение этой проблемы во многом определяет состояние и уровень развития теории реакционной способности химических соединений, которая позволяет в свою очередь прогнозировать протекание различных процессов и устанавливать пути направленного синтеза новых химических веществ с заданными свойствами. Проблема идентификации кинетических моделей имеет и важное практическое значение, ибо как быстрое и эффективное

внедрение новых процессов в промышленности, так и проектирование отдельных химических производств невозможно без установления структуры и динамики химического превращения, причем чем детальнее изучена кинетика химической реакции, тем точнее результаты всех последующих промышленных прогнозов. Однако несмотря на свое значение для определения механизма химической реакции, эта проблема остается до сих пор изученной не до конца.

В настоящее время быстрое развитие вычислительной математики, средств вычислительной техники, а также кибернетики химико-технологических процессов открывает принципиально новые возможности для решения проблемы идентификации кинетических моделей. В последние годы в этом направлении выполнен ряд теоретических работ¹⁻⁵ по разработке методов идентификации кинетических моделей и, в частности, статистических методов планирования и анализа эксперимента, проверки гипотез о механизме протекания химического процесса, установления адекватной кинетической модели и прецизионной оценки ее параметров. Применение этих методов открывает принципиально новые возможности при изучении сложных физико-химических явлений, так как позволяет эффективно выявлять более тонкие детали механизма протекания реакций, чем при использовании традиционных классических методов.

Для простоты изложения материала в данном обзоре общее решение проблемы идентификации кинетической модели сложной химической реакции разбито на четыре последовательных этапа.

1. Выдвигается система гипотез о механизме химической реакции, которые необходимо проверить на соответствие экспериментальным данным. Каждый гипотетический механизм представляется в виде некоторой совокупности элементарных реакций (стадий).

2. С использованием имеющейся информации об изучаемой химической системе, иногда с учетом выбранного критерия оптимальности, строится «стартовый» план эксперимента; по результатам измерений производится оценка констант в кинетических моделях.

3. Оценки кинетических констант, полученные на втором этапе, обычно не удовлетворяют необходимым требованиям точности и поэтому уточняются с помощью последовательных методов планирования эксперимента.

4. Строится последовательный план дискриминирующего эксперимента таким образом, чтобы при проведении каждого единичного наблюдения конкурирующие модели находились в некоторых критических условиях (с точки зрения их описательной силы). По степени согласия с опытными данными принимается решение о пригодности той или иной модели. В результате выбирается одна модель, лучше остальных отражающая эксперимент; механизм реакции, соответствующий этой кинетической модели, предполагается наиболее полно представляющим структуру химического превращения.

II. ПОСТРОЕНИЕ СИСТЕМЫ КОНКУРИРУЮЩИХ ГИПОТЕЗ О МЕХАНИЗМЕ РЕАКЦИИ

Роль первого этапа общей схемы решения проблемы идентификации кинетических моделей исключительно велика. Практика показывает, что экспериментатор, как правило, не может, исходя из интуитивных соображений, правильно выбрать достаточно полную систему конкурирующих гипотез, в особенности если речь идет о многостадийных реакциях.

При протекании многостадийной химической реакции происходит количественное и качественное изменение состава участвующих в реакции реагентов за счет перераспределения составляющих их атомов и отдельных неизменяемых в процессе реакции молекулярных структур. (Подобные атомы и молекулярные структуры называются атомными видами.) Все химические реагенты представимы в виде некоторых целочисленных ассоциаций атомных видов. Исходя из необходимости выполнения закона сохранения массы реагирующей системы можно разработать формализованный метод построения системы конкурирующих гипотез, базирующийся на стехиометрическом анализе реагирующей системы.

Задача стехиометрического анализа реагирующей системы формулируется следующим образом; предполагается, что известны (или постулируются) из теоретических предпосылок, а также из физических измерений исходные реагенты, промежуточные вещества и продукты реакций, которые называются молекулярными видами M_i^B ($i = 1, 2, \dots, N$). Требуется определить возможные химические реакции \mathcal{R}_r ($r = 1, 2, \dots, Q$), происходящие среди N молекулярных видов M_i^B и построить системы конкурирующих гипотез о механизме протекания реакции. Для решения этой задачи каждому из N химических реагентов M_i^B ставится в соответствие некоторый M -мерный вектор $\alpha_i = (\alpha_{i1}, \dots, \alpha_{ij}, \dots, \alpha_{iM})^T$, а каждой из Q реакций \mathcal{R}_r — N -мерный вектор $\beta_r = (\beta_{r1}, \dots, \beta_{ri}, \dots, \beta_{rN})^T$ ⁶⁻¹³.

$$M_i^B = \sum_{j=1}^M \alpha_{ij} M_j^A, \quad (1.1)$$

$$\mathcal{R}_r = \sum_{i=1}^N \beta_{ri} M_i^B, \quad (1.2)$$

где α_{ij} — целые неотрицательные числа, M_j^A ($j = 1, \dots, M$) — атомные виды, β_{ri} — рациональные числа, называемые стехиометрическими коэффициентами. Принимается, что любой молекулярный вид, имеющий положительный стехиометрический коэффициент, является продуктом реакции, а молекулярный вид с отрицательным стехиометрическим коэффициентом — исходным реагентом⁷.

$$\text{Две химические реакции } \mathcal{R}_r = \sum_{i=1}^N \beta_{ri} M_i^B = 0 \text{ и } \mathcal{R}'_r = \sum_{i=1}^N \gamma_{ri} M_i^B = 0$$

являются тождественными при любом рациональном γ . Отсюда следует, что для любой химической реакции определяющим является отношение стехиометрических коэффициентов. Реакции \mathcal{R}_l ($l = l_1, \dots, l_R$) из общего множества реакций $\{\mathcal{R}_r\}$ считаются независимыми, если не существует таких не равных нулю множителей γ_l ($l = l_1, \dots, l_R$), что

$$\sum_{l=l_1}^{l_R} \gamma_l \mathcal{R}_l = 0.$$

Очевидно, что общее число независимых реакций определяется рангом матрицы стехиометрических коэффициентов $B = \{\beta_{ri}\}$ для множества реакций $\{\mathcal{R}_r\}$.

Реакция \mathcal{R}_{r_1} , протекающая среди молекулярных видов M_1^B, \dots, M_k^B , определяется как стехиометрически простая, если между последними не может происходить никаких реакций из заданного множества $\{\mathcal{R}_r\}$.

На основе (1.1) и (1.2) выводится ряд полезных соотношений. Подставив (1.1) в систему уравнений (1.2), имеем:

$$\mathcal{R}_r = \sum_{i=1}^N \beta_{ri} \sum_{j=1}^M \alpha_{ij} M_j^A = \sum_{j=1}^M \left(\sum_{i=1}^N \beta_{ri} \alpha_{ij} \right) M_j^A = 0.$$

Следовательно,

$$\sum_{i=1}^N \beta_{ri} \alpha_{ij} = 0 \quad (1.3)$$

для всех $r = 1, \dots, Q; j = 1, \dots, M$. Или в матричном виде

$$BA = 0, \quad (1.4)$$

где $A = \{\alpha_{ij}\}$ — молекулярная матрица, $B = \{\beta_{ri}\}$ — матрица стехиометрических коэффициентов. Необходимо отметить, что эта система уравнений характеризует не что иное, как закон сохранения массы реагирующей системы.

Так как элементы молекулярной матрицы A известны, то любая из возможных реакций может быть получена в результате решения однородной линейной системы уравнений (1.3), которая называется основной системой уравнений (ОСУ); справедливо также и обратное утверждение — каждое решение ОСУ представляет собой возможную реакцию. Пусть ранг матрицы A равен u_A . Тогда максимальное число независимых реакций, которые могут протекать в рассматриваемой реагирующей системе, равно, согласно правилу Гиббса, $u_B = N - u_A$.

Для вычисления стехиометрических коэффициентов β_{ri} независимых реакций найдем в молекулярной матрице A какой-либо минор A ранга u_A и проведем соответствующую переиндексацию столбцов и строк матриц B и A , так чтобы индексы строк минора A изменялись от 1 до u_A включительно. Тогда уравнения (1.3) можно записать в виде:

$$\sum_{i=1}^{u_A} \beta_{ri} \alpha_{ij} = - \sum_{i=u_A+1}^N b_{ri} \alpha_{ij}, \quad r = 1, \dots, u_B; \quad j = 1, \dots, u_A, \quad (1.5)$$

где β_{ri} ($i = 1, \dots, u_A$) — искомые неизвестные, b_{ri} — стехиометрические коэффициенты, значения которых задаются произвольно. Любое решение системы (1.5) определяется с точностью до линейных комбинаций параметров b_{ri} . При этом решение называется базисным, если параметры b_{ri} выбираются следующим образом:

$$b_{ri} = \delta_{u_A+r, i} = \begin{cases} 0, & \text{если } i \neq u_A + r; \\ 1, & \text{если } i = u_A + r. \end{cases} \quad (1.6)$$

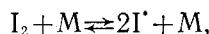
Тогда система (1.5) преобразуется к виду:

$$\sum_{i=1}^{u_A} \beta_{ri} \alpha_{ij} = \sum_{i=u_A+1}^N \delta_{u_A+r, i} \alpha_{ij}. \quad (1.7)$$

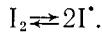
Нетрудно показать, что каждое базисное решение ОСУ определяет некоторую стехиометрически простую реакцию. Следовательно, если определены все базисные решения, то тем самым найдены все возможные в рассматриваемой реагирующей системе стехиометрически простые химические реакции. Для нахождения всех базисных решений необходимо решать ОСУ при ограничениях (1.6) для каждого неособенного

минора А ранга u_A . Естественно, что определенные таким путем стехиометрически простые реакции будут линейно зависимые. Из полученного множества стехиометрически простых реакций выбрасываются реакции абсурдные с точки зрения химической теории, и вводятся отдельные дополнительные элементарные реакции, которые не определяются с помощью указанной выше процедуры без ее существенного усложнения, но с очевидностью следуют из множества стехиометрически простых реакций (обычно количество таких дополнительных элементарных реакций не превышает 1—2).

Для пояснения приведем следующий пример. Пусть изучается реакционная система, в которой протекает следующая элементарная реакция



где I_2 , I^\cdot , M — соответственно молекула иода, радикал иода и нейтральная молекула. Описанная выше процедура определит стехиометрически простую реакцию



Располагая информацией о том, что радикал иода в газовой фазе образуется при соударении I_2 с нейтральной молекулой, можно записать первую из этих реакций на основе второй.

Необходимо отметить, что рассматриваемая процедура будет малоэффективной, если участниками элементарных химических реакций являются изомерные и стереоизомерные молекулы.

Установление возможных механизмов процесса на основе полученной совокупности элементарных реакций (стадий) сводится к выбору соответствующих множеств стехиометрических чисел¹⁴. Численные значения последних не произвольны, а подчинены ряду ограничений. Так, сумма уравнений элементарных химических стадий, умноженных на стехиометрические числа соответствующих стадий, должна дать суммарную брутто-реакцию (итоговое уравнение)^{15—18}. Следовательно, если в результате проведенного стехиометрического анализа оказалось, что имеется S элементарных стадий, включающих I промежуточных веществ и P реагентов и продуктов реакции, то (например, для одномаршрутной реакции) существуют S стехиометрических чисел v_r таких, которые удовлетворяют соотношениям (1.8) и (1.9):

$$\sum_{r=1}^S \beta_{rs} v_r = 0, \quad s = 1, \dots, I, \quad (1.8)$$

$$\sum_{r=1}^S \beta_{rs} v_r = b_{s'}, \quad s' = 1, \dots, P', \quad (1.9)$$

где $b_{s'}$ — стехиометрический коэффициент s' -того реагента или продукта реакции в итоговом уравнении.

Естественно, что любое множество стехиометрических чисел $\{v_r\}$, удовлетворяющих соотношениям (1.8) и (1.9), соответствует определенному механизму реакции. Так как обычно количество стехиометрических чисел больше, чем независимых линейных уравнений (1.8) и (1.9), то существуют значительные совокупности гипотетических одномаршрутных механизмов реакции. Необходимые ограничения могут быть получены с привлечением дополнительной априорной информации о реагирующей системе, например, такой как наличие определенной лимитирующей элементарной реакции или консективности определенных стадий.

дий¹⁹. Обобщения, получаемые для случая многомаршрутных реакций, очевидны.

Для каждого гипотетического механизма химической реакции строится соответствующая ему кинетическая модель. Методы их построения достаточно хорошо разработаны^{16, 18} и поэтому в данном обзоре не рассматриваются.

III. ОЦЕНКА КОНСТАНТ В КИНЕТИЧЕСКИХ МОДЕЛЯХ

Полученные кинетические модели, представляющие собой системы дифференциальных, алгебраических или дифференциально-алгебраических уравнений, содержат неизвестные константы, которые должны быть оценены по экспериментальным данным. В тех случаях, когда используются эмпирические модели или же теоретические кинетические модели линейны по параметрам, подробно разработанные методы линейного регрессионного анализа применяются для получения точечных и интервальных оценок параметров. Теория линейного регрессионного анализа излагается в монографиях²⁰⁻²⁴.

Однако подавляющее большинство моделей нелинейны по параметрам и процедура их оценки сильно усложняется. Рассмотрим ее более подробно. Пусть m кинетических моделей для каждой из m кинетических гипотез M_j ($j=1, \dots, m$) представимы в виде

$$\frac{d\eta^{(j)}}{dt} = \varphi(\mathbf{x}_u, \theta_j); \quad y_u^{(j)} = \eta_u^{(j)} + \varepsilon_u \text{ или } y_u^{(j)} = f(\mathbf{x}_u, \theta_j) + \varepsilon_u, \quad (2.1)$$

где $f(\mathbf{x}_u, \theta_j) = \eta^{(j)}(0) + \int_{t_0}^{\tau} \varphi(\bar{\mathbf{x}}_u, \theta_j) dt$, t — время протекания реакции,

θ_j — p_j -мерный вектор пространства неизвестных параметров, \mathbf{x}_u — q -мерный вектор пространства управляемых переменных, ε_u — ошибка воспроизведимости наблюдений, u — номер опыта. Для простоты ограничимся вначале анализом однооткликовых моделей; обобщения, необходимые для многооткликовых моделей, очевидны.

В кинетических исследованиях при оценке неизвестных параметров в моделях наиболее употребительным является метод максимального правдоподобия. При его использовании обычно дополнительно предполагается, что вид плотности распределения ошибок наблюдений известен. При этом неизвестные параметры в плотности распределения вероятностей могут быть оценены наряду с неизвестными параметрами моделей. Однако совместная оценка параметров в функции плотности распределения и в кинетической модели имеет смысл только для одно-, двух-, максимум — трехоткликовых моделей, так как в противном случае поиск оценок закорелированных друг с другом констант приходится вести в пространстве очень большой размерности, что делает решение задачи невозможным даже с использованием ЭВМ третьего поколения.

Положим теперь, что для модели M_j некоторым способом получены оценки θ_j^+ ; тогда соотношение (2.1) записывается в виде

$$e_u^{(j)} = y_u^{(j)} - f^{(j)}(\mathbf{x}_u, \theta_j^+), \quad (2.2)$$

где $e_u^{(j)}$ — оценки ошибки эксперимента ε_u для данных θ_j^+ и M_j .

Пусть проведено n опытов. Обозначим через $P(\varepsilon_u, \psi)$ плотность распределения вероятностей случайной величины, а через $P(\varepsilon, \psi)$ совмест-

ную плотность распределения случайного вектора $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$, где ψ — вектор пространства параметров плотности распределения, содержащий, в частности, величины математического ожидания и дисперсии воспроизводимости. Тогда функция правдоподобия выборки $L^{(j)}(\theta_j^+, \psi^+)$, получаемая в результате подстановки величины $e_u^{(j)}$ из соотношения (2.2) в выражение $P(\varepsilon, \psi)$, имеет вид

$$L^{(j)}(\theta_j^+, \psi^+) = P(e_j^{(j)}, \psi^+). \quad (2.3)$$

При независимых случайных величинах ε_u ($u=1, \dots, n$) имеет

$$L^{(j)}(\theta_j^+, \psi^+) = \prod_{u=1}^n P(e_u(\theta_j^+), \psi^+). \quad (2.4)$$

Согласно методу максимального правдоподобия, наилучшими оценками параметров являются оценки, приписывающие максимальные вероятности тем значениям наблюдений, которые получены в действительности. Следовательно, задача оценки параметров сводится к определению таких θ_j^\times и ψ^\times , которые удовлетворяют условию

$$L^{(j)}(\theta_j^\times, \psi^\times) = \max_{\theta_j^+, \psi^+} L^{(j)}(\theta_j^+, \psi^+). \quad (2.5)$$

Конкретный вид функции $L^{(j)}(\theta_j, \psi)$ зависит от плотности распределения вероятностей ε . В частности, если случайные величины ε_u ($u=1, \dots, n$) независимы и нормально распределены с нулевым средним и известными дисперсиями σ_u^2 , то

$$L^{(j)}(\theta_j^+, \psi^+) = \frac{1}{(2\pi)^{n/2} \left(\prod_{u=1}^n \sigma_u^2 \right)^{1/2}} \exp \left(-\frac{1}{2} \sum_{u=1}^n \frac{(y_u - f^{(j)}(x_u \theta_j^+))^2}{\sigma_u^2} \right). \quad (2.6)$$

Тогда оценки, получаемые с использованием принципа максимального правдоподобия, эквивалентны оценкам, минимизирующими взвешенную сумму квадратов (м. н. к. оценкам):

$$\Phi^{(j)}(\theta_j^\times) = \min_{\theta_j^+} \Phi^{(j)}(\theta_j^+) = \min_{\theta_j^+} \sum_{u=1}^n \frac{e_u^{(j)}(\theta_j^+)^2}{\sigma_u^2}. \quad (2.7)$$

При неизвестных, но равных дисперсиях наблюдений выражение (2.7) преобразуется к виду:

$$F^{(j)}(\theta_j^\times) = \min_{\theta_j^+} F^{(j)}(\theta_j^+) = \min_{\theta_j^+} \sum_{u=1}^n [e_u^{(j)}(\theta_j^+)]^2. \quad (2.8)$$

Отметим в заключение, что при нормально распределенных ошибках наблюдений оценки, полученные методами максимального правдоподобия и наименьших квадратов θ_j , совпадают. Они обладают рядом оптимальных свойств, если модели $f^{(j)}(\bar{x}, \bar{\theta}_j)$ линейны по параметрам, вне зависимости от того, распределены величины ε_u ($u=1, \dots, n$) нормально или нет^{24, 25}.

Для многооточных кинетических моделей функция правдоподобия выборки $L^{(j)}(\theta_j^+, \psi^+)$ при независимых нормально распределенных

ошибках наблюдений записывается следующим образом:

$$\begin{aligned} L^{(j)}(\theta_j^+, \Psi^+) &= \prod_{u=1}^n P(e_u^{(j)}(\theta_j^+, \Psi^+)) = \\ &= (2\pi)^{-Qn/2} \det(\Sigma)^{-n/2} \exp \left[-\frac{1}{2} \sum_{k=1}^Q \sum_{l=1}^Q \sigma^{kl} \sum_{u=1}^n e_{uk}^{(j)} e_{ul}^{(j)} \right] = \\ &= (2\pi)^{-Qn/2} \det(\Sigma)^{-n/2} \exp [-\frac{1}{2} \operatorname{Sp}(\Sigma^{-1} A(\theta_j^+))], \end{aligned} \quad (2.9)$$

где $e_u^{(j)}(\theta_j^+) = y_u - f^{(j)}(\mathbf{x}_u, \theta_j^+) = (e_{u1}^{(j)}(\theta_j^+), \dots, e_{uQ}^{(j)}(\theta_j^+))^T$ — Q -мерный вектор измерений; $f^{(j)}(\mathbf{x}_u, \theta_j^+)$ — Q -мерная вектор-функция — кинетическая модель химической реакции, соответствующая гипотезе M_j , $\Sigma = \{\sigma_{kl}\}$ — дисперсионно-ковариационная матрица измерений; T — индекс транспонирования, $A(\theta_j^+) = \sum_{u=1}^n e^{(j)}(\theta_j^+) e_u^{(j)}(\theta_j^+)^T$; $\Sigma^{-1} = \{\sigma^{kl}\}$.

Согласно методу максимального правдоподобия, оценки θ_j^X при известной Σ максимируют $L^{(j)}(\theta_j^+, \Psi^+)$, если они минимизируют величину

$$SS_1^{(j)}(\theta_j^X) = \operatorname{Sp}(\Sigma^{-1} A(\theta_j^+)) = \min_{\bar{\theta}_j^+} \operatorname{Sp}[\Sigma^{-1}(A(\theta_j^+))]. \quad (2.10)$$

Если матрица Σ^{-1} диагональна, то $\operatorname{Sp}(\Sigma^{-1} A(\theta_j^+))$ представляет собой взвешенную сумму квадратов остатков. Очевидно, что при $Q=1$ выражено (2.10) совпадает с (2.7).

Если же дисперсионно-ковариационная матрица ошибок наблюдений Σ неизвестна априори, то с использованием байесовского подхода к оценке параметров показано²⁶, что оценки максимального правдоподобия θ_j^X получаются минимизацией по параметрам $\det A(\theta_j^+)$:

$$SS_2(\theta_j^X) = \det A(\theta_j^+) = \min_{\bar{\theta}_j^+} \det A(\theta_j^+). \quad (2.11)$$

При этом отмечено²⁷, что для устранения возможных ошибок в выражение (2.11) включаются только отклики, линейно независимые между собой. Авторы²⁷ приводят необходимый математический аппарат для установления и исключения линейно зависимых откликов.

В ряде случаев, особенно при отличных от нормальных распределениях ошибок наблюдений, использование принципа максимального правдоподобия приводит к иным соотношениям, характеризующим степень близости расчетных и экспериментальных данных, нежели (2.7), (2.8), (2.10), (2.11). В частности, если ошибка распределена по Лапласу, то необходимо использовать для однооткликовых ситуаций метод наименьших модулей и соответственно соотношение

$$SS_3(\theta_j^X) = \sum_{u=1}^n |e_u(\theta_j^X)| = \min_{\bar{\theta}_j^+} \sum_{u=1}^n |e_u(\theta_j^+)|. \quad (2.12)$$

Очевидно, что нахождение оценок максимального правдоподобия еще не гарантирует получения адекватной модели, она проверяется дополнительно по статистическим критериям на соответствие экспериментальным данным^{28, 29}.

IV. ПРОВЕРКА АДЕКВАТНОСТИ КИНЕТИЧЕСКИХ МОДЕЛЕЙ

Основные методы проверки адекватности кинетических моделей, наиболее употребительные при проведении экспериментальных исследований, базируются в основном на приемах дисперсионного анализа и анализа остатков.

Дисперсионный анализ моделей используется для проведения сравнения между собой величин разностей $e_u^{(j)} = y_u - f_u^{(j)}(\mathbf{x}_u, \theta_j)$ с величинами ε_u , характеризующими ошибку измерений. Посредством подобного сравнения можно установить как общую адекватность модели, так и способы ее дальнейшего упрощения путем выбрасывания из модели статистически незначимых отдельных ее членов или кинетических параметров.

Для достижения этой цели вычисляются величины сумм квадратов $SS(1) = \sum_{u=1}^n y_u^2$ и $SS(2) = \sum_{u=1}^n \hat{e}_u^{(j)2} = \sum_{u=1}^n f_u^{(j)2}$, характеризующие соответственно разброс экспериментальных данных и разброс рассчитанных по модели значений откликов. Разности $e_u^{(j)} = y_u - f_u^{(j)}$, называемые остатками, представляют собой меру неспособности модели точно описать экспериментальные данные. Очевидно, что если испытываемая модель истинна, то остатки фактически есть оценки ошибки измерений. Ввиду этого общая мера $SS(3)$ несоответствия модели результатам эксперимента представима в виде:

$$SS(3) = \sum_{u=1}^n (y_u - f_u^{(j)})^2.$$

В литературе величина $SS(1)$ называется общей суммой квадратов, $SS(2)$ — суммой квадратов, обусловленной регрессией и $SS(3)$ — остаточной суммой квадратов. При этом

$$SS(1) = SS(2) + SS(3),$$

что является прямым следствием свойств метода наименьших квадратов.

При проведении дисперсионного анализа каждому отдельному измерению отклика приписывается одна степень свободы. Следовательно, при постановке n опытов для однооткликовой ситуации общая сумма квадратов $SS(1)$ обладает n степенями свободы: $SS(3)$ имеет $n - p_j$ и $SS(2)$ имеет p_j степеней свободы (p_j — число параметров в модели, с использованием оценок которых вычисляется $SS(2)$).

При реализации нескольких повторных измерений при одинаковых условиях эксперимента сумма квадратов $SS(4) = \sum_{u=1}^N (y_u - \bar{y})^2$, где

$\bar{y} = \frac{1}{N} \sum_{u=1}^N y_u$, содержит в себе всю необходимую информацию об ошибках измерений, так как она пропорциональна дисперсии $\sigma_{\text{оп}}^2$. Поэтому величина $SS(5)$

$$SS(5) = \sum_{u=1}^n (y_u - f_u^{(j)})^2 - \sum_{u=1}^n (y_u - \bar{y})^2,$$

равная разности между $SS(3)$ и $SS(4)$, определяет меру неспособности

испытываемой модели отражать результаты эксперимента, т. е. иначе **сумма квадратов $SS(5)$** характеризует степень неадекватности модели.

Если проведено по \tilde{n} повторных опытов при каждом из q различных условий проведения эксперимента, то сумма квадратов $SS(4)$ имеет $\tilde{n}-1$ степень свободы (одна степень свободы используется для оценки \bar{y}), в то время как сумма квадратов $SS(5)$ обладает $n-p_j-q(\tilde{n}-1)$ степенями свободы; последнее число определяется как разность между числом степеней свободы остаточной суммы квадратов и суммы квадратов ошибок измерений.

Суммы квадратов, обусловленные различными источниками, будучи поделенными на соответствующие числа степеней свободы, определяют средние квадраты. Поскольку средний квадрат ошибки измерений представляет собой оценку дисперсии ошибки измерений, то очевидно, что адекватность модели может определяться отношением величин среднего квадрата неадекватности модели к среднему квадрату ошибки измерений (F -статистика). Если это отношение велико (по крайней мере $\gg 1$), то имеются достаточно веские доводы в пользу того, что испытываемая модель не отражает результаты эксперимента.

Как известно, F -статистика определяется как отношение сумм квадратов независимых нормальных случайных величин, и испытание адекватности кинетических моделей фактически сводится к проверке справедливости неравенства

$$\frac{SS(5)/(n-p_j-q(\tilde{n}-1))}{SS(4)/(q(\tilde{n}-1))} < F_{\alpha}[n-p_j-q(\tilde{n}-1), q(\tilde{n}-1)]$$

при заранее заданном уровне значимости α . Величины F -статистики для заданных уровня значимости α числа степеней свободы в числителе и знаменателе затабулированы и приведены в руководствах^{24, 29}.

Для иллюстрации использования приемов дисперсионного анализа при установлении адекватности кинетических моделей рассмотрим результаты кинетического исследования реакции изомеризации пентана, проведенного в работах^{30–32}. Авторы выдвинули для испытания два механизма протекания реакции и соответственно получили две конкурирующие кинетические модели. По экспериментальным данным оценены параметры в моделях и проведен дисперсионный анализ для проверки адекватности последних. Основные результаты этой проверки приведены в табл. 1. Из данных табл. 1, в частности, следует, что первая модель в достаточной мере хорошо отражает эксперимент, в то время как вторая находится на «границе адекватности». Последнее не позволяет с уверенностью ни принять вторую модель, ни отвергнуть ее как неадекватную. Поэтому необходимо проведение дальнейших статистических испытаний второй модели.

Изложенная выше процедура проверки адекватности модели тем не менее еще не гарантирует статистической значимости отдельных ее членов. Следовательно, необходимо проведение более детального анализа составляющих модели, что достигается прямым обобщением методов дисперсионного анализа. Для этой цели дополнительно проводится разложение суммы квадратов, обусловленных регрессией, на ряд составляющих. При этом обычно для простоты анализа вычисляются суммы квадратов, обусловленные общей регрессионной моделью и упрощенной моделью с выброшенным одним или группой членов. Разность между этими двумя суммами квадратов представляет собой сумму квадратов, характеризующую влияние испытываемого компонента кинетической модели. Так как известно, что для адекватных моделей

ТАБЛИЦА 1
Результаты дисперсионного анализа кинетических моделей изомеризации пентана

Источник рассеяния	Сумма квадратов	Число степеней свободы	Средние квадраты	Отношение средних квадратов	F=Статистика	
					95%	97,5%
Первая модель						
Общий	763,54	24	—	—	—	—
Регрессия	754,29	4	188,6	—	—	—
Отклонение относительно уравнения регрессии (модели)	9,25	20	0,462	—	—	—
Неадекватность модели	8,58	17	0,528	5,9	8,7	14,3
Отклонение среди повторных наблюдений	0,27	3	0,090	—	—	—
Исключение одной константы модели	123,2	1	123,2	1370	246	985
Вторая модель						
Общий	28894,0	24	—	—	—	—
Регрессия	26896,8	4	6724,2	—	—	—
Отклонение относительно уравнения регрессии (модели)	1997,20	20	99,86	—	—	—
Неадекватность модели	1964,56	17	115,5	10,62	8,7	14,3
Отклонение среди повторных наблюдений	32,65	3	10,88	—	—	—

средний квадрат остатков характеризует дисперсию воспроизводимости, то условие

$$\overline{SS}(6)/\overline{SS}(5) > F_{\alpha}(1, n-p_i), \quad (3.1)$$

где $\overline{SS}(6)$ — средний квадрат, обусловленный испытываемым компонентом модели; $\overline{SS}(5)$ — средний квадрат остатков, определяет значимость испытываемого компонента модели. Очевидно, что подобные испытания следует проводить для всех членов (компонентов) кинетической модели³³.

Пример подобного анализа значимости одного из членов первой кинетической модели для реакции изомеризации пентана представлен в табл. 1. Из табл. 1 следует, что ввиду большого различия в величинах средних квадратов общей и упрощенной модели неравенство (3.1) выполняется: следовательно, испытываемый член является значимым и его необходимо сохранить в модели.

В заключение отметим, что результаты дисперсионного анализа позволяют сделать вывод лишь об общей пригодности модели или о значимости ее отдельных членов. Тем не менее неадекватность модели может иметь место, даже если критерии типа критерия Фишера указывают на соответствие модели экспериментальным данным. Поэтому требуется проведение более детального испытания моделей, которое осуществляется с помощью методов анализа остатков.

Как известно, остатки $e_u = y_u - f^{(j)}(\mathbf{x}_u, \boldsymbol{\theta}_j)$ как случайные величины обладают вполне определенной функцией распределения вероятностей. Причем в большинстве встречающихся на практике случаев она представляет собой нормальную функцию распределения с нулевым математическим ожиданием и дисперсией $\sigma_{\text{оп}}^2$. Очевидно, что установление адекватности кинетической модели только по какой-либо одной характеристике функции распределения (для критерия Фишера такой характеристикой является $\sigma_{\text{нед}}$) не может дать полной гарантии адекватности модели. Ввиду изложенного выше для комплексной проверки соответствия кинетической модели экспериментальным данным необходимо использовать или целиком функции распределения вероятностей, что требует постановки большого числа опытов, или же ее основные характеристики. Среди них необходимо отметить следующие^{24, 34-36}.

1. *Анализ нормальности распределения.* Для этой цели строятся гистограммы распределения нормированных частот появления остатков в зависимости от их численных значений. Подобные гистограммы должны приближенно отвечать нормальному закону распределения. При этом гипотеза о нормальности может быть проверена по различным статистическим критериям. Наряду с ней дополнительно проверяется также гипотеза о равенстве нулю математического ожидания выборочного распределения, что осуществляется с использованием как графических методов, так и методов линейного или нелинейного регрессионного анализа.

2. *Анализ отсутствия в остатках неслучайных составляющих.* Подобный анализ на первом этапе его проведения осуществляется с помощью построения и изучения графической зависимости остатков от предсказанных значений откликов, что дает возможность установить соответствие модели экспериментальным данным. Так, в частности, результаты дисперсионного анализа, проведенного для первой кинетической модели реакции изомеризации пентана, показывают полную адекватность модели, в то время как из графика остатков (рис. 1) с очевидностью следует, что общая адекватность модели достигается за счет сбалансирования остатков e_u при малых и больших величинах откликов при наличии их закономерного изменения с увеличением величин откликов. Следовательно, первую модель необходимо отвергнуть как неадекватную.

Анализ графической зависимости остатков от предсказанных значений откликов позволяет также получить дополнительную информацию о соблюдении ряда исходных статистических посылок по характеристике ошибок измерений и, в частности, по соблюдению условия постоянства дисперсии воспроизводимости в выбранной области экспериментирования (рис. 2). При этом, если, например, на таких графиках разброс величин остатков монотонно увеличивается или монотонно уменьшается, то очевидно, что дисперсия ошибки воспроизводимости является переменной величиной. В таком случае для оценки параметров моделей необходимо использовать метод наименьших квадратов с переменными весовыми коэффициентами, либо для сохранения постоянства дисперсии σ^2 провести преобразование зависимой переменной $\eta = f^{(j)}(\mathbf{x}, \boldsymbol{\theta}_j)$.

Аналогичным образом построение графических зависимостей остатков от управляемых переменных и от времени, а также графических зависимостей оценок параметров от управляемых переменных позволяет получить важную информацию о возможных скрытых неадекватностях в модели, которые устанавливаются путем обнаружения трендов остатков от уровней независимых управляемых переменных. Детальное исследование подобных зависимостей позволяет провести качественный

анализ соответствия конкурирующих моделей экспериментальным данным, а также наметить пути устранения возможных неадекватностей.

Установление адекватности многооткликовых моделей является более сложной процедурой и требует привлечения значительной по объему экспериментальной информации, ибо в противоположность случаю однотакливой модели требуется проверить гипотезу не о равенстве двух

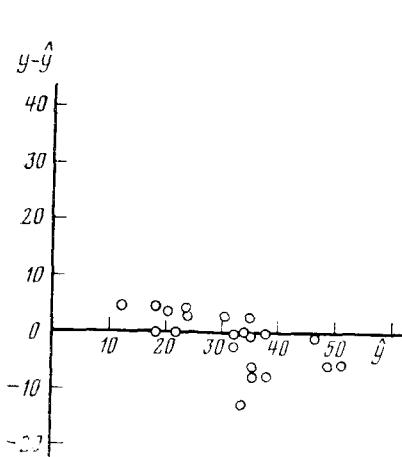


Рис. 1

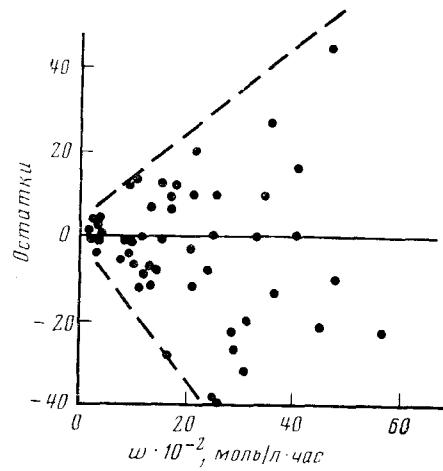


Рис. 2

Рис. 1. Зависимость остатков от предсказанного значения отклика

Рис. 2. Зависимость остатков от предсказанных значений скорости реакции w

дисперсий, а о равенстве двух ковариационных матриц Σ_1 и Σ . При этом $\Sigma_1 = A(\theta_j)/(n-p_j)$, где n — общее число измерений, p_j — минимальное число опытов, необходимое для оценки кинетических параметров,

$A(\theta_j) = \sum_{u=1}^n e_u(\theta_j) e_u(\theta_j)^T$; Σ — выборочная ковариационная матрица измерений

$\Sigma = \sum_{u=1}^{n_2} (\mathbf{y}_u - \bar{\mathbf{y}})(\mathbf{y}_u - \bar{\mathbf{y}})^T / (n_2 - 1) = A_2 / (n_2 - 1)$, где $\mathbf{y}_1, \dots, \mathbf{y}_{n_2}$ — повтор-

ная выборка объема n_2 , $\bar{\mathbf{y}} = \sum_{u=1}^{n_2} \mathbf{y}_u / n_2$.

Для проверки гипотезы

$$\Sigma_1 = \Sigma$$

используется статистика T_4 ²⁸:

$$T_4 = \frac{1}{2} \left\{ (n - p_j) \operatorname{Sp} \left[\Sigma_1 \left(\frac{A(\theta_j) + A_2}{n - p_j + n_2 - 1} \right)^{-1} - I \right]^2 + \right. \\ \left. + n_2 \operatorname{Sp} \left[\Sigma \left(\frac{A(\theta_j) + A_2}{n - p_j + n_2 - 1} \right)^{-1} - I \right]^2 \right\}, \quad (3.2)$$

где I — единичная матрица размерности $Q \times Q$, Q — размерность вектора наблюдений \mathbf{y}_u ($u=1, \dots, n_2$).

При предположении, что гипотеза верна, распределение статистики T_4 , выраженное как линейная комбинация χ^2 распределений с различными степенями свободы, для больших выборок имеет вид:

$$\begin{aligned} P(T_4 \leq x) = & P_f + m^{-1} \left[\frac{1}{12} \{ \tilde{\rho} Q (4Q^2 + 9Q + 4) - 12Q(Q+1)^2 + \right. \\ & + 4Q(Q^2 + 3Q + 4) \} P_{f+6} + \frac{1}{8} \{ -\tilde{\rho} Q (6Q^2 + 13Q + 9) + 16Q(Q+1)^2 + \\ & + 3Q(2Q^2 + 5Q + 5) \} P_{f+4} + \frac{1}{4} (2\tilde{\rho} + 2) Q(Q+1)^2 P_{f+2} + \\ & \left. + \frac{1}{24} (1 - \tilde{\rho}) Q(2Q^2 + 3Q - 1) P_f + O(n^{-2}) \right], \end{aligned} \quad (3.3)$$

где $\tilde{\rho} = \frac{n + n_2 - p_j - 1}{n - p_j} \cdot \frac{n + n_2 - p_j - 1}{n_2 - 1}$, $m = n + n_2 - p_j - 1$, $P_f = P(\chi_f^2 \leq x)$

с $f = Q(Q+1)$ степенями свободы.

Следовательно, если наблюдаемая величина T_4 , как мера отклонения от нулевой гипотезы, не значима согласно распределению (3.3), то принимается гипотеза о равенстве Σ_1 и Σ .

V. ПЛАНИРОВАНИЕ ПРЕЦИЗИОННЫХ ЭКСПЕРИМЕНТОВ

Оценка параметров кинетических моделей обычно должна удовлетворять необходимым требованиям точности для того, чтобы созданная модель обладала достаточными прогнозирующими возможностями, делающими ее пригодной для использования в целях проектирования, оптимизации и управления процессами. Поэтому, как правило, требуется проведение дополнительного последовательно планируемого эксперимента для уточнения параметров. Многочисленные исследования, посвященные этой проблеме^{37, 38}, ограничиваются в основном линейными моделями; существенно меньше внимания уделяется нелинейным. Поэтому вполне естественно, что задача планирования эксперимента для оценки параметров нелинейных моделей до сих пор изучена не до конца, и используемые для ее решения методы являются по сути дела прямым обобщением методов планирования эксперимента для линейных моделей.

При этом процедура построения стратегии последовательно экспериментирования будет определяться конкретными задачами кинетического исследования на данном этапе. Среди них важнейшими являются: 1) оценка с заданной точностью одного параметра или подвектора пространства параметров, что необходимо, например, для подтверждения механизма реакции или для характеристики отдельных стадий химического превращения; 2) минимизация коэффициентов корреляции между двумя параметрами или группой параметров. Величины коэффициентов корреляций оказывают заметное влияние на скорость сходимости итерационных методов оценки параметров и на экстраполяционные свойства получаемых моделей; 3) оценка вектора пространства параметров в моделях.

Для решения подобных задач оказывается достаточным проанализировать некоторые функционалы от информационной матрицы, а также ее отдельные элементы и подматрицы. Так, в качестве функционала, характеризующего точность оценок параметров для однооткликовых моделей, в³⁹ предложено использовать детерминант информационной матрицы $\Delta = \det X^T X$, квадратный корень величины которого обратно-

пропорционален объему $V_{\text{д.э.}}$ доверительного эллипсоида. Действительно,

$$V_{\text{д.э.}} = C \prod_{i=1}^p \left(\frac{\beta}{\lambda_i} \right)^{1/2} = C \frac{\beta^{p/2}}{(\det X^T X)^{1/2}}, \quad (4.1)$$

где λ_i ($i=1, \dots, p$) — собственные значения матрицы $X^T X$:

$$X_{n \times p} = \{x_{ul}\} = \left\{ \frac{\partial f(x_u, \theta)}{\partial \theta_l} \right\};$$

C — постоянная, зависящая от размерности p пространства параметров; β — постоянная, зависящая от дисперсии воспроизведимости σ^2 , числа степеней свободы $n-p$ и уровня значимости α .

Проблема планирования эксперимента при выборе критерия Δ будет состоять в последовательном определении таких условий эксперимента, при которых наблюдается максимальное уменьшение объема доверительной области (соответственно максимальное увеличение $\det X^T X$). Цикл планирования эксперимента продолжается до тех пор, пока параметры не достигнут желаемой степени точности. Естественно, что так как модель нелинейна по параметрам, то необходима их переоценка после каждого вновь поставленного опыта.

Имеется значительное число модификаций критерия Бокса — Лукаса для различных практических ситуаций, в частности:

а) для многооткликовых моделей при отсутствии априорной информации о параметрах⁴⁰:

$$\Delta = \sum_{i=1}^Q \sum_{j=1}^Q \sigma^{ij} (X^{(i)})^T X^{(j)}, \quad (4.2)$$

где Δ — информационная матрица Q -откликовой системы;

$$X^{(i)} = \{x_{ul}^{(i)}\} = \left\{ \frac{\partial f^{(i)}(x_u, \theta_l)}{\partial \theta_l} \right\} \quad (i = 1, \dots, Q; l = 1, \dots, p);$$

i, l, u — индексы, определяющие номер отклика, параметра и опыта;

б) для многооткликовых моделей при наличии априорной информации о параметрах⁴¹:

$$\Delta_k = \sum_{i=1}^Q \sum_{j=1}^Q \sigma^{ij} (X^{(i)})^T X^{(j)} + \Omega^{-1}, \quad (4.3)$$

где Ω — ковариационная матрица оценок параметров, известная заранее, до постановки уточняющих экспериментов;

в) для оценки заданного подмножества параметров⁴²:

$$\Delta_s = \det (Y_{11} - Y_{12} Y_{22}^{-1} Y_{21}), \quad (4.4)$$

$$X^T X = \begin{bmatrix} Y_{11} & Y_{21} \\ Y_{12} & Y_{22} \end{bmatrix},$$

где Y_{11} — подматрица размерности $s \times s$ матрицы $X^T X$ размерности $p \times p$; s — число оцениваемых параметров, причем $s < p$.

В работах^{43, 44} проведены дальнейшие обобщения для случаев непрерывных измерений, для одновременного планирования более чем m ($m > 1$) экспериментов, для планирования эксперимента, характеризующегося зависимостью от условий проведения опытов дисперсионной матрицей.

Критерии типа (4.2) — (4.4) носят название детерминантных. Бокс⁴⁵ проанализировал основные достоинства и недостатки детерминантных

критериев и обсудил многие их свойства. Практическое использование этих критериев, особенно при оценке параметров в многооткликовых моделях и при проведении небольшого по объему эксперимента, показало их исключительную эффективность даже для достаточно сложных случаев⁴⁶.

Однако структура кинетических моделей, как правило, такова, что оценки кинетических параметров сильно коррелированы друг с другом. Это ведет к тому, что функции, характеризующие степень совпадения экспериментальных и расчетных данных, обнаруживают в пространстве параметров в окрестности точки минимума наличие оврагов, которые затрудняют определение точечных оценок констант. Детерминантные критерии значительно уменьшают объем доверительного эллипсоида, не изменяя, однако, величин коэффициентов корреляций и, следовательно, не исправляя овражной ситуации.

В этом отношении критерий формы*, максимизирующий наименьшее собственное значение информационной матрицы $X^T X$, представляется более предпочтительным, так как стремится придать доверительной области большую сферичность посредством минимизации длины большой полуоси доверительного эллипса. Необходимо заметить также, что практическое использование критерия формы позволяет добиться и увеличения детермианта информационной матрицы, ибо

$$\det X^T X = \prod_{i=1}^p \lambda_i, \quad (4.5)$$

где λ_i ($i = 1, \dots, p$) — собственные значения $X^T X$; возрастание наименьшего собственного значения ведет к возрастанию $\det X^T X$. Следовательно, оба критерия не вполне независимы друг от друга.

Сравнение эффективности указанных выше критериев проводилось в работах^{47, 48}. На ряде примеров показано, что детерминантный критерий уменьшает объем доверительной области в основном за счет ее сжатия по малым полуосям гиперэллипса, которые по порядку величин существенно меньше больших полуосей. Критерий формы, напротив, стремится приблизить совместную доверительную область к сфере. Однако истинной сферичности достичь не удается, что обусловлено в основном различной чувствительностью модели к изменению параметров и ограниченными размерами области проведения эксперимента.

Для иллюстрации изложенного выше рассмотрим следующий пример⁴⁷. При изучении ферментативной реакции гидролиза сахарозы в¹⁴ получено кинетическое уравнение (4.6), в котором оценивались кинетические константы

$$R_c^b = - \frac{k_1 k_2 c_c}{1 + k_2 c_c + k_3 c_r + k_4 c_\phi}; \quad (4.6)$$

где R_c^b — скорость расходования сахарозы, c_c , c_r , c_ϕ — концентрации сахарозы, глюкозы и фруктозы соответственно. Их начальные концентрации использовались в качестве независимых управляемых переменных. На рис. 3 представлен стартовый план эксперимента, а также условия проведения 23-х дополнительно планируемых экспериментов по детерминантному критерию и критерию формы. Результаты обработки этих опытов представлены в табл. 2. Из табл. 2 следует, что критерий

* Критерий формы в литературе носит название *E*-критерия оптимальности плана, детерминантный критерий — *D*-критерия.

ТАБЛИЦА 2
Результаты планирования уточняющего эксперимента

Оцениваемые константы	Оценки констант	$\sqrt{\frac{\text{стандартные отклонения}}{\sigma_{k_i}^2}}$	Детерминантный критерий		Критерий формы	
			оценки констант	$\sqrt{\frac{\text{стандартные отклонения}}{\sigma_{k_i}^2}}$	оценки констант	$\sqrt{\frac{\text{стандартные отклонения}}{\sigma_{k_i}^2}}$
После стартового эксперимента				После 23-х уточняющих экспериментов		
k_1	4,61	1,16	2,89	0,09	2,97	0,06
k_2	92,29	24,41	57,47	1,94	58,94	1,56
k_3	16,04	4,22	9,23	0,54	9,24	1,07
k_4	33,89	8,58	18,88	0,86	21,24	1,50
Коэффициенты корреляционные матрицы оценок	$\begin{bmatrix} 1 & 0,99 & 0,98 & 0,99 \\ 0,99 & 1 & 0,97 & 0,99 \\ 0,98 & 0,97 & 1 & 0,37 \\ 0,99 & 0,99 & 0,37 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0,98 & 0,56 & 0,66 \\ 0,98 & 1 & 0,48 & 0,60 \\ 0,56 & 0,48 & 1 & 0,46 \\ 0,66 & 0,60 & 0,46 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0,96 & 0,43 & 0,41 \\ 0,96 & 1 & 0,06 & 0,02 \\ 0,43 & 0,06 & 1 & 0,01 \\ 0,11 & 0,02 & 0,01 & 1 \end{bmatrix}$			

формы позволяет точнее оценить константы k_1 и k_2 , чем k_3 и k_4 . По-видимому, это обусловлено тем, что среди совокупности оцениваемых констант константы k_1 и k_2 обладают большей параметрической чувствительностью. В то же время анализ парных корреляций оценок показывает, что корреляции существенно больше по абсолютной величине для детерминантного критерия. Следовательно, в целом при рассмотрении данного конкретного примера критерий формы представляется более предпочтительным.

Результаты сравнения этих критериев⁴⁷ не обладают общностью, и не следует ожидать получения аналогичных результатов во всех случаях. Результаты сильно зависят от типа кинетической модели, от конкретных численных оценок констант и их параметрической чувствительности, от протяженности интервалов варьирования различных управляемых переменных. Поэтому выбор того или иного критерия определяется целями кинетического исследования, а также общим балансом времени, отпущенном на эксперимент и проведение необходимых расчетов на ЭВМ.

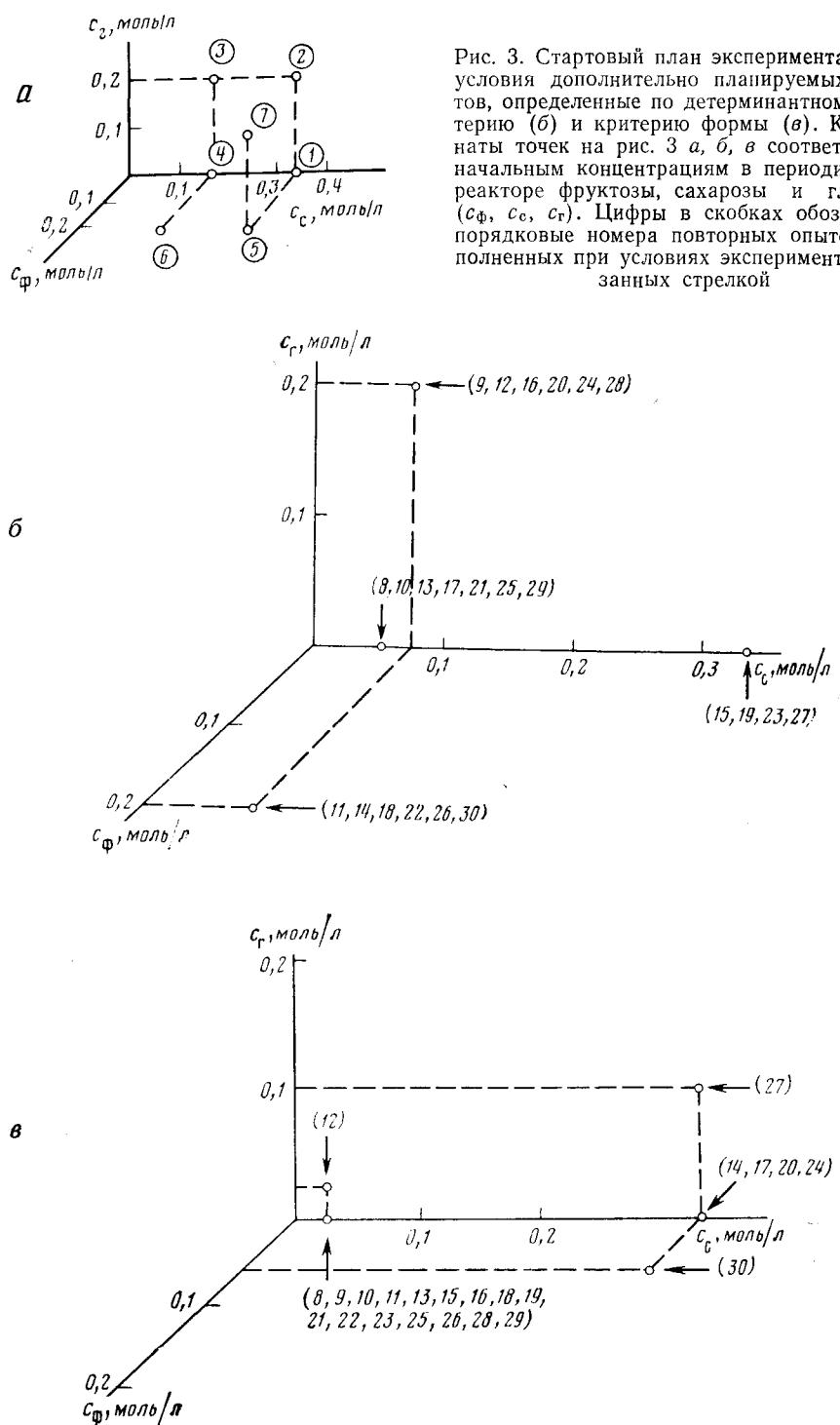
Естественно, что план эксперимента, оптимальный относительно одного критерия, может быть существенно неоптимальным относительно любого другого. Ввиду этого представляется целесообразным характеризовать оптимальные свойства планов, используя не один, а широкий класс основных критериев. Часто бывает выгодно, по аналогии со статистической теорией принятия решений, несколько поступиться оптимальностью плана относительно выбранного критерия, существенно улучшив при этом оптимальные свойства планов, оцененные по другим критериям. Следовательно, вместо использования частных критериев желательно построить один комплексный критерий.

Кифер⁴⁹ предложил использовать для оценки оптимальности планов семейство функционалов от информационной матрицы

$$\Phi_k(\xi) = \left\{ p^{-1} \operatorname{Sp} [\Delta^{-k}(\xi)] \right\}^{1/k} = \left\{ p \sum_{i=1}^p \lambda_i^{-k} \right\}^{1/k},$$

$$\Phi_0(\xi) = \lim_{k \rightarrow +0} \Phi_k(\xi) = \{\det \Delta(\xi)\},$$

$$\Phi_\infty(\xi) = \lim_{k \rightarrow +\infty} \Phi_k(\xi) = \max \{\lambda_i^{-1}\},$$
(4.7)



которые включают в себя три наиболее общих критерия оптимальности A , D и E , получаемые, в частности, при $k=1$, 0 или ∞ . Свойства семейства функционалов (4.7) детально проанализированы Кифером; многие из них он интерпретирует как различные характеристики размеров доверительного эллипсоида.

Имеются очевидные примеры, когда семейство функционалов Φ_k ($k=0, 1, \dots, \infty$) может не исправлять тех или иных недостатков процедур построения планов. Однако использование Φ_k безусловно полезно в большинстве случаев для характеристики свойств планов. При этом необходимо отметить, что эффективность процедуры планирования уточняющего (прецизионного) эксперимента для нелинейных по параметрам моделей будет зависеть также от исходных величин оцениваемых параметров, подлежащих уточнению, и, следовательно, от стартового плана эксперимента.

Как показывает практика, как правило, интуитивно выбранные в возможной области экспериментирования стартовые опыты приводят к неудовлетворительным — вырожденным или почти вырожденным — планам эксперимента; отметим, что почти вырожденный план эксперимента* в ряде случаев может привести к получению неверных результатов. Известно^{50, 51}, что стратегия стартового экспериментирования зависит как от конкретного типа модели, так и от численных значений ее параметров. Причем зависимость стратегии от первого фактора существенно более сильная, поэтому планирование стартовых экспериментов целесообразно даже в тех случаях, когда экспериментатор располагает априори весьма приближенными оценками параметров моделей. Естественно, что при полном отсутствии априорной информации о величинах параметров традиционный подход, при котором эксперимент планируется интуитивно, представляется единственным возможным.

Рассмотрим ситуацию, когда имеется некоторая априорная информация о параметрах, причем стоимость единичного наблюдения достаточно велика и поэтому требуется построить точный D -оптимальный план эксперимента для заранее заданного числа опытов N , ненамного превышающего число оцениваемых параметров p . В последнем случае теорема эквивалентности для точных локально оптимальных планов не выполняется, и оказывается справедливой только система неравенств

$$\det(\mathbf{x}, \varepsilon'_N) \leq \det(\mathbf{x}, \varepsilon_N) \leq \frac{Np}{N+1-p}, \quad (4.8)$$

где ε'_N , ε_N — соответственно планы эксперимента, минимизирующие дисперсию отклика $d(\mathbf{x}, \varepsilon_N)$ и максимизирующие детерминант информационной матрицы $M(\varepsilon_N)$. При этих условиях процедура построения точных локально оптимальных планов для однооточных моделей, в частности, может осуществляться следующим образом:

- 1) для испытываемой кинетической модели задаются p величин кинетических параметров θ ;
- 2) выбирается в области экспериментирования G план $\varepsilon_N^{(1)} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ с $\det M(\varepsilon_N) \neq 0$;
- 3) вычисляются условия проведения $N+1$ -го опыта, максимизирующие значение $d_N(\mathbf{x}, \varepsilon_N^{(1)})$;
- 4) формируется план $\varepsilon_{N+1}^{(1)} = \varepsilon_N^{(1)} \cup \mathbf{x}_{N+1}$;

* Под невырожденным планом эксперимента понимается план, детерминант информационной матрицы которого существенно отличен от нуля.

5) на основе плана $\varepsilon_{N+1}^{(1)}$ последовательным перебором среди $N+1$ -го N -точечного плана находится план $\varepsilon_N^{(1)}$, для которого детерминант информационной матрицы наибольший.

На s -том шаге построения точного плана неравенство

$$\det \{M(\varepsilon_N^{(s)})\} > \det \{M(\varepsilon_N^{(s-1)})\} \quad (4.9)$$

будет всегда выполняться, если

$$\sup_{x \in G} d(x, \varepsilon_N^{(s)}) = d(\varepsilon_N^{(s)}) > \frac{Np}{N+1-p}. \quad (4.10)$$

Процесс построения плана повторяется до тех пор, пока для заданной величины δ на некотором шаге не окажется, что

$$d(\varepsilon_N^{(k)}) \leq \frac{Np}{N+1-p} + \delta$$

для всех $k > s$.

Рассмотренная выше процедура показывает достаточно быструю сходимость к точному локально оптимальному плану. Так, для построения точного плана обычно оказывается достаточным проведение пяти-шести итераций.

Следует отметить, что наиболее ответственным моментом, определяющим оптимальность стратегий стартового (и в последующем уточняющего) эксперимента является определение числа констант в кинетической модели, допускающих оценку. В первом приближении оно равно рангу информационной матрицы $M(\varepsilon_N)$ плана ε_N ; этот ранг зависит как от численных значений θ , так и от плана эксперимента ε_N . При этом число констант, допускающих оценку, обычно меньше общего числа параметров в модели, так как часть из них — например, из-за использования приближений о квазистационарном протекании реакции или квазиравновесии быстрых стадий — определяется в виде некоторых комбинаций. В общее число параметров не входят параметры, статистическая незначимость которых установлена методами дисперсионного анализа при проверке адекватности кинетических моделей.

Поэтому при проведении последовательного планирования уточняющих экспериментов необходимо наряду с максимизацией выбранного критерия оптимальности плана максимизировать и ранг информационной матрицы, чтобы не потерять информацию о численных значениях параметров и не прийти к неверным выводам при выборе адекватной процессу кинетической модели.

VI. ДИСКРИМИНАЦИЯ МОДЕЛЕЙ

Обычно перед заключительным этапом установления вероятного механизма химической реакции оказывается, что несколько кинетических моделей соответствуют имеющимся экспериментальным данным, и поэтому возникает задача их дискриминации. Существует два основных метода проверки кинетических моделей — метод отношения правдоподобия и байесовский метод. Хотя они во многом подобны, тем не менее у них есть принципиальные отличия.

*Метод отношения правдоподобия*⁵² является эффективным методом проверки кинетических гипотез, особенно в случае, когда функция правдоподобия известна, т. е. известны плотность распределения ошибок измерений, а также численные значения параметров как для кинетических моделей, так и для функции распределения. Тогда величина отношения правдоподобия

$$u_{ij} = L_i / L_j \quad (5.1)$$

характеризует меру предпочтительности i -той модели по сравнению с j -той. Очевидно, если $u_{ij} > 1$, то $L_i > L_j$ и, следовательно, результаты эксперимента лучше отражает модель i , чем j . При $u_{ij} > 10$, а тем более при $u_{ij} > 100$, имеются веские доводы в пользу принятия i -той модели.

Однако чаще всего параметры θ_i и θ_j , от которых зависит величина отношения правдоподобия u_{ij} (θ_i , θ_j), неизвестны. Один из возможных приемов исключения «мешающих» параметров состоит в использовании оценок максимального правдоподобия в функции отношения правдоподобия

$$u_{ij} (\theta_i^X, \theta_j^X) = \max_{\theta_i} L_i (\theta_i) / \max_{\theta_j} L_j (\theta_j). \quad (5.2)$$

В этом случае $u_{ij} (\theta_i^X, \theta_j^X)$ представляет собой отношение правдоподобия двух индивидуально наилучших моделей. При неизвестных заранее параметрах, оцениваемых из эксперимента, интерпретация отношения правдоподобия и принятие решения в пользу той или иной модели становятся гораздо более произвольными. При одинаковом числе параметров в моделях величина отношения правдоподобия $u_{ij} (\theta_i, \theta_j)$, при которой отдается предпочтение i -той модели перед j -той, выбирается в пределах от 10 до 100.

Если в $u_{ij} (\theta_i, \theta_j)$ входят векторы θ_i и θ_j различной размерности, то испытания обычно выигрывает модель, содержащая большее число параметров. Этот факт должен обязательно учитываться при проверке кинетических гипотез.

Если число испытываемых гипотез больше двух ($m > 2$), то проводится их попарная проверка для всех возможных пар. Предпочтительней однако использование критерия обобщенного отношения правдоподобия⁵³:

$$u_j = \frac{L_j (\theta_j)}{\left(\prod_{k=1}^m L_k (\theta_k) \right)^{1/m}}. \quad (5.3)$$

В данном случае, аналогично изложенному выше, выигравшей испытания считается p -тая модель, если величины u_j ($j = 1, \dots, p-1, p+1, \dots, m$) не превышают значений 0,01—0,1; при этом предполагается, что $u_p > 0,1$.

Критерий (5.3) все же наиболее целесообразно использовать в последовательной процедуре установления наилучшей кинетической модели, сущность которой состоит в следующем. Эксперимент проводится последовательно, пока величина отношения правдоподобия одной из моделей, например l -той, не окажется в интервале 0,01—0,1; тогда l -тая модель выбрасывается из рассмотрения, и количество моделей сокращается на единицу. Испытания продолжаются до тех пор, пока не останется только одна модель, которая и принимается как наиболее соответствующая экспериментальным данным.

Байесовский подход к решению задачи дискриминации моделей основывается на теореме Байеса и субъективной интерпретации априорной информации. При байесовском подходе решение задачи дискриминации начинается с установления априорной плотности распределения вероятностей $P(M_j)$, $j = 1, \dots, m$, построенной для множества m конкурирующих гипотез M_j . Она задается исходя из предыдущих экспериментальных исследований или литературных данных и представляет собой субъективную интерпретацию исследователем имеющейся информа-

мации. Плотность распределения вероятностей $P(M_j)$ обычно дискретна ввиду конечности множества проверяемых гипотез, и ее вид не оказывает влияния на результат окончательного решения задачи испытания гипотез, хотя от ее выбора и будет зависеть как длительность процедуры дискrimинации, так и объем необходимого эксперимента.

Так как функцию потерь при принятии неверного решения построить затруднительно, обычно максимизируется априорная плотность распределения $P(M_j | \mathbf{y})$, которая определяется по результатам поставленного эксперимента \mathbf{y} и по априорной плотности распределения $P(M_j)$ исходя из теории Байеса:

$$P(M_j | \mathbf{y}) = \frac{P(\mathbf{y} | M_j) P(M_j)}{P(\mathbf{y})} = \gamma P(\mathbf{y} | M_j) P(M_j), \quad (5.4)$$

где $P(\mathbf{y} | M_j)$ — плотность распределения выборки, получения при предположении о справедливости гипотезы; $P(\mathbf{y})$ — частная плотность распределения выборки. Константа пропорциональности γ в (5.4) определяется из условия равенства единице суммы априорных вероятностей для всех моделей.

Если функция правдоподобия $P(\mathbf{y} | \theta_j, M_j)$ содержит «мешающие» параметры θ_j , то они должны быть исключены из рассмотрения путем усреднения функции $P(\mathbf{y} | \theta_j, M_j)$ по всем возможным значениям θ_j :

$$P(\mathbf{y} | M_j) = \gamma \int P(\mathbf{y} | \theta_j, M_j) P(\theta_j | M_j) d\theta_j, \quad (5.5)$$

где $P(\theta_j | M_j)$ — априорная плотность распределения параметров θ_j . Аналогично $P(M_j)$ выбирается по имеющейся исходной информации о величинах θ_j .

Отметим, что при использовании байесовского подхода к задаче дискrimинации параметры θ_j рассматриваются как случайные величины с некоторой плотностью распределения вероятностей.

Существует несколько способов практической реализации байесовского приближения для решения задачи дискrimинации моделей, нелинейных по параметрам. Наиболее общий из них основан на разложении нелинейных моделей в пространстве параметров в ряд Тейлора в окрестности значений полученных ранее оценок θ_j . При этом, ограничиваясь только первыми членами разложения имеем:

$$\mathbf{d}_j = X_j(\theta_j - \theta_j^+) + \boldsymbol{\varepsilon}, \quad (5.6)$$

где $\mathbf{d}_j = (d_{ju}, \dots, d_{jn})^T$; $d_{ju} = y_u - f^{(j)}(\mathbf{x}_u, \theta_j^+)$; $u = 1, \dots, n$ — индексы, характеризующие номер опыта. $\theta_j = (\theta_{j1}, \dots, \theta_{jp_j})$ — вектор пространства параметров размерности $(p_j \times 1)$; $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$ — вектор ошибок воспроизведимости; $X_j = \{x_{ul}^{(j)}\}$ — матрица размерности $n \times p_j$, для которой u, l -ый элемент можно представить в виде

$$x_{ul}^{(j)} = \left\{ \frac{\partial f^{(j)}(\mathbf{x}_u, \theta_j)}{\partial \theta_{jl}} \right\}_{\theta_j = \theta_j^+}. \quad (5.7)$$

Дополнительно предполагается, что 1) вектор ошибок воспроизведения $\boldsymbol{\varepsilon}$ есть нормальный случайный вектор с нулевым вектором математических ожиданий и ковариационной матрицей Σ , т. е. $\boldsymbol{\varepsilon} \sim N(0, \Sigma)$; 2) априорная плотность распределения вероятностей θ_j также является нормальной с вектором средних $\tilde{\theta}_j$ и известной ковариационной матрицей Ω_j размерности $p_j \times p_j$.

Нетрудно видеть, что при сделанных допущениях случайный вектор \mathbf{d}_j распределен нормально с $E \mathbf{d}_j = X_j(\theta_j - \theta_j^+)$ и ковариационной матрицей $D \mathbf{d}_j = X_j \Omega_j X_j^T + \Sigma$; здесь E — символ математического ожидания. Тогда выражение (5.5) примет вид:

$$P(\mathbf{y} | M_j) = \frac{\exp\left\{-\frac{1}{2}[\mathbf{d}_j - X_j(\tilde{\theta}_j - \theta_j^+)]^T(X_j \Omega_j X_j^T + \Sigma)^{-1}[\mathbf{d}_j - X_j(\tilde{\theta}_j - \theta_j^+)]\right\}}{(2\pi)^{n/2} [\det(X_j \Omega_j X_j^T + \Sigma)]^{1/2}}. \quad (5.8)$$

Хотя вывод выражения (5.8) проведен без использования (5.5), однако он вполне ему эквивалентен⁵².

Апостериорные вероятности принятия гипотез вычисляются по формуле (5.4). Следовательно, имеем

$$P(M_j | \mathbf{y}) = \frac{P(M_j) P(\mathbf{y} | M_j)}{\sum_{k=1}^m P(M_k) P(\mathbf{y} | M_k)}. \quad (5.9)$$

На определенном шаге экспериментирования принимается, что l -тая модель выиграла испытания среди совокупности конкурирующих, если $P(M_l | \mathbf{y})$ принимает значения, достаточно близкие к единице.

Аналогично байесовский метод принятия решений может быть применен для оценки кинетических параметров. В этом случае апостериорное распределение параметров θ также определяется по теореме Байеса

$$P(\theta_j | \mathbf{y}) = \frac{P(\mathbf{y} | \theta_j) P(\theta_j)}{\int_{\Omega} P(\mathbf{y} | \theta_j) P(\theta_j) d\theta_j} = \frac{P(\mathbf{y} | \theta_j) P(\theta_j)}{P(\mathbf{y})}, \quad (5.10)$$

где $P(\mathbf{y})$ и $P(\theta_j)$ — частные плотности распределения случайных векторов \mathbf{y} и θ_j . Как и ранее, $P(\theta_j)$ содержит в себе исходные сведения о величинах θ_j . Если $P(\mathbf{y} | \theta_j)$ и $P(\theta_j)$ — нормальные плотности распределения вероятностей, то $P(\theta_j | \mathbf{y})$ — условная апостериорная плотность распределения параметров θ_j — также нормальна с

$$E(\theta_j) = (X_j^T \Sigma^{-1} X_j + \Omega_j^{-1})(X_j^T \Sigma^{-1} X_j \theta_j^+ + \Omega_j^{-1} \tilde{\theta}_j) \quad (5.11)$$

и с ковариационной матрицей

$$D(\theta_j) = [(X_j^T \Sigma^{-1} X_j + \Omega_j^{-1})]^{-1}.$$

При кинетических исследованиях сложных химических систем часто необходимо для подтверждения механизма реакции оценить с заданной точностью только часть констант из общей совокупности, например чтобы затем сравнить их величины со значениями констант, полученными по ранее проведенным независимым измерениям. Следовательно, требуется оценить только некоторый подвектор вектора параметров θ_j , апостериорная плотность распределения которого есть частная апостериорная плотность распределения подвектора интересующих нас в данном случае параметров. Последняя, при нормальной $P(\theta_j | \mathbf{y})$, также нормальна с вектором математических ожиданий, равным соответствующему подвектору выделенных параметров вектора $E(\theta_j)$, и с ковариационной матрицей, которая составлена из элементов $D(\theta_j)$, стоящих на пересечении столбцов и строк соответствующих выделенных параметров.

Заметим также, что использование апостериорной плотности распределения кинетических параметров для характеристики точности их оценок несет в себе несравненно больше информации о параметрах,

чем точечные и интервальные оценки, что является характерным для байесовского подхода преимуществом.

Рассмотренная выше байесовская процедура проверки кинетических гипотез, предусматривающая линеаризацию кинетических моделей по параметрам, корректна только тогда, когда подвергаемые испытаниям на соответствие результатам эксперимента модели линейны по параметрам и правильны основные статистические предпосылки относительно характера плотности распределения ошибок измерений и справедливости использования методов регрессионного анализа. Поэтому применение подобных процедур для дискриминации существенно нелинейных по параметрам моделей правомерно лишь в случае достаточно хорошей аппроксимации нелинейных моделей линеаризованными.

Для повышения надежности байесовских методов, предусматривающих линеаризацию кинетических моделей, необходимо в первую очередь улучшение самой процедуры линеаризации. Это может быть достигнуто, в частности, подходящим выбором оценок θ_j^+ , в окрестности которых модели аппроксимируются усеченным рядом Тейлора. В качестве таких оценок обычно целесообразно использовать оценки максимального правдоподобия или оценки, соответственно равные вектору математических ожиданий апостериорной плотности распределения параметров.

Другой способ улучшения процедуры линеаризации заключается в предварительной репараметризации моделей и приведении их к виду, который уже не содержит существенных нелинейностей. Тем самым преобразованные кинетические модели более удобны для последующей линеаризации. Так, в работе⁵⁴ показано на ряде примеров, что непосредственное применение байесовских методов дискриминации к непараметризованным моделям приводит к противоречивым результатам, в то время как предварительно проведенная репараметризация конкурирующих моделей устраняет возможные ошибки. При этом предлагаемые алгоритмы репараметризации кинетических моделей во многом аналогичны алгоритмам Бокса и Кокса⁵⁵, Бокса и Тидвелла⁵⁶, разработанным для замены зависимых и независимых переменных в регрессионных уравнениях.

Следует отметить, что репараметризация позволяет заметно повысить не только надежность байесовских процедур дискриминации моделей, но и точность расчета дисперсионных матриц и соответственно доверительных интервалов и доверительных областей для оценок параметров⁵⁷.

Метод проверки кинетических гипотез, исключающий процедуру линеаризации моделей и использующий в процессе принятия решений кинетические модели как таковые, предложен в⁵⁸. Согласно этому методу для каждой конкурирующей модели непрерывная область параметрического пространства, в которой осуществляется поиск наилучших оценок параметров модели, покрывается конечным множеством дискретных точек, равноотстоящих друг от друга. Каждой l_j -той параметрической точке θ_j приписывается некоторая априорная вероятность $P(\theta_j | M_j)$, а выражение для вычисления апостериорных вероятностей принятия гипотез (5.9) заменяется его дискретным аналогом:

$$P(M_j | \mathbf{y}) = \frac{P(M_j) \sum_{l=1}^{s_j} P(\theta_{l_j} | M_j) P(\mathbf{y} | \theta_{l_j}, M_j)}{\sum_{i=1}^m P(M_i) \sum_{l=1}^{s_i} P(\theta_{l_i} | M_i) P(\mathbf{y} | \theta_{l_i}, M_i)}, \quad (5.12)$$

где s_i — общее число параметрических точек θ_{l_i} , покрывающих область параметрического пространства для i -той модели.

По выборке \mathbf{y} и априорным вероятностям $P(\theta_{l_j} | M_j)$ оценивают величины апостериорных вероятностей

$$P(\theta_{l_j} | \mathbf{y}) = \frac{P(\theta_{l_j} | M_j) P(\mathbf{y} | \theta_{l_j}, M_j)}{\sum_{i=1}^m \sum_{l=1}^{s_i} P(\theta_{l_i} | M_i) P(\mathbf{y} | \theta_{l_i}, M_i)}. \quad (5.13)$$

На следующем шаге испытаний конкурирующих кинетических моделей, т. е. перед постановкой новой серии опытов и получением новой выборки \mathbf{y} , апостериорные вероятности $P(\theta_{l_j} | \mathbf{y})$ и $P(M_j | \mathbf{y})$ считаются априорными и цикл исследований повторяют до тех пор, пока апостериорная вероятность $P(M_k | \mathbf{y})$ некоторой k -той модели не будет близка к единице; k -тая модель принимается как наиболее соответствующая результатам эксперимента.

Практическая реализация предложенного в⁵⁸ метода предполагает обязательное использование современных ЭВМ с большим объемом памяти и быстродействием. Действительно, если, например, кинетическая модель содержит 5 параметров и каждый параметр имеет 10 дискретных значений, то в памяти ЭВМ необходимо хранить 10^5 значений функции правдоподобия. Следовательно, метод⁵⁸ эффективен — из-за так называемого «проклятия размерности» — только в тех случаях, когда число параметров в конкурирующих моделях невелико и их предварительные оценки известны с достаточно высокой точностью, что позволяет выбрать интервал возможного изменения параметров достаточно узким и соответственно ограничиться при проверке гипотез небольшим числом дискретных значений кинетических параметров.

В заключение рассмотрим один полезный метод дискриминации, использующий в своей основе χ^2 -критерий⁵⁹; этот метод по существу нельзя отнести ни к байесовским методам, ни к методам отношения правдоподобия. Основная идея рассматриваемого метода состоит в том, что отношение взвешенной суммы квадратов остатков к соответствующему числу степеней свободы для однооткликовых моделей является несмешенной оценкой ошибки воспроизводимости только для истинной кинетической модели; для всех остальных моделей эта оценка смешена. Метод проверяет однородность оценок дисперсий воспроизводимости для всех конкурирующих моделей с использованием распределения Бартлетта

$$\chi^2_c = \frac{(\ln \bar{s})^2 \sum_{i=1}^m p_{s_i} - \sum_{i=1}^m p_{s_i} \ln s_i^2}{1 + \frac{1}{3(m-1)} \left[\sum_{i=1}^m \frac{1}{p_{s_i}} - \frac{m}{\sum_{i=1}^m p_{s_i}} \right]}, \quad (5.14)$$

где m — число конкурирующих моделей; s_i^2 — оценка дисперсий воспроизводимости, вычисленная по i -той модели; p_{s_i} — число степеней свободы оценки; \bar{s}^2 — усредненная оценка ошибки воспроизводимости. Если при проверке m гипотез окажется, что χ^2_c превышает табличное значение, то модель, дающая наибольшую величину дисперсии ошибки

воспроизводимости, отбрасывается. Количество гипотез сокращается на единицу, а χ^2_e вычисляется вновь.

При применении этого критерия для решения задач дискриминации гипотез следует иметь в виду, что χ^2_e -критерий весьма чувствителен к отклонению данных от нормальности и требует статистической независимости оценок дисперсии, относительно которых проверяется гипотеза однородности. Установить, насколько точно выполняются подобные предположения, часто не представляется возможным. Однако практические примеры использования χ^2_e -критерия в химических исследованиях показали его эффективность и надежность⁵⁹.

VII. ПЛАНИРОВАНИЕ ДИСКРИМИНИРУЮЩИХ ЭКСПЕРИМЕНТОВ

В изложенных выше процедурах проверки кинетических гипотез использованы для дискриминации моделей результаты некоторого эксперимента, в частности пассивного. Однако эффективность процедур проверки значительно возрастает, если эксперимент проводится активно. При этом задача планирования дискриминирующих экспериментов, в отличие от задачи планирования прецизионных экспериментов, характеризуется отсутствием единобразия применяемых статистических критериев. Однако их можно подразделить на две категории: к первой относятся критерии, опирающиеся на различные меры информации, а ко второй — эмпирические критерии^{60—63}.

В кинетических экспериментальных исследованиях среди критериев первой категории наиболее употребительным является энтропийный, использующий в качестве меры информации энтропию S . Она определяется для всех $P(M_j) \neq 0$ ($j=1, \dots, m$) следующим образом:

$$S = - \sum_{j=1}^m P(M_j) \ln P(M_j), \quad (6.1)$$

где $P(M_j)$ — априорная вероятность принятия M_j -той гипотезы. Если $P(M_i) = 0$ для некоторого i , то произведение $P(M_i) \ln P(M_i)$ приравнивается нулю.

Очевидно, что если для k -той модели $P(M_k) = 1$, то $P(M_j) = 0$ для $j=1, \dots, k-1, k+1, \dots, m$. Тогда энтропия S принимает наименьшее из возможных значений, равное нулю. Напротив, при равных $P(M_j) = 1/m$ ($j=1, \dots, m$) значение S достигает максимума. Следовательно, S характеризует меру полученной информации об изучаемой системе и увеличение информации сопровождается обязательным понижением энтропии.

Так как при проведении исследований желательно максимально уменьшить степень неопределенности представлений о химической системе, то следует стремиться перейти от неинформативной ситуации ($P(M_j) = 1/m$, $j=1, \dots, m$; информация скучна и отдать предпочтение какой-либо модели невозможно) к наиболее информативной ($P(M_k) = 1$, $j=1, \dots, k-1, k+1, \dots, m$; информации достаточно, чтобы была доказана — в «вероятностном смысле» — справедливость k -той модели). Для этой цели в области условий проведения эксперимента G необходимо реализовать только такие опыты, которые максимально понижают величину энтропии S .

Для установления условий проведения таких опытов Бокс и Хилл⁶⁴ предложили для однооткликовых моделей критерий $D(x)$, определяющий верхнюю границу ожидаемого уменьшения энтропии при планируе-

мой постановке опыта в точке \mathbf{x} области экспериментирования G :

$$D(\mathbf{x}) = \frac{1}{2} \sum_{j=1}^m \sum_{i=j+1}^m P(M_i) P(M_j) \left\{ \frac{(\sigma_i^2 - \sigma_j^2)}{(\sigma^2 + \sigma_i^2)(\sigma^2 + \sigma_j^2)} \times \right. \\ \left. \times [\hat{\eta}_i(\mathbf{x}) - \hat{\eta}_j(\mathbf{x})]^2 \left(\frac{1}{\sigma^2 + \sigma_i^2} + \frac{1}{\sigma^2 + \sigma_j^2} \right) \right\}, \quad (6.2)$$

где \mathbf{x} — вектор пространства независимых управляемых переменных; $\tilde{\eta}_i(\mathbf{x})$ — значение отклика в точке \mathbf{x} , предсказанное по i -той модели; σ^2 — дисперсия ошибки воспроизведимости; σ_i^2 — дисперсия предсказываемого значения отклика i .

По $D(\mathbf{x})$ -критерию наиболее выгоден такой эксперимент \mathbf{x}^* , который максимизирует $D(\mathbf{x})$, т. е. выполняется условие

$$D(\mathbf{x}^*) = \max_{\mathbf{x} \in G} D(\mathbf{x}). \quad (6.3)$$

Энтропийный критерий⁶⁰ обобщен в⁵ применительно к многооткликовым моделям. Впоследствии Рейлли⁵² вновь усовершенствовал критерий $D(\mathbf{x})$, дав алгоритм для непосредственного вычисления величины ожидаемого уменьшения энтропии S .

Другой важной мерой информации, используемой при дискриминации моделей, является функция обобщенного отношения правдоподобия. В этом случае минимуму информации о системе гипотез соответствует равенство величин обобщенного отношения правдоподобия, а максимуму информации — максимальная разность между ними. Отсюда непосредственно следует, что для достижения в процессе принятия решения наиболее благоприятной информативной ситуации необходимо планировать эксперимент таким образом, чтобы от опыта к опыту максимизировать ожидаемый средний прирост суммы величин логарифма обобщенного отношения правдоподобия^{12, 53}.

$$\Delta L(\mathbf{x}_{N+1}) = \sum_{j=1}^m \int_{-\infty}^{+\infty} [\ln u_j(N+1) - \ln u_j(N)] P_j(y_{N+1} | \mathbf{x}) d y_{N+1} = \\ = \frac{1}{2m} \left\{ \sum_{j=1}^m \sum_{k=1}^m \text{Sp}(\Sigma \vdash S_j)(\Sigma + S_k)^{-1} - m \text{Sp} I + \right. \\ \left. + \sum_{j=1}^m \sum_{k=1}^m (\hat{\eta}_{N+1}^{(k)} - \hat{\eta}_{N+1}^{(j)})^T [(\Sigma + S_j)^{-1} (\hat{\eta}_{N+1}^{(k)} - \hat{\eta}_{N+1}^{(j)})] \right\}, \quad (6.4)$$

где $D_j(N)$ — ковариационная матрица оценок констант j -той модели, вычисленная по N опытам; $P_j(y_{N+1} | \mathbf{x}_{N+1})$ — плотность распределения, предсказываемого по j -той модели $N+1$ -го результата измерения отклика в точке \mathbf{x}_{N+1} ; $L_j(N)$ — обобщенная функция правдоподобия, равная

$$L_j(N) = \frac{\det \Sigma^{-N/2} \det D_j(N)^{1/2}}{(2\pi)^{nN/2-p_j/2}} \cdot \exp \left\{ -\frac{1}{2} \left[\sum_{r=1}^n \sum_{s=1}^n \sigma^{rs} \sum_{l=1}^N \times \right. \right. \\ \left. \left. \times (y_{rl}^{(j)} - \hat{\eta}_{rl}^{(j)})(y_{sl}^{(j)} - \hat{\eta}_{sl}^{(j)}) \right] \right\};$$

$$u_j(N) = \frac{L_j(N)}{\left(\prod_{k=1}^m L_k(N) \right)^{1/m}};$$

$S_j = X_j D_j (N) X_j^T$; X_j — матрица размерности $n \times p_j$ с общим элементом вида $x_{ri}^{(j)} = \partial \hat{\eta}_{r,N+1}^{(j)} / \partial \theta_i^{(j)}$; $\hat{\eta}^{(j)}$ — вектор размерности $n \times 1$ прогнозируемых значений отклика по j -той модели; $\Sigma = \{\sigma_{ij}\}$ — ковариационная матрица измерений; $\Sigma^{-1} = \{\sigma^{rs}\}$; индексы j, i, r характеризуют соответственно номера модели, параметра и отклика.

Применение критерия (6.4) при решении ряда практических задач химической кинетики показало его эффективность и надежность^{64, 65}. Так, в частности, использование (6.4) при проверке на соответствие эксперименту двух конкурирующих гипотез о механизме реакции диспропорционирования толуола на цеолитных катализаторах показало, что достаточно поставить всего два последовательно планируемых эксперимента, чтобы испытания выиграла единственная гипотеза. Подобная картина часто имеет место при дискриминации многооткликовых моделей, в то время как для однооткликовых необходимое число дискриминирующих экспериментов обычно существенно больше.

На рис. 4 графически представлена функция обобщенного отношения правдоподобия ΔL , полученная при расчете условий проведения первого дискриминирующего опыта. Многоэкстремальный и овражный характер функции ΔL в зависимости от управляемых переменных указывает на возникающие вычислительные трудности при решении задач дискриминации многооткликовых нелинейных по параметрам моделей. Следует отметить, что дискриминирующие эксперименты обычно расположены на границе области возможного экспериментирования.

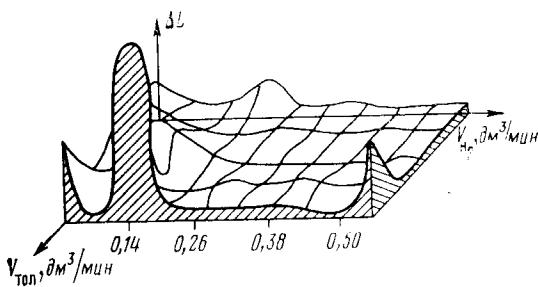


Рис. 4. Поверхность отклика функции отношения правдоподобия ΔL в зависимости от управляемых переменных — объемных скоростей подачи водорода и толуола; объемная скорость подачи бензола постоянна и равна $0,05 \text{ дм}^3/\text{мин}$

Известны и другие критерии планирования дискриминирующих экспериментов, опирающиеся на иные меры информации. Так, в⁶⁶ предложено использовать куллбаковскую меру, а в⁶⁷ — ожидаемое увеличение наибольшей апостериорной вероятности принятия моделей. Сравнение критериев и оценка эффективности их работы проведено в работе⁶⁸. Однако в практике кинетических исследований критерии, предложенные в^{66, 67}, пока не использовались.

Все рассмотренные выше критерии⁶⁹ позволяют выявить опыты в области экспериментирования, которые позволяют поставить конкурирующие модели в критические условия с точки зрения их описательной силы. При этом, как правило, учитывается неопределенность в предсказаниях откликов; предпочтение может быть отдано условиям проведения эксперимента, характеризуемым небольшими расхождениями в предсказанных с малыми коридорами ошибок значениях откликов по сравнению с условиями, сопровождающимися большими расхождениями в предсказанных с большими коридорами ошибок значениях откликов.

Практически дисперсии прогнозируемых значений откликов вычисляются с большими трудностями и поэтому при построении эмпириче-

ских критериев ими, как правило, пренебрегают. Наиболее употребительными являются критерий $C(x)$ ⁶⁷

$$C(x) = \sum_{i=1}^m \sum_{j=i+1}^m |\hat{\eta}_i(x) - \hat{\eta}_j(x)| \quad (6.5)$$

и критерий $B(x)$ ⁵⁸

$$B(x) = \sum_{i=1}^m \sum_{j=i+1}^m |P(M_i) + P(M_j)| \cdot |\hat{\eta}_i(x) - \hat{\eta}_j(x)|. \quad (6.6)$$

Обычно при применении критериев (6.5) и (6.6) дискриминирующие опыты располагаются только в одной, достаточно ограниченной подобласти области экспериментирования G . Поэтому в⁷⁰ предложен комплексный критерий, предусматривающий наряду с максимизацией $C(x)$ максимизацию расстояния новой планируемой дискриминирующей точки от всех остальных, ранее найденных точек плана. Практическая реализация этого принципа осуществлена в работе⁷⁰.

К настоящему времени имеется уже значительное число публикаций, в которых методы планирования и анализа кинетического прецизионного и дискриминирующего эксперимента нашли практическое применение. В большинстве из них описаны реальные практические ситуации^{61-63, 71-80}, в других — лишь моделируемые на ЭВМ. В основном в этих публикациях изложены результаты сравнения различных критериев и проведена оценка эффективности их работы⁷⁶⁻⁸⁰.

ЛИТЕРАТУРА

1. Н. М. Эмануэль, Д. Г. Кнопре, Курс химической кинетики, «Высшая школа», М., 1974.
2. Г. К. Боресков, Вестн. АН СССР, 5, 47 (1965).
3. В. В. Кафаров, Методы кибернетики в химии и химической технологии, «Химия», М., 1976.
4. В. В. Налимов, Теория эксперимента, «Наука», М., 1971.
5. Д. М. Химмельблau, Анализ процессов статистическими методами, «Мир», М., 1973.
6. В. Н. Писаренко, А. Г. Погорелов, Планирование кинетических исследований, «Наука», М., 1969.
7. R. Aris, R. Mah, Ind. Engng. Chem. Fund., 2, 90 (1963).
8. A. Pethe, Chem. Engng. Sci., 20, 791 (1965).
9. G. J. Pings, E. B. Nebecker, Ind. Engng. Chem. Fund., 4, 376 (1965).
10. Д. А. Франк-Каменецкий, Успехи химии, 10, 373 (1941).
11. В. В. Кафаров, В. Н. Писаренко, Л. А. Масчева, О. Н. Смирнова, ДАН СССР, 216, 342 (1974).
12. В. В. Кафаров, В. Н. Писаренко, Т. Б. Жукова, Л. А. Масчева, сб. докладов IV Всес. конф. по планированию и автоматизации эксперимента в научных исследованиях, «Сов. радио», М., 1974.
13. N. N. Semenoff, J. Chem. Phys., 7, 683 (1939).
14. J. Horiuti, J. Res. Inst. Catalysis, Hokkaido Univ., 5, № 1, 1 (1957).
15. М. И. Темкин, ДАН СССР, 152, 156 (1963).
16. М. И. Темкин, сб. Механизм и кинетика сложных каталитических реакций, «Наука», М., 1970, стр. 57.
17. Е. Христиансен, сб. Катализ. Исследование гетерогенных процессов, ИЛ, М., 1956, стр. 193.
18. Ю. С. Снаговский, Г. М. Островский, Моделирование кинетики гетерогенных каталитических реакций, «Химия», М., 1976.
19. А. И. Кузьмичев, Б. П. Гусев, В. Н. Писаренко, В. В. Кафаров, Изв. АН СССР, сер. хим., 1974, 1393.
20. В. В. Налимов, Н. А. Чернова, Статистические методы планирования экстремальных экспериментов, «Наука», М., 1965.
21. C. Daniel, F. S. Wood, Fitting Equation to Data, Wiley, New York, 1971.
22. O. L. Davies, The Design and Analysis of Industrial Experiments, Oliver and Boyd, New York, 1956.
23. R. L. Plackett, Principles of Regression Analysis, Wiley, Oxford, 1960.
24. H. Дрейпер, Г. Смит, Прикладной регрессионный анализ, «Статистика», М., 1973.

25. Ш. Закс, Теория статистических выводов, «Мир», М., 1975.
26. G. E. P. Box, N. R. Draper, *Biometrika*, 52, 355 (1965).
27. G. E. P. Box, W. G. Hunter, McGregor, J. Erjavec, *Technometrics*, 15, 33 (1973).
28. H. Nagao, *Ann. Stat.*, 1, 400 (1973).
29. O. L. Davies, *Statistical Methods in Research and Production*, N. Y., Hafner, 1960.
30. N. L. Corr, *Ind. Eng. Chem.*, 52, 391 (1960).
31. R. A. Johnson, N. A. Standel, R. Mezaki, *Ind. Eng. Chem. Fund.*, 7, 181 (1968).
32. J. R. Kittrel, J. Erjavec, *Ind. Eng. Chem. Proc. Design Develop.*, 7, № 3, 321 (1966).
33. R. L. Kabel, L. N. Johnson, A. J. Ch. E. J., 8, 621 (1962).
34. F. J. Anscombe, Proc. Berkely Symp. Math. Statist. Probability, 4-th, I, 19, 1963.
35. F. J. Anscombe, J. W. Tukey, *Technometrics*, 5, 141 (1963).
36. G. E. P. Box, *Ann. N. Y. Acad. Sci.*, 86, 792 (1960).
37. R. W. Anderson, T. A. Bancroft, *Statistical Theory in Research*, N. Y., 1952.
38. T. W. Anderson, *Ann. Math. Stat.*, 33, 266 (1962).
39. G. E. P. Box, H. Lucas, *Biometrika*, 46, 77 (1959).
40. N. R. Draper, W. G. Hunter, Там же, 53, 525 (1966).
41. N. R. Draper, W. G. Hunter, Там же, 54, 147 (1967).
42. W. G. Hunter, W. J. Hill, T. L. Henson, *Canad. J. Chem. Eng.*, 47, 76 (1969).
43. A. C. Atkinson, W. G. Hunter, *Technometrics*, 10, 271 (1968).
44. M. J. Box, *The Future of Statistics*, Acad. Press, N. Y., 1968.
45. M. J. Box, *Technometrics*, 12, 569 (1970).
46. G. F. Froment, R. Mezaki, *Chem. Eng. Sci.*, 25, 293 (1970).
47. L. H. Hosten, Там же, 29, 2247 (1974).
48. L. H. Hosten, G. Emig, Там же, 30, 1357 (1975).
49. J. Kiefer, *Biometrika*, 62, 277 (1975).
50. В. В. Кафаров, В. Н. Писаренко, Т. Б. Жукова, А. В. Солохин, сб. II конф. по кинетике катализитических реакций, т. 2, Новосибирск, Изд-во СО АН СССР, 1975, стр. 137.
51. В. Н. Писаренко, В. В. Кафаров, А. В. Солохин, Т. Б. Жукова, Труды МХТИ, Неорганические материалы, процессы и аппараты, Кибернетика химич. процессов, 1975, вып. 85, стр. 172.
52. P. M. Reilly, *Canad. J. Chem. Eng.*, 48, 168 (1970).
53. В. В. Кафаров, В. Н. Писаренко, Т. Б. Жукова, ДАН СССР, 218, 386 (1974).
54. J. Beamer, G. E. Blau, Statistical Technique Applied to Catalytic Reaction Model Identification, AIChE Meeting, San Francisco, 1971.
55. G. E. P. Box, D. R. Cox, *J. Roy. Stat. Soc., Ser. B*, 26, 211 (1964).
56. G. E. P. Box, P. W. Tidwell, *Technometrics*, 4, 531 (1962).
57. E. M. Beale, *J. Roy. Stat. Soc., Ser. B*, 22, 41 (1960).
58. T. Hsiang, P. M. Reilly, *Canad. J. Chem. Eng.*, 49, 865 (1971).
59. L. H. Hosten, Proc. 4-th Internat. Symp. on Chemical Reaction Engineering, Heidelberg, 1976.
60. G. E. P. Box, W. J. Hill, *Technometrics*, 9, 57 (1967).
61. В. В. Налимов, Г. И. Голиков, Логические основания планирования эксперимента, «Металлургия», М., 1976.
62. В. В. Федоров, Теория оптимального эксперимента, «Наука», М., 1972.
63. В. Г. Горский, сб. Вопросы кибернетики, ВИНИТИ, М., 1972, стр. 103.
64. А. И. Кузьмичев, Р. Ф. Мержанова, Л. А. Масцева, Т. Б. Жукова, В. Н. Писаренко, В. В. Кафаров, Н. Ф. Кононов, Зав. лаб., 41, 1235 (1975).
65. В. В. Кафаров, Б. С. Мортков, В. Н. Писаренко, Х. М. Миначев, Сб. Всес. конф. по механизму гетерогенно-катализитических реакций, препринт 48, «Картолитография», М., 1974.
66. H. Chernoff, *Ann. Math. Stat.*, 30, 755 (1959).
67. B. M. Borth, Ph. D. Thesis, University of Waterloo, 1972.
68. D. Meeter, W. Piril, W. Blot, *Technometrics*, 12, 307 (1965).
69. В. В. Налимов, Зав. лаб. 1978, № 3, 325.
70. J. D. Shaw, T. W. Hoffman, P. M. Reilly, C. E. P. Symp. Ser., 40, 118 (1974).
71. R. Mezaki, J. Happel, *Catalysis Rev.*, 3, 241 (1969).
72. L. H. Hosten, G. F. Froment, *Ind. Eng. Chem. Proc. Des. Dev.*, 10, 280 (1971).
73. В. В. Эшновьев, В. И. Тимошенко, Э. Г. Малыгин, М. Г. Слинько, Д. Г. Кнорре, Молекулярная биология, 7, 82 (1973).
74. В. Н. Бибин, Б. И. Попов, сб. всес. конф. по кинетике катализитических реакций, 2, 184 (1975).
75. H. Nagao, *J. Sci. Hiroshima Univ. Ser A-1*, 34, 153 (1970).
76. N. Sugiura, H. Nagao, *Ann. Math. Stat.*, 39, 1686 (1968).
77. G. F. Froment, R. Mezaki, *Chem. Eng. Sci.*, 25, 293 (1970).
78. P. P. Sane, J. M. Woods, R. E. Eckert, *Chem. Eng. Sci.*, 28, 1609 (1973).
79. D. J. Pritchard, D. W. Bacon, *Canad. J. Chem. Eng.*, 52, 103 (1974).
80. R. Bajramovic, P. M. Reilly, Там же, 55, 82 (1977).